# Rational Design of G-Quadruplex Topologies: The impact of Nucleoside Analogs on the Folding of G-rich Oligonucleotides

I n a u g u r a l d i s s e r t a t i o n

zur

Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften (Dr. rer. nat.)

der

Mathematisch-Naturwissenschaftlichen Fakultät

der

Universität Greifswald

vorgelegt von

Beatrice Karg

geboren am 25.11.1989

in Lauchhammer

Greifswald
July 26, 2019

Dekan: Prof. Dr. Werner Weitschies

1. Gutachter: Prof. Dr. Klaus Weisz

2. Gutachter: Prof. Dr. Guido Clever

Tag der Promotion: 24.07.2019

# Contents

## Abbreviations

| | |
|---|---|
| A | adenosine |
| $^{Br}$G | 8-bromo-2'-deoxyguanosine |
| C | cytidine |
| CD | circular dichroism |
| G | guanosine |
| G4 | G-quadruplex |
| GRO | guanosine rich oligonucleotide |
| $^{F}$G | 2'-fluoro-2'-deoxyguanosine |
| LNA | locked nucleic acid |
| $^{LNA}$G | 2′-O,4′-C-methyleneguanosine |
| NMR | nuclear magnetic resonance |
| rG | riboguanosine |
| RNA | ribonucleic acid |
| ROS | reactive oxygen species |
| T | thymidine |
| TBA | thrombin binding aptamer |

# 1 Scope and Outline

Central to this thesis are so-called G-quadruplex (G4) nucleic acids. These unusual structures have recently moved into the scientific limelight - mostly due to their prevalence in the human genome. Incidentally, the vast majority of G4-prone sequences is found in telomeric regions and in the promoter sequences of a large number of cancer-related genes.
Furthermore, recent studies suggest a wide applicability of these structures as therapeutic and functional agents, though the technology is still in its infancy with only a few oligonucleotides in clinical trials. Notably, G-quadruplexes are highly polymorphous, exhibiting different topologies and conformations based on sequence, solution condition and molecularity. Therefore, rational design of such structures with specific, topology-encoded functions demands a comprehensive understanding of the underlying folding parameters.
As the folding process is the result of a whole orchestra of parameters with synergistic effects, the herein proposed approach to understand the G4 structural arrangement concentrates on native G4-forming sequences with well-defined topologies. Perturbations of these structures by rational nucleotide substitutions allow for the observation of discrete effects on the folding pathway and on the resulting overall topology.
The method chosen for primary investigation in the following studies on G4 architectures was Nuclear Magnetic Resonance (NMR) as it is the most powerful tool for structure elucidation in liquids. Unique to this technique, it permits the observation of discrete species in mixtures by distinct perturbations at the atomic level as well as valuable insights into the molecular dynamics.
The listed publications study the effects of site-specific bromine substitutions on native quadruplex scaffolds, thereby successfully inducing new structures. These expand the G4 structural landscape but also enhance our understanding of the driving forces in G4 folding.

**Publication I** **Observation of a Dynamic G-Tetrad Flip in Intramolecular G-Quadruplexes**
Karg, B., Haase, L., Funke, A., Dickerhoff, J. & Weisz, K. *Biochemistry*, 2016, 55, 6949–6955.

In this study, $^{Br}$G substitutions are used as a means to redirect tetrad polarity, thus inverting the native hydrogen bond donor-acceptor pattern while retaining the overall fold. Using NMR

spectroscopy, several important parameters could be extracted, including the energy barrier of a tetrad flip and the exact exchange kinetics between the two forms. As the flip is proposed to happen in native systems but may not observable due to low populations, reported results contribute to an expansion of the quadruplex conformational space and provide for a unique glimpse into the mechanisms of structural rearrangements.

**Publication II** **Loop Length Affects Syn-Anti Conformational Rearrangements in Parallel G-Quadruplexes**
Karg, B. & Weisz, K. *Chemistry - A European Journal* 2018, 24, 10246–10252.

Based on the tetrad polarity flip, this study concentrates on the driving forces of the tetrad flip in parallel G4s. Various substitution patterns and concatenating sequences are probed for their propensity to induce the flip. In conjunction with the high-resolution structure of the flipped species, evasive motions as a response to steric clashes are revealed to be the main driving forces behind a tetrad flip.

**Publication III** **Duplex-Guided Refolding into Novel G-Quadruplex (3+1) Hybrid Conformations**
Karg, B., Mohr, S. & Weisz, K. *Angewandte Chemie*, 2019, DOI: 10.1002/anie.201905372.

Using a rational design approach, a novel G4 topology could be successfully induced by introduction of several fold-defining parameters such as bromine substitutions and Watson-Crick pairing overhangs. The new topology can be achieved with two different stacking patterns which were both characterized in detail. Both structures were shown to coexist with a parallel fold when employing a non-substituted sequence.

**Publication IV** **Manipulating DNA G-Quadruplex Structures Through Guanosine Analogs**
Haase, L., Karg, B. & Weisz, K., *ChemBioChem* 2018, 20, 985-993.

The review article presents an overview on the nucleoside analogs used for a directed folding of quadruplexes, thereby illuminating the driving forces of tetrad flips, strand reorientations and some assorted complete refolds as induced by specific base- and sugar-modified nucleosides.

# 2 Introduction

As the bearer of genetic information, DNA is often viewed as a constant in organisms, but the binary picture of only a passive single and a dormant double strand cannot live up to the dynamic structural range of this versatile biomolecule. Just as its counterpart, RNA, well-known to form complex three-dimensional structures with biological functions and enzymatic activity, DNA can fold into a variety of non-canonical structures - contrary to the picture of the simple four-letter library. Watson-Crick paired nucleosides, adenosine and thymidine as well as guanosine and cytidine stacked upon another in a double-helical arrangement have become common knowledge (Fig. 2.1 **A**). Though, there is another pattern, Hoogsteen pairing, that gives rise to unusual structures such as triplexes and quadruplexes (G4).
In the case of G4, stacks of four guanine bases are arranged in a quartet joined by Hoogsteen-hydrogen bonds (Fig. 2.1 **B**). To compensate for the negative potential in the center of that tetrad stack an additional counterion is usually complexed. Depending on size and charge, though monovalent alkali metals are highly preferred, the counterion is either arranged in the center of one tetrad or between two tetrads (Fig. 2.1 **C**).

## 2.1 G-Quadruplexes *in vivo*

With the help of algorithms about 300.000 probable G4-folding sites have been detected in the human genome.[1] Modern sequencing pushed this figure to about 700.000 with most G4-folding sites occurring in oncogenes and tumor suppressor genes.[2] However, these are all *in*

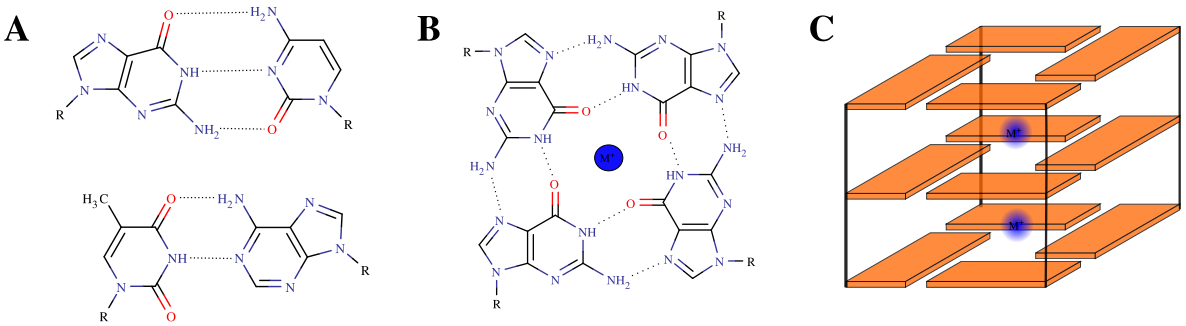



**Fig. 2.1:** Watson-Crick base pairs A-T and C-G (**A**); Hoogsteen-paired guanine-tetrad with central cation (**B**); scheme of a tetramolecular quadruplex with cations complexed between tetrads (**C**).

*vitro* approaches towards the G4 phenomenon. Recent efforts could prove the existence of quadruplexes *in vivo* by both antibody binding[3] and endogenous chromatin sequencing.[4]
The most apparent G4 sites are the telomeric ends of the chromosomes. Here bases are found in unpaired single-stranded form which facilitates G4-formation *in vivo*.[3] In fact, the human telomeres consist of the well-studied TTAGGG-repeat prone to fold into G4s and directly related to telomerase activity and thus cell immortality.[5,6] Efforts to counteract the unrestrained proliferation of cancer cells are concentrated on these G4 sites in telomeric regions, employing low molecular weight ligands to stabilize G4s.[7] Thus, the enzyme telomerase, responsible for maintaining telomere length and coincidentally overexpressed in most cancer cells, is effectively blocked in the presence of G4 structures which induces senescence and eventually death in cancer cells.[8]
G4 sites in the double-stranded regions have been found abundantly in origins of replication[9] but formation has best been studied as a response to the partial melting of the DNA during transcription. As the RNA-polymerase progresses along the strand and segments are unwound, negative supercoiling is induced behind the transcription bubble.[10] Upstream elements, like the NHE $III_1$ element of the cancer-associated MYC transcription factor[11] are strongly influenced by this negative superhelicity.[12] Specifically, this region shows both enhanced G4 and i-motif formation on the corresponding strand, though in cells these two structures are observed to be mutually exclusive and part of a complex regulatory molecular switch (Fig. 2.2).[13]
Other prominent regulatory G4 sites have been found in the immunoglobulin switch region,[14] the promoter region of the proliferation-related KRAS switch[15] and in the anti-apoptotic BCL-2 family.[16] Furthermore, quadruplexes are strongly involved in epigenetic processes such as methylation.[17] The G4-structure is found to be a primary target and also inhibitor of the DNA methyltransferase. Consequently, G4s are potentially protecting off-site elements from methylation.
Yet another suggested regulatory task of G4s, is related to reactive oxygen species (ROS) as these produce DNA-modifications such as 8-oxoguanine which can directly interact in G4-formation. Following a complex signaling pathway including a base excision repair mechanism in G4s, a recent study could prove the introduction of 8-oxoguanine in G-rich sites to be an important gene activator.[18]

## 2.2 G-Quadruplexes for technological applications

In spite of all possible biological targets for ligand-induced de-/stabilization, G4s have moved beyond their role in biology. Due to their unique structural features, they can be found as knots in nanostructures[19] and as components in long conducting nanowires.[20] Their ability to change conformation upon a variety of signals has made them very attractive as parts of

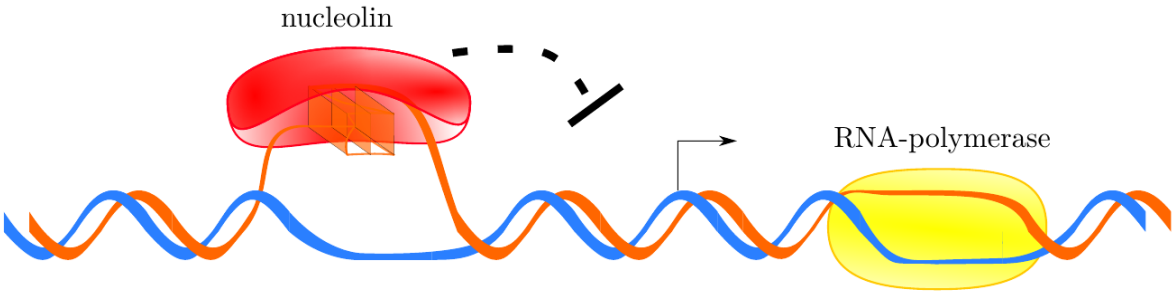



**Fig. 2.2:** Scheme of G4-related MYC-transcription: nucleolin binds and stabilizes G4 thus down-regulating MYC expression.

nanomachines.[21–23] These aspects are also used in ligand-driven switches[24] where G4s act in perfect sensing mechanisms[25,26] mostly in combination with light-up probes detecting G4-formation or even specific G4-structures in easy-to-use assays. Systematic Evolution of Ligands by Exponential Enrichment (SELEX) has produced a vast database of quadruplex-based sensing of small molecules like thyroxin[27] but also proteins[28] and straightforward oligonucleotide recognition.[29,30]

Apart from that G4s can be found as catalysts, either as enantioselective templates for chemical reactions[30] or as peroxidase mimicking DNAzymes.[31–33] Recently, a G4-DNAzyme was found that selectively repairs UV-damaged DNA, specifically cyclobutane Thymine dimers, in the same orders of magnitude as some lyase enzymes.[34]

The quadruplex structure is also often the scaffold in selected antiproliferative nucleic acids, prominently in so-called G-rich oligonucleotides (GRO).[35,36] In general, their activity is linked to competitive nucleolin occupation, a protein that specifically targets G4s.[9] It is also associated with neurodegenerative cell cascades[37] and is speculated to represent a target for viruses.[38] The most advanced quadruplex-based aptamer in the field is the long G4 AS1411, which reached phase II clinical trials[39] as a TNF$\alpha$[40] and NF$-\kappa$B inhibitor.[41] Likewise, G4s have been found in antiviral aptamers against HIV integration[42,43] and transcription[44] as well as SARS helicase[45] and hepatitis C polymerase.[46]

The workhorse of G4-aptamer design and development is the Thrombin Binding Aptamer (TBA) revealed by SELEX as a highly potent anticoagulant.[47–49] The small 15-nt oligonucleotide binds at concentrations as low as 10 nM, thereby inhibiting thrombin-mediated cleavage of fibrinogen to fibrin and eventually clot formation.[50] Unfortunately, due to high dosages needed in humans, phase II clinical trials have been terminated.[51] Nevertheless, on the basis of crystallographic and NMR data of the thrombin-G4 complex,[52,53] there are recent efforts to improve efficacy by insertion of site-specific modifications that both prevent cellular degradation of the structure and enhance binding propensity to the target.[54,55]

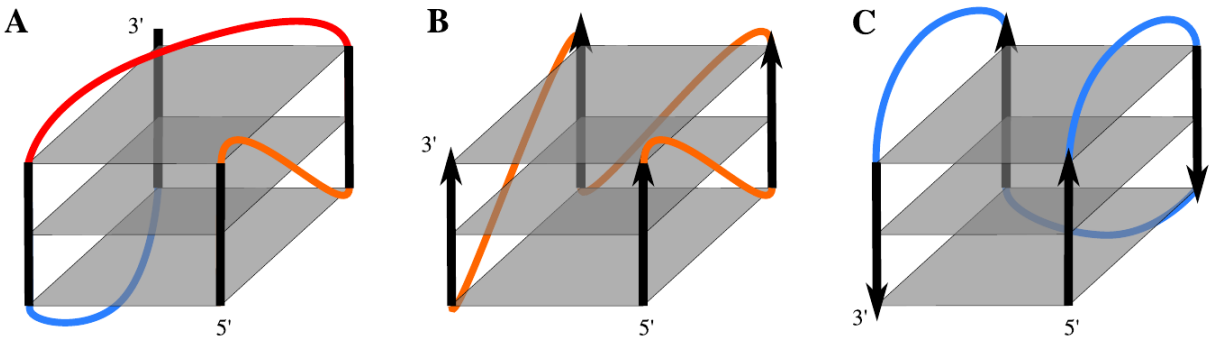



**Fig. 2.3:** **A**: Scheme of a hybrid structure with most common loops in monomolecular G4s from 5' to 3': propeller (orange), diagonal (red), lateral (blue) (**A**). Scheme of a parallel G4 with propeller loops (**B**). Scheme of an antiparallel G4 with lateral loops (**C**).

## 2.3 Understanding G-Quadruplex Structures

Rational design and strategic enhancement of certain features of functional G4s demands a comprehensive understanding of the governing independent folding principles that create a complex network of interacting forces.

First, it needs to be mentioned that even though quadruplexes can be constructed of two,[56] three[57] or four oligonucleotides,[58] the monomolecular variant is most often encountered in biology and technological application. If all strands are connected, the bridging regions, called loops, become a structural determinant (Fig. 2.3 **A**). Short intervening sequences tend to form as propeller loops which traverse G4 grooves from 5' to 3' and thereby connect parallel strands (Fig. 2.3 **B**), hence their synonym 'double chain reversal loops'.[59–61] Antiparallel strands are connected by either lateral (edgewise) or diagonal loops (Fig. 2.3 **C,A**). Thus, three folding patterns are possible all-parallel, antiparallel and hybrid structures, usually described as (3+1) for the three parallel tracts accompanied by one tract in antiparallel orientation (Fig. 2.3 **A**). However, G-runs can also be interrupted by bulges,[62] connected with zero-nucleotide loops[63] or be not even restricted to right-handedness.[64] With such a variety, straightforward prediction from sequence to structure is not yet reliably possible. The rather simple fixation on G-tracts and loops also neglects a wide range of additional stabilizing factors such as alternative base stacks of A,[65] C in tetrads[66,67] or hexads[68] or water-mediated mismatches.[69] Even long double-helices in loop sequences are possible.[70,71] This variety of impacting elements comes as a hindrance when trying to find possible G4-sites in a genome thereby constantly demanding ever more complex algorithms.

For a consistent prediction of structures from known sequences, the described external driving forces need to be combined and the groove width has to be considered, an additional and often-neglected factor in simple double-helical DNA. In an antiparallel arrangement of G-tracts, linked guanosine bases have to accomodate alternating *syn*- and *anti*-glycosidic torsion angles

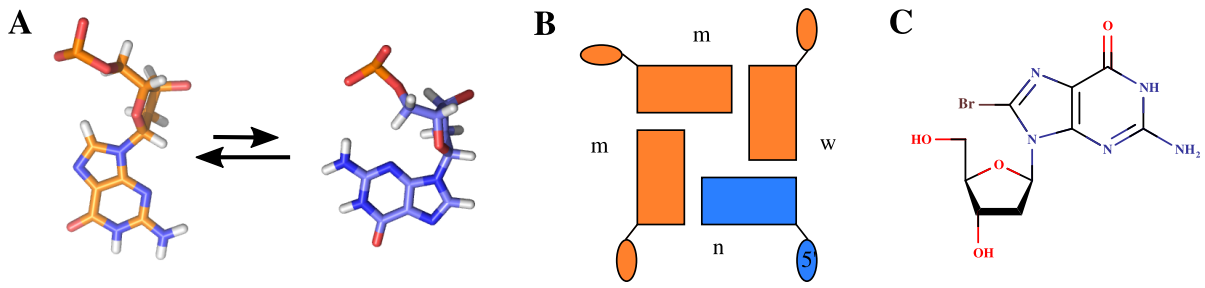



**Fig. 2.4:** Model of *anti*- and *syn*-conformation of guanosine (orange and blue, respectively) (**A**). Groove width in a monomolecular G4 depends on guanosine conformation (w-wide, m-medium, n-narrow) (**B**). Modified nucleoside: 8-bromo-2'-deoxyguanosine ($^{Br}$G, **B**).

to retain the hydrogen bonding network (Fig. 2.4 **A**).[72] This will lead to diverging groove sizes for bases with the same torsion angle (medium) and between bases with different torsion angles (narrow and wide, Fig. 2.4 **B**). Groove width in turn influences the requirements for loop sequences and torsion angles are relevant to the overall stacking pattern. For example, a *syn-syn* step of 5'-G to 3'-G is energetically inferior to an *anti-anti* or *syn-anti* step.[73]

A perfect case study of such variability is given by the human telomeric sequence, with the specific sequence $AG_3(TTAG_3)_3$. In $Na^+$ solution the sequence adopts a well-defined antiparallel structure, whereas potassium as a central cation will lead to a variety of coexisting structures.[74,75] Variations in the terminal overhang sequences and different loop nucleotides will favor specific folding patterns.[76–78] Moreover, external conditions such as crowding agents can influence the overall topology.[79]

In fact, endeavors to determine the possible three-dimensional structures of G-rich genomic sequences rarely yield a single well-defined quadruplex. This might be a result of the very different conditions *in vitro* and *in vivo* which are known to influence the overall fold.[80] It might also be the lack of the genomic context in terms of overhang sequences or the neglect of the many other factors previously described. Though, most intuitively, G4 variability can be better assessed by following the G4 folding pathway. Quadruplex folding is not straightforward, but happens on an energy landscape of well-separated competing minima each related to structurally different conformations. While basins are transiently populated until a final, thermodynamically stable conformer is achieved, the basins can function as kinetic traps prolonging the folding process and guiding its directions.[81]

## 2.4 Base modifications

A way to rationally affect the energetic barriers between different minima are base substitutions. These are commonly used to enforce a desired folding pattern and thus to isolate single species from mixtures. Among the frequently used substitutes are 8-bromoguanosine ($^{Br}$G, Fig. 2.4 **C**), 8-methylguanosine, locked residues (LNA) and 2'-fluoro-guanosine ($^{F}$G). Specifically, $^{Br}$G has long been known to adopt *syn*-conformation and was first observed in the tightly packed crystal structures of the hydrated nucleoside without tetrad-stabilizing counterions.[82] Based on its advantageous electron acceptor potential, $^{Br}$G has previously been used to show an efficient electron transport in G4-scaffolds upon light excitation. *In vivo*, the bromination of G at position 8 is known to be induced by excess amounts of hypobromous acid, a potent oxidant released under inflammatory conditions by neutrophils.[83] As inflammation is known to be a major cancer predisposition, $^{Br}$G has been detected in human liver and urine of cancer patients[84] and hence is proposed to function as a diagnostic indicator for early stage detection of carcinoma.[85]
In regard to G4 studies, $^{Br}$G has extensively been used to induce *syn*-conformation in G-residues, therefore not only stabilizing specific structures in mixtures[86–88] and allowing the identification of major species of native G4-forming sequences, but also expanding the G4 conformational space by inducing unprecedented structures.[89,90]
For a general understanding of the quadruplex folding landscape, such novel structures are essential as they represent possible transient states or minor species that are not observable in the experimental timescale or with the available sensitivity of biophysical methods.

# 3 Observation of a Dynamic G-Tetrad Flip in Intramolecular G-Quadruplexes

In this study, $^{Br}$G analogs have been used to probe the *syn*/*anti* propensity of G4 residues in an overall parallel folding topology. Parallel quadruplexes are usually composed of all-*anti* guanosines and have been known to be the most stable of all possible G4-folds. In opposition, $^{Br}$G preferentially adopts *syn*-conformation and the introduction of $^{Br}$G into an all-*anti* structure should perturb the overall native folding pattern significantly. Compared to an all-*anti* tetrad as observed in the parallel structures, an all-*syn* tetrad would invert the hydrogen bonding polarity. In context with G4, the changed polarity would result in a different stacking pattern introducing a shift from the native homopolar to heteropolar tetrad stacking.
Two modified c-MYC sequences were studied with either three or all four 5'-terminal tetrad G-residues of the parallel c-MYC ($TGAG_3TG_3TAG_3TG_3TAA$) substituted with $^{Br}$G (MYC-3 and MYC-4, respectively). While MYC-4 exclusively adopts a new fold with an inverted all-*syn* tetrad, MYC-3 exhibits a mixture of two species according to 1D NMR data. Thorough NMR

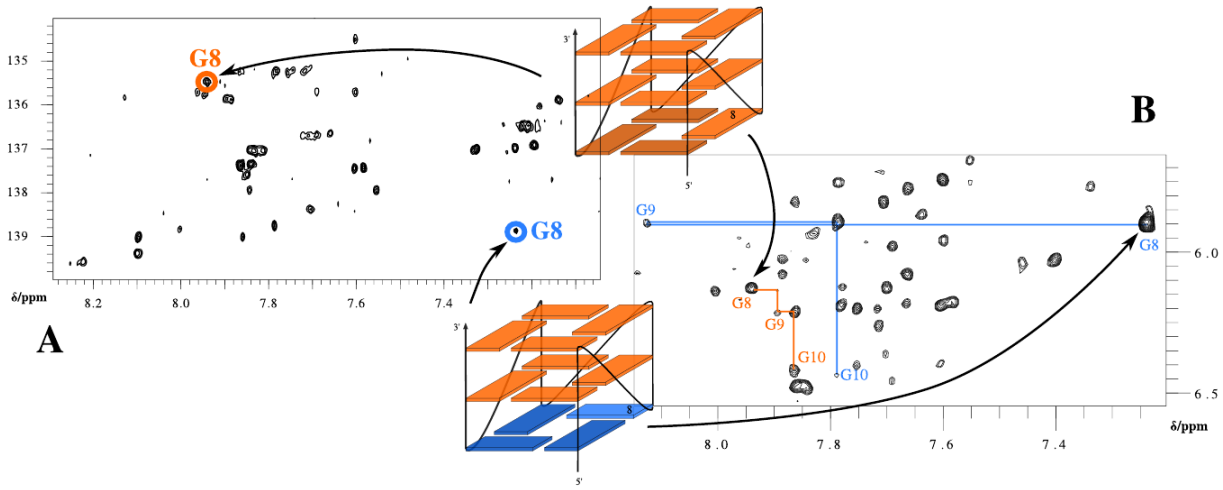



**Fig. 3.1:** Scheme of MYC-3s and MYC-3a linked to their respective H8-C8 (**A**) and H8-H1' (**B**) correlations of the G8 residue in HSQC and 2D NOE spectra of the MYC-3 mixture; the strong intranucleotide crosspeak for H1'-H8 is due to their close proximity in a *syn*-conformation. Nucleosides are colored corresponding to their torsion angles, either *anti* or *syn* (orange and blue, respectively). Darker shades indicate substitution sites with $^{Br}$G.

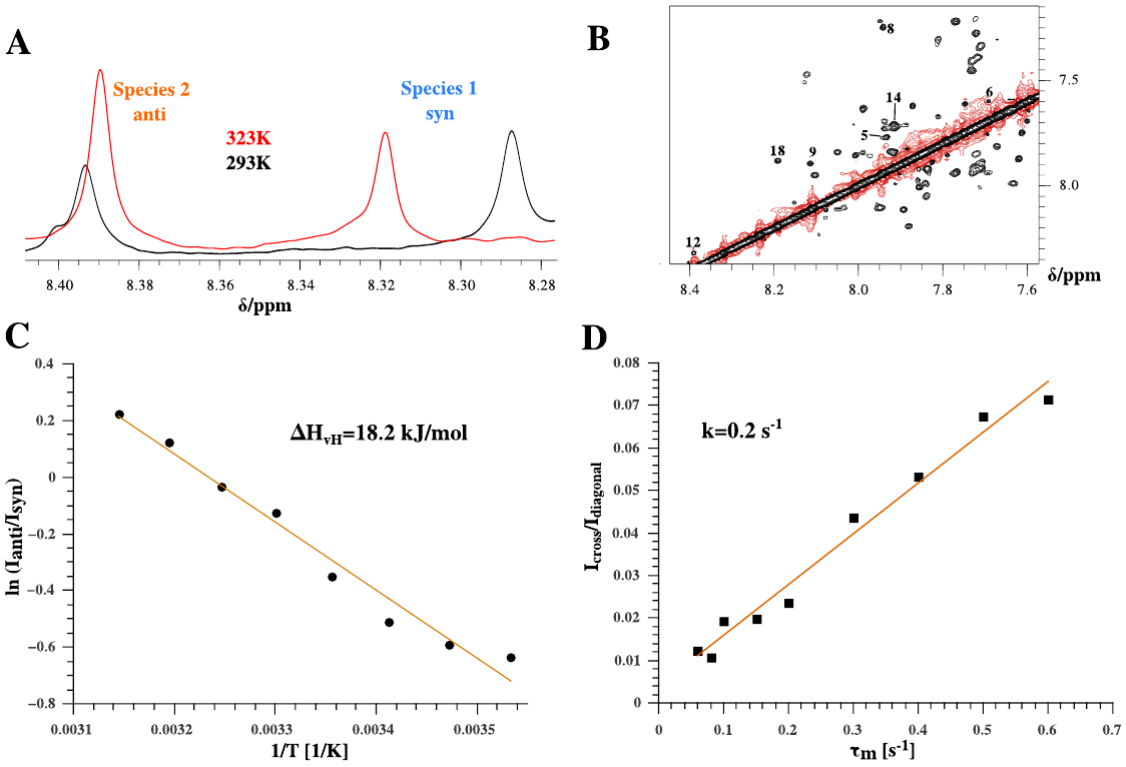



**Fig. 3.2:** 1D-NMR spectra of MYC-3 show temperature dependent populations (**A**). EASY-ROESY crosspeaks of aromatic protons in MYC prove dynamic exchange of MYC-3s and MYC-3a (**B**). Van't Hoff analysis of MYC-3 yields enthalpy change of the tetrad flip (**C**). EASY-ROESY spectra acquired with different mixing times allow calculation of the exchange rate between the two species of MYC-3 (**D**).

studies demonstrated these to be a quadruplex with an all-*anti* conformation, MYC-3a, similar to the native sequence, and another parallel structure with an all-*syn* 5'-terminal tetrad as in MYC-4. Specifically, these two structures can be discerned by different chemical shift patterns displayed by G8, the non-brominated residue of the 5'-terminal tetrad. A strong upfield shift of protons can be attributed to a *syn*-conformer, whereas strong upfield shifted carbon resonances are indicative of *anti*-guanosines. (Fig. 3.1)

A striking feature of the MYC-3 mixture is its dynamic interconversion. The dynamic exchange of both species could be observed on the NMR timescale and, by varying the mixing times in ROESY experiments, peak intensities could be used to calculate a $k_{ex} = 0.2 - 0.3$ s$^{-1}$ (Fig. 3.2 **B,D**). Given this relatively fast exchange rate, the observed tetrad flip was estimated to be the result of a single tetrad dissociation.

This assumption was strengthened by thermodynamic studies. Since varying temperatures affected population ratios and hence equilibrium constants, the mixture could be subjected to a van't Hoff analysis. Using a 40 °C temperature range an enthalpy change for the flip could be calculated as $\Delta H^\circ = 18.2$ kJ/mol (Fig. 3.2 **A,C**). This rather low energy barrier is not

compatible with a complete melting event as a prerequisite for the tetrad inversion. Rather, thermal breathing motions of the terminal tetrad are likely to allow for the conformational change. In that case, the tetrad flip might occur as well in a non-modified sequence but is probably not observable due to low populations of flipped species.
Another important observation needs to be mentioned with regards to thermal stability and the energetics of base stacking. Both MYC-4 and MYC-3 sequences show the same melting transition at around 70 °C. In the case of MYC-3, this denotes a single melting transition for both the MYC-3s and MYC-3a fold. In regard to base stacking, the folds differ in having exclusively homopolar or also heteropolar interactions. This is in agreement with previous calculations, as a *syn-anti* step in MYC-3s and an *anti-anti* step in MYC-3a are expected be about equal in energetic terms as shown by the similar thermal stability of both folds.

# 4 Loop Length Affects Syn-Anti Conformational Rearrangements in Parallel G-Quadruplexes

Based on the previously observed tetrad flip in MYC-derived parallel quadruplexes the impact of $^{Br}$G in nonmatching *anti*-positions has been studied in detail by systematic substitutions. Various patterns have been chosen in both the 3'- and 5'-terminal tetrad to further study the site-dependent *syn*/*anti* susceptibility of C8-brominated G-residues.
Several substitutions in the 3'-terminal tetrad yielded mixtures of species with significantly decreased thermal stabilities as implied by NMR and CD spectroscopic measurements. Therefore, it can be assumed that upon the formation of an all-*syn* 3'-tetrad grave energetic penalties are evoked, in line with previous observations for tetrameric G4s.

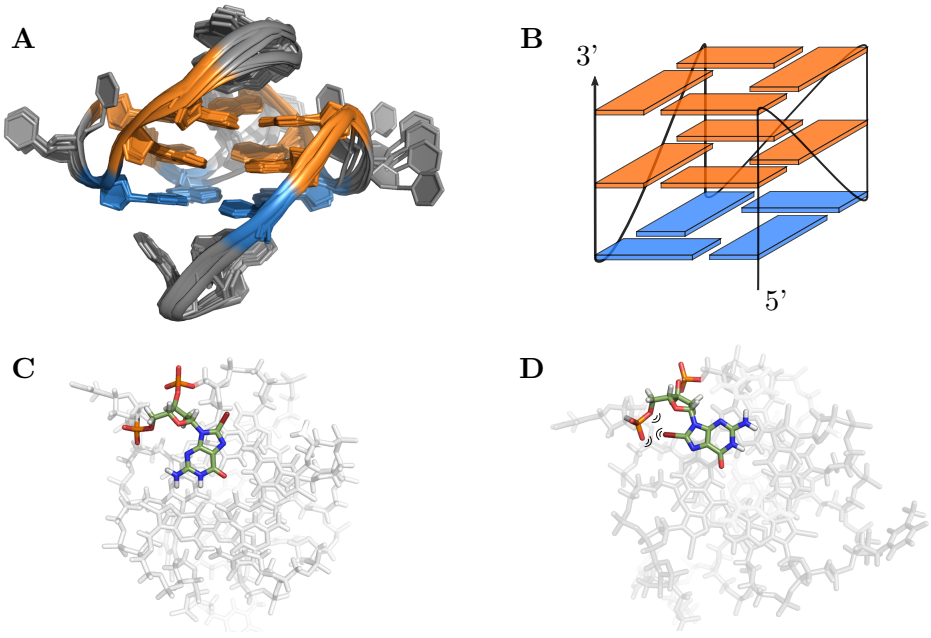



**Fig. 4.1:** High-resolution structure of MYC-4. Ten lowest-energy structures are superimposed, *syn*- and *anti*-residues are presented in blue and orange, respectively (**A**). Scheme of MYC-4, residues colored in blue are both $^{Br}$G-substituted and in *syn*-conformation (**B**). Bromine substitution at position 17 in *syn*- (**C**) and in *anti*-conformation (**D**) with indication of potential steric clashes.

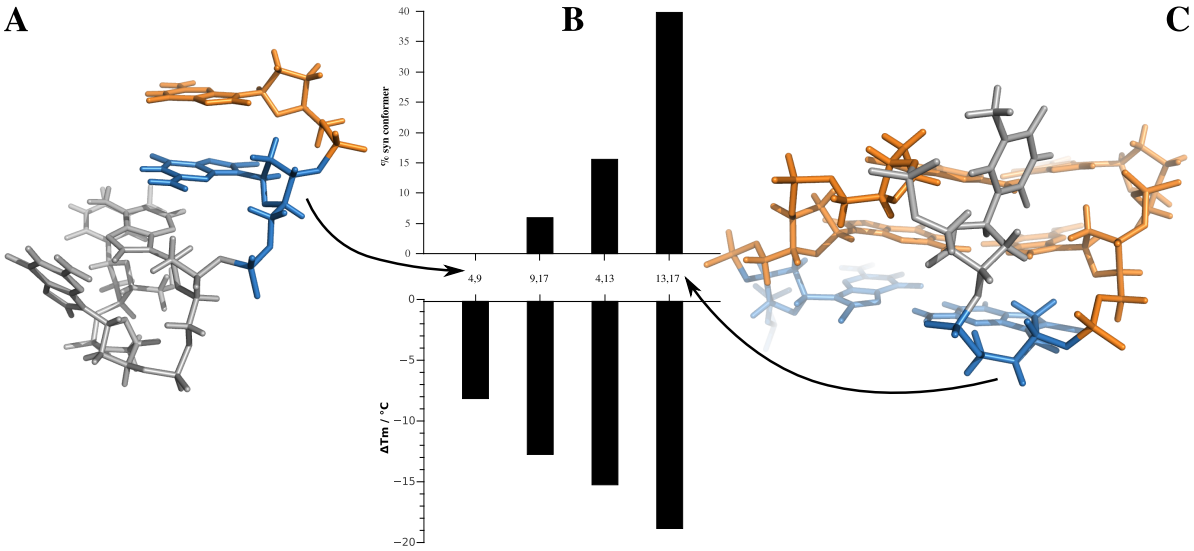



**Fig. 4.2:** Bromine substitutions at 5'-terminus (**A**) and after a short 1-nt propeller loop (**C**) result in different syn populations and thermal stabilities (**B**) in 2-$^{Br}$G substituted MYC-derivatives.

Consequently, the focus was shifted onto distinct substitution patterns with three $^{Br}$G sites in the 5'-terminal tetrad. NMR experiments could confirm a mixture of either all-*anti* or all-*syn* 5'-tetrads for all differently modified sequences. Ratios of populations could be calculated by straightforward integration of the well-resolved adenine H2 resonances, yielding a relative efficiency of the tetrad flip for each substituted residue. Strikingly, flip inducing sites are in direct neighborhood of 1-nt propeller loops indicating steric effects to drive the flip. Using only sequences with 2-site substitutions this correlation could be validated: $^{Br}$G-substitutions connected to short loops yielded 40 % of the 5'-terminal *syn*-conformer whereas the antithetic sequence remained exclusively in *anti*-conformation.

Structure elucidation of the 100 % flipped MYC-4 with fourfold substitution sheds light on this steric contribution (Fig. 4.1 **A,B**). Whereas the calculated structure with all $^{Br}$G residues in *syn*-conformation shows no steric hindrances at all, a modeled structure with *anti*-conformers exhibits significant clashes between the bromine and 5'-phosphate oxygen atoms (Fig. 4.1 **C,D**, respectively) that probably serve as an additional driving force for *anti*→*syn* transitions.

The loop-length related flexibility of single residues can also be observed in imino proton exchange rates with water which are accessible via NMR using ROESY exchange spectroscopy. These should directly correlate with the first step of the proposed tetrad flip, the opening of base pairs, as imino protons loose hydrogen bonding and readily exchange with water. Indeed, the exchange rate of the residue neighboring the flexible 2-nt propeller loop is greatly increased, whereas the 5'-terminal G following past a 1-nt loop is significantly reduced.

For proof of principle, a new parallel sequence was designed with shuffled loop lengths and comparable substitution patterns, shifting from a 1-2-1 to a 2-1-1 nucleotide loop sequence. The

anticipated degree of tetrad inversion can be observed by following adenine resonance integrals (Fig. 4.2 **B**, upper), indicative of the loop length as a major determinant in flip efficiency. Astonishingly, this effect is also observable in thermal stability, as substitutions with higher flip tendency exhibit decreased melting temperatures (Fig. 4.2 **B**, lower). Thus, these differences directly reflect relative energies of position-dependent steric interactions of $^{Br}$G *anti* conformers. In conclusion, a potential guideline for bromine substitutions can be formulated as follows: for highest *anti*→*syn* transition efficiency, bromination sites should best be located within a more restricted residue environment whereas $^{Br}$G residues in *anti*-conformation are better tolerated if the sterically demanding bromine substituent is located within a more flexible environment.

# 5 Duplex-Guided Refolding into Novel G-Quadruplex (3+1) Hybrid Conformations

Of all possible G4-structures, only a small fraction has been observed *in vitro* up until now. A prominent example of the elusive structures is a hybrid fold with the specific loop pattern: 5'-lateral-propeller-propeller-3' (lpp). In an effort to expand the quadruplex conformational space, the parallel MYC structure has been employed as a parent sequence for the rational design of such a novel topology.

For a hybrid structure to form, one G-tract has to change directionality compared to the native all-parallel structure. This antiparallel strand would be constituted of all-*syn* guanosine residues if the hydrogen bonding pattern is retained. Furthermore, as longer intervening sequences

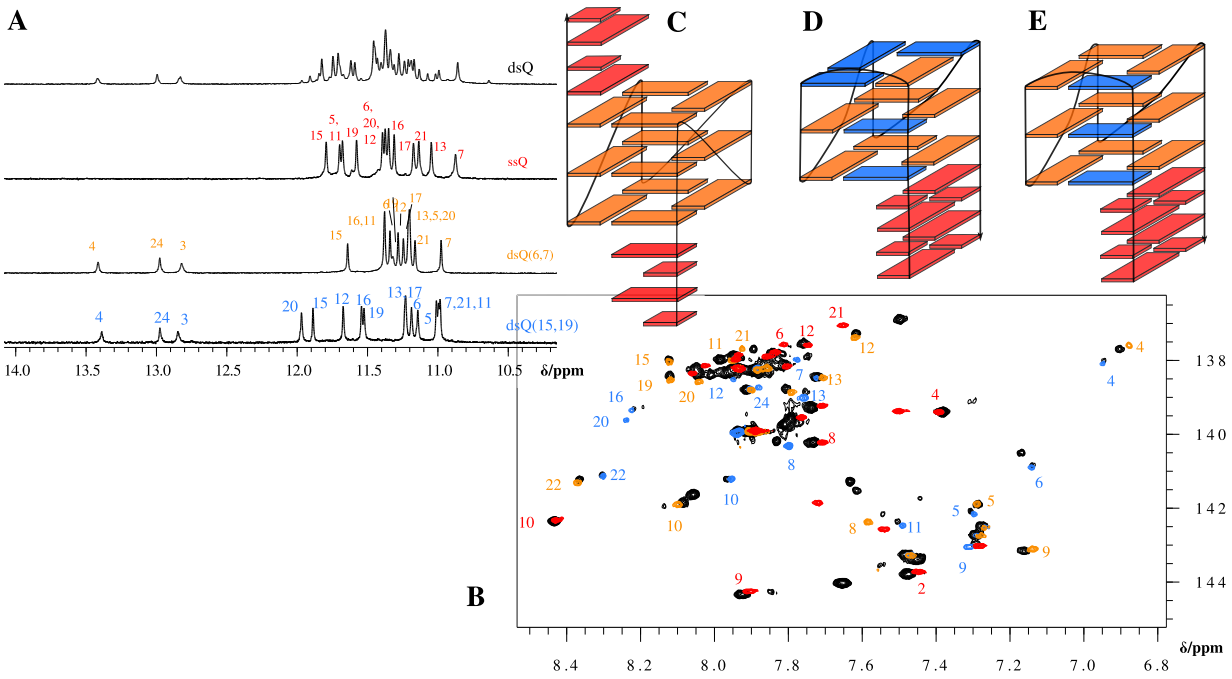



**Fig. 5.1:** Imino proton spectra of dsQ and modified sequences ssQ, dsQ(6,7) and dsQ(15,19) (**A**). HSQC overlay of the dsQ mixture (black) with all modification-enforced comformers ssQ (red), dsQ(6,7) (orange) and dsQ(15,19) (blue) (**B**). Schematic representations of the parallel structure ssQ with non-matching single stranded termini (**C**) and the bromine-substituted dsQ(15,19) (**D**) and dsQ(6,7) (**E**) both with paired overhangs; *syn* residues are colored in blue.

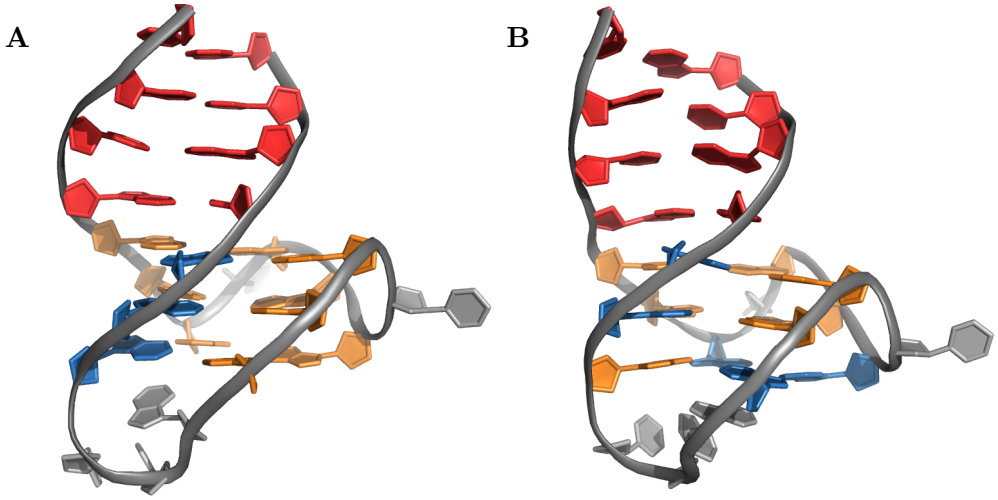



**Fig. 5.2:** Three-dimensional structure of dsQ(6,7) with 3 *syn*-guanosines (**A**) and dsQ(15,19)-like fold with 5 *syn*-guanosines (**B**).

are preferred for lateral and diagonal loops a 3-1-1 loop nucleotide pattern for the desired lateral-propeller-propeller loop sequence was introduced. Also, additional duplex-pairing termini were added to enforce a single well-resolved species dsQ(6,7) with two $^{Br}$G-substitution sites for *syn*-G stabilization.

Detailed NMR-investigations revealed a hybrid structure with the first strand inverted compared to the native MYC-structure, most prominently indicated by Watson-Crick base pairs in the overhang regions (Fig. 5.1 **A**). The desired lpp pattern and resulting all-*syn* G-tract could be shown through sequential NOE contacts of *syn*-guanosine imino to sugar protons as well as contacts from loops into the G-tracts. Due to its relationship to the known hybrid-1 structure, this fold was termed hybrid-1R.

The attempt to retain the conformation without bromine substitutions resulted in a mixture of several species (dsQ). One of those species could be identified as the previously discussed hybrid-1R fold with a *syn-syn-syn* tract. Another species was shown to be a parallel fold with unpaired single-stranded termini similar to the native MYC, as indicated by a comparison with a sequence with 4-nt non-matching 5'- and 3'-overhangs but with preserved 3-1-1 loop pattern (Fig. 5.1 **B**).

A third species was found by varying bromine substitution patterns with $^{Br}$G residues translocated from the first strand to the 5'-terminal tetrad (dsQ15,19). This stabilizes a hybrid structure with lpp loop sequence but with only two *syn*-residues in the first G-tract, further denoted as hybrid-1R'. For the last guanosine of this tract to retain *anti*-conformation, the three remaining tetrad-Gs have to attain *syn*-conformation. The resulting structure could be proven by detailed NMR structural analyses, again using C8 chemical shifts to detect *syn*-conformers and with

NOE-connectivities for full resonance assignments. Comparison of HSQC-spectra allowed the identification of the final minor species of dsQ as being the corresponding hybrid-1R' fold.
In terms of base stacking, the hybrid-1R' fold is expected to be more stable than hybrid-1R as it contains a smaller number of disfavored *syn-syn* stacking interactions than the latter. However, with a relative population of only 12 % hybrid-1R' in the mixture of dsQ, it can be assumed that a smaller number of *syn*-G residues in the hybrid-1R quadruplex outweighs the larger number of energetically unfavorable *syn-syn* steps.
It should be noted that the parallel fold is the most populated in the dsQ mixture. Thus, the additional four base pairs do not provide sufficient energetic compensation to direct the folding to a new global minimum. However, the additional formation of only four base pairs allows for a disfavored fold to effectively compete with the preferred parallel topology.
Based on these results, a possible folding pathway can be postulated with a first step consisting of a fast transient base pairing of the overhang sequences and subsequent preorganization of the final G4. Thus, the structure can be kinetically trapped and the two hybrid-1R and hybrid-1R' structures (Fig. 5.2) correlate to two basins with strand-flipped species on the folding landscape but with the parallel structure still being the thermodymically most stable form.

# 6 Manipulating DNA G-Quadruplex Structures by Using Guanosine Analogs

For a better understanding of quadruplex folding, a great many of influencing factors have to be taken into account. A special tool emerged recently in the form of modified nucleosides serving as a magnifying glass for single residues. This review provides as an overview of induced changes in folding upon exclusive incorporation of nucleoside analogs.

As strand directionality and glycosidic torsion angles are intimately connected, they can easily be used to influence one another. Bulky base substituents, as described before, usually drive a *syn*-glycosidic conformation (Fig. 6.1 **A**). In contrast, sugar modifications like 2'-fluorine- ($^{F}$G), ribo-sugars (rG) or locked guanosines ($^{LNA}$G) induce *anti*-conformation. This effect is generally related to their preference or, in case of $^{LNA}$G, compulsion to adopt a North puckering, a conformation of the furanose ring with C3' above the plane and oriented towards C5' - as opposed to the C2' above-plane South pucker favored by unmodified 2'-deoxynucleosides (Fig.

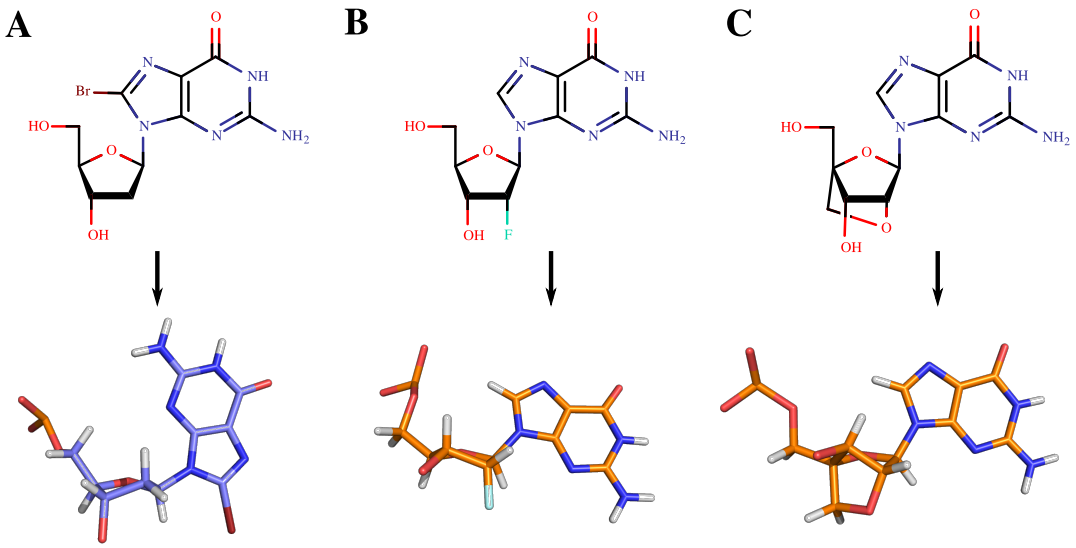



**Fig. 6.1:** Nucleoside analogs and their respective 3D representations: 8-bromo-2'-deoxyguanosine ($^{Br}$G) with a 3D-representation in *syn*-conformation and with South pucker (**A**); 2'-fluoro-2'-deoxyguanosine ($^{F}$G) in *anti*-conformation and with North pucker (**B**); 2′-O,4′-C-methyleneguanosine ($^{LNA}$G) in *anti*-conformation and with North pucker (**C**).

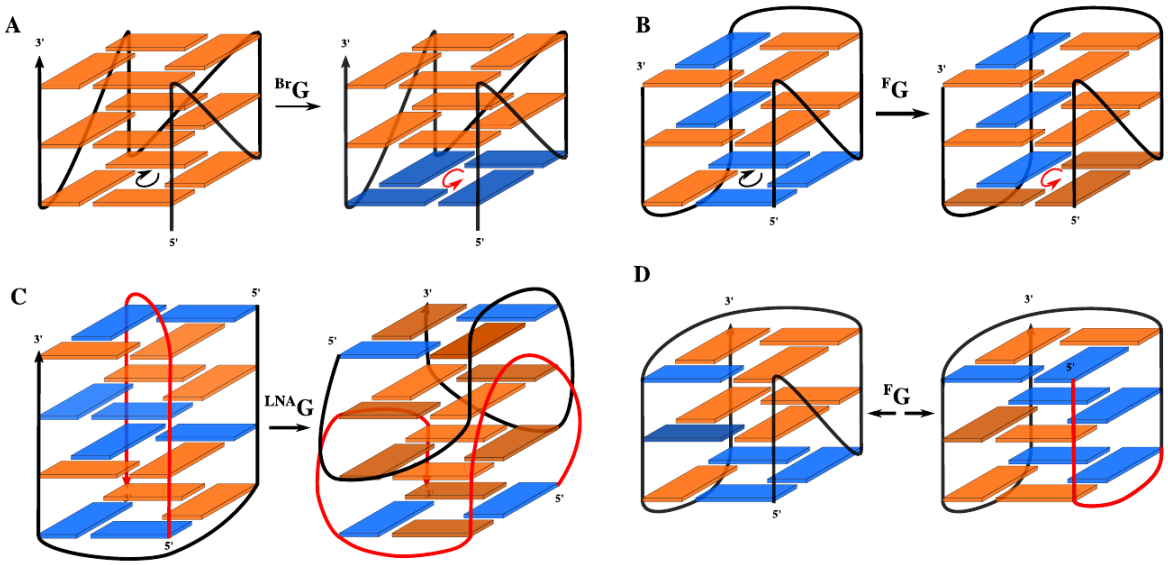



**Fig. 6.2:** Overview of modified structures with *syn*-(blue)/*anti*-(orange) conformational rearrangements induced by modified residues (darker shade). Introduction of $^{Br}$G in 5'-terminal tetrad of parallel MYC induces a tetrad flip (**A**). Tetrad flip by substitution of $^{F}$G in non-matching *syn*-positions of hybrid-type human telomeric sequence (**B**). Insertion of $^{LNA}$G residues into an antiparallel dimer results in an unusual V4-fold (**C**). Modified $^{F}$G at non-matching *syn*-position of the central tetrad leads to two coexisting species including an inverted 5'-terminal G-tract (**D**).

6.1 **B,C**). Upon adoption of a North-pucker the *anti*-conformation in turn prevents clashes between N3 and H3' and O5' shifting the glycosidic torsion angle equilibrium.

Modifications like these have been used to induce tetrad inversions either from all-*anti* to all-*syn* tetrads in parallel tetra- and monomolecular systems or in hybrid structures by substituting non-matching *syn*-positions with sugar-modified residues (Fig. 6.2 **A**,**B**). Apparently, these substitutions tend to also influence unmodified residues of the same tetrad in order to retain hydrogen bonding. Parameters such as preceding loop lengths still have to be taken into account when determining the inversion efficiency of each substitution site to achieve a full tetrad flip without coexisting species.

Compared to tetrad inversions, complete refolds are often difficult to predict. For example, induction of 5'-*syn-anti-syn-anti*-3' stabilizing modifications within a tetramolecular and exclusively parallel sequence, resulted in a new antiparallel structure in coexistence with several other species of similar stability. Introduction of pucker-locked residues ($^{LNA}$G) in the bimolecular Oxytricha nova telomeric sequence ($G_4T_4G_4$) into matching *anti*-positions induced a completely new fold probably driven by unfavorable narrow groove interactions of the $^{LNA}$G residues(Fig. 6.2 **C**). Another example included a hybrid structure with $^{F}$G substitutions at a single non-matching position in the central tetrad. This specific modification resulted in a mixture of two species, the first with unaltered topology and $^{F}$G in the unfavorable *syn*-conformation and the second

with a strand polarity inversion, thus allowing for the preferred $^{F}$G *anti*-conformer (Fig. 6.2 **D**). Substitutions can also induce additional favorable interactions such as unconventional hydrogen bonds. These have been proven to occur intranucleotidic between 2'-fluorine-substituents in South-puckered ribose sugars and guanine H8, as well as in a sequential, internucleotidic form from 2'-F to H8 of the 3'-neighboring guanine base.
Sugar modified residues have also been observed to retain unexpected sugar puckers to avoid unfavorable interactions in narrow grooves. This has been observed in hybrid structures with a 2'-fluorine or 2'-hydroxy substituent oriented away from the narrow groove and into the medium groove by a suitable sugar conformation.
Taken together, all these substitution-induced structures and their respective driving forces add to our understanding of general folding principles in quadruplex nucleic acids. These can be used to better predict G4 folding patterns and may ultimately enable a rational design of G4 scaffolds for specific structure-related tasks.

# 7 Conclusion

This thesis serves as a glimpse into the vast structural landscape of quadruplexes. Utilizing site-specific modified nucleoside substitution with $^{Br}$G, a variety of structures could be induced yielding novel folding patterns. Tese, in comparison with their native counterparts, hold valuable information about the driving forces of G4 folding.

While initial studies (publication I) demonstrated that $^{Br}$G is feasible to achieve glycosidic torsion angle flips in parallel quadruplexes with a resulting inversion of the hydrogen bonding pattern, following investigations (publication II) could show that the main driving force of such $^{Br}$G-driven tetrad polarity inversions is based on steric clashes involving the unfavorable nucleoside conformer and directly related to the overall G4 structure by loop flexibility. Taken together, a pathway for the observed tetrad flip can be formulated as follows: First, thermal breathing motions initialize tetrad opening followed by a concerted conformational transition and the reformation of the hydrogen bonding pattern. During the transition the *syn-anti* equilibrium can be fine-tuned with the herein described tool, i.e. the $^{Br}$G analog.

Subsequent studies (publication III) build on former results and reveal how the $^{Br}$G-tool can be applied. Supported by a new strategy, the duplex-guided strand directionality, the studies demonstrate the rational design of a novel G4-topology. Incidentally, two different unprecedented variations of this novel topology have been found and characterized, elucidating the energetic contributions of favorable glycosidic conformers and stacking interactions. The coexistence of both, hybrid-1R and hybrid-1R' structures, together with the energetically more favorable parallel form lacking the terminal double helical extension, revealed the Watson-Crick pairing to primarily function as a kinetic trap in the G4 formation, adding a new method to the toolbox.

Finally, the review (publication IV) summarizes principles gathered in the previous publications and contextualizes them with other nucleoside modifications such as $^{F}$G and $^{LNA}$G analogs with regards to tailored quadruplex folding. Thereby valuable insights into the G4 folding process could be gained and efficient strategies for rational G4 design could be crystallized.

In summary, the engineered structures illustrate the possibilities of rational quadruplex design, applying lessons learned from the growing G4 topological landscape and the characterization of novel structures. The assembled toolbox consists of various nucleoside analogs and kinetic traps, paving the way of a true rational design of G4 architectures with desired structural characteristics for a growing number of G4-based applications.

# Bibliography


[1] J. L. Huppert and S. Balasubramanian, "Prevalence of quadruplexes in the human genome," *Nucleic Acids Research*, vol. 33, no. 9, pp. 2908–2916, 2005.

[2] V. S. Chambers, G. Marsico, J. M. Boutell, M. D. Antonio, G. P. Smith, and S. Balasubramanian, "High-throughput sequencing of DNA G-quadruplex structures in the human genome," *Nature Biotechnology*, vol. 33, no. 8, pp. 877–881, 2015.

[3] G. Biffi, D. Tannahill, J. McCafferty, and S. Balasubramanian, "Quantitative visualization of DNA G-quadruplex structures in human cells," *Nature Chemistry*, vol. 5, no. 3, pp. 182–186, 2013.

[4] R. Hänsel-Hertsch, D. Beraldi, S. V. Lensing, G. Marsico, K. Zyner, A. Parry, M. D. Antonio, J. Pike, H. Kimura, M. Narita, D. Tannahill, and S. Balasubramanian, "G-quadruplex structures mark human regulatory chromatin," *Nature Genetics*, vol. 48, no. 10, pp. 1267–1272, 2016.

[5] T. M. Fletcher, D. Sun, M. Salazar, and L. H. Hurley, "Effect of DNA secondary structure on human telomerase activity," *Biochemistry*, vol. 37, no. 16, pp. 5536–5541, 1998.

[6] W. N. Kim, M. A. Piatyszek, K. R. Prowse, C. B. Harley, M. D. West, P. L. C. Ho, G. M. Coviello, W. E. Wright, S. L. Weinrich, and J. W. Shay, "Specific association of human telomerase activity with immortal cells and cancer," *Science*, vol. 266, no. 5193, pp. 2011–2015, 1994.

[7] R. Hänsel-Hertsch, M. Di Antonio, and S. Balasubramanian, "DNA G-quadruplexes in the human genome: Detection, functions and therapeutic potential," *Nature Reviews Molecular Cell Biology*, vol. 18, no. 5, pp. 279–284, 2017.

[8] V. V. Negrutska, L. V. Dubey, M. M. Ilchenko, and I. Y. Dubey, "Design and study of telomerase inhibitors based on G-quadruplex ligands," *Biopolymers and Cell*, vol. 29, no. 3, pp. 169–176, 2013.

[9] X. Xu, F. Hamhouyia, S. D. Thomas, T. J. Burke, A. C. Girvan, W. G. McGregor, J. O. Trent, D. M. Miller, and P. J. Bates, "Inhibition of DNA replication and induction of S phase cell cycle arrest by G-rich oligonucleotides," *Journal of Biological Chemistry*, vol. 276, no. 46, pp. 43221–43230, 2001.

[10] L. F. Liu and J. C. Wang, "Supercoiling of the DNA template during transcription," *Proceedings of the National Academy of Sciences*, vol. 84, no. 20, pp. 7024–7027, 2006.

[11] C. V. Dang, "c-Myc target genes involved in cell growth, apoptosis, and metabolism," *Molecular and Cellular Biology*, vol. 19, no. 1, pp. 1–11, 1999.

[12] D. Sun and L. H. Hurley, "The importance of negative superhelicity in inducing the formation of G-quadruplex and i-motif structures in the c-Myc promoter: implications for drug targeting and control of gene expression," *Journal of medicinal chemistry*, vol. 52, no. 9, pp. 2863–74, 2009.

[13] C. Sutherland, Y. Cui, H. Mao, and L. H. Hurley, "A mechanosensor mechanism controls the G-quadruplex/i-motif molecular switch in the MYC promoter NHE $III_1$," *Journal of the American Chemical Society*, vol. 138, no. 42, pp. 14158–14151, 2016.

[14] E. D. Larson, M. L. Duquette, W. J. Cummings, R. J. Streiff, and N. Maizels, "MutS$\alpha$ binds to and promotes synapsis of transcriptionally activated immunoglobulin switch regions," *Current Biology*, vol. 15, no. 237, pp. 470–474, 2005.

[15] S. Cogoi, V. Rapozzi, S. Cauci, and L. E. Xodo, "Critical role of hnRNP A1 in activating KRAS transcription in pancreatic cancer cells: A molecular mechanism involving G4 DNA," *Biochimica et Biophysica Acta (BBA) - General Subjects*, vol. 1861, no. 5B, pp. 1389–1398, 2016.

[16] P. Sengupta, S. Chattopadhyay, and S. Chatterjee, "G-Quadruplex surveillance in BCL-2 gene: a promising therapeutic intervention in cancer treatment," *Drug Discovery Today*, vol. 22, no. 8, pp. 1165–1186, 2017.

[17] S. Q. Mao, A. T. Ghanbarian, J. Spiegel, S. Martínez Cuesta, D. Beraldi, M. Di Antonio, G. Marsico, R. Hänsel-Hertsch, D. Tannahill, and S. Balasubramanian, "DNA G-quadruplex structures mold the DNA methylome," *Nature Structural and Molecular Biology*, vol. 25, no. 10, pp. 951–957, 2018.

[18] V. V. Cheong, C. J. Lech, B. Heddi, and A. T. Phan, "Inverting the G-tetrad polarity of a G-quadruplex by using xanthine and 8-oxoguanine," *Angewandte Chemie - International Edition*, vol. 55, no. 1, pp. 160–163, 2016.

[19] C. Zhou, Z. Tan, C. Wang, Z. Wei, Z. Wang, C. Bai, J. Qin, and E. Cao, "Branched nanowire based guanine rich oligonucleotides," *Journal of Biomolecular Structure and Dynamics*, vol. 18, no. 6, pp. 807–812, 2001.

[20] A. Calzolari, R. Di Felice, and E. Molinari, "Electronic properties of guanine-based nanowires," *Solid State Communications*, vol. 131, pp. 557–564, 2004.

[21] R. P. Fahlman, M. Hsing, C. S. Sporer-Tuhten, and D. Sen, "Duplex pinching: A structural switch suitable for contractile DNA nanoconstructions," *Nano Letters*, vol. 3, no. 8, pp. 1073–1078, 2003.

[22] P. Alberti and J.-L. Mergny, "DNA duplex–quadruplex exchange as the basis for a nanomolecular machine," *Proceedings of the National Academy of Sciences*, vol. 100, no. 4, pp. 1569–1573, 2003.

[23] J. Jianwei, L. Tan, and W. Tan, "A single DNA molecule nanomotor," *Nano Letters*, vol. 2, no. 4, pp. 315–318, 2005.

[24] R. Rodriguez, G. D. Pantoş, D. P. N. Gonçalves, J. K. M. Sanders, and S. Balasubramanian, "Ligand-driven G-quadruplex conformational switching by using an unusual mode of interaction," *Angewandte Chemie - International Edition*, vol. 46, no. 28, pp. 5405–5407, 2007.

[25] H. Sun, J. Xiang, W. Gai, Q. Shang, Q. Li, A. Guan, Q. Yang, Y. Liu, Y. Tang, and G. Xu, "Visual detection of potassium by a cyanine dye supramolecular aggregate responsive to G-quadruplex motif transition," *The Analyst*, vol. 137, no. 1, pp. 5713–5715, 2012.

[26] E. Largy, A. Marchand, S. Amrane, V. Gabelica, and J. L. Mergny, "Quadruplex turncoats: cation-dependent folding and stability of quadruplex-DNA double switches," *Journal of the American Chemical Society*, vol. 138, no. 8, pp. 2780–2792, 2016.

[27] D. Lévesque, J.-D. Beaudoin, S. Roy, and J.-P. Perreault, "In vitro selection and characterization of RNA aptamers binding thyroxine hormone," *Biochemical Journal*, vol. 403, no. 1, pp. 129–138, 2007.

[28] Y. Xiao, B. D. Piorek, K. W. Plaxco, and A. J. Heegert, "A reagentless signal-on architecture for electronic, aptamer-based sensors via target-induced strand displacement," *Journal of the American Chemical Society*, vol. 127, no. 51, pp. 17990–17991, 2005.

[29] A. Bourdoncle, A. E. Torres, C. Gosse, L. Lacroix, P. Vekhoff, T. Le Saux, L. Jullien, and J. L. Mergny, "Quadruplex-based molecular beacons as tunable DNA probes," *Journal of the American Chemical Society*, vol. 128, no. 34, pp. 11094–11105, 2006.

[30] S. Nakayama and H. O. Sintim, "Colorimetric split G-quadruplex probes for nucleic acid sensing: improving reconstituted DNAzyme's catalytic efficiency via probe remodeling," *Journal of the American Chemical Society*, vol. 131, no. 29, pp. 10320–10333, 2009.

[31] E. Golub, R. Freeman, and I. Willner, "A hemin/G-quadruplex acts as an NADH oxidase and NADH peroxidase mimicking DNAzyme," *Angewandte Chemie - International Edition*, vol. 2, no. 50, pp. 11710–11714, 2011.

[32] P. Travascio, Y. Li, and D. Sen, "DNA-enhanced complex peroxidase activity of a DNA aptamer-hemin," *Chemistry & Biology*, vol. 5, no. 9, pp. 505–517, 1998.

[33] T. Shibata, Y. Nakayama, Y. Katahira, H. Tai, Y. Moritaka, Y. Nakano, and Y. Yamamoto, "Characterization of the interaction between heme and a parallel G-quadruplex DNA formed from d(TTGAGG)," *Biochimica et Biophysica Acta (BBA)- General Subjects*, vol. 1861, no. 5, pp. 1264–1270, 2017.

[34] A. Barlev and D. Sen, "DNA's Encounter with Ultraviolet Light: An Instinct for Self-Preservation?," *Accounts of Chemical Research*, vol. 51, no. 2, pp. 526–533, 2018.

[35] V. Dapić, P. J. Bates, J. O. Trent, A. Rodger, S. D. Thomas, and D. M. Miller, "Antiproliferative activity of G-quartet-forming oligonucleotides with backbone and sugar modifications," *Biochemistry*, vol. 41, no. 11, pp. 3676–3685, 2002.

[36] P. J. Bates, J. B. Kahlon, S. D. Thomas, J. O. Trent, and D. M. Miller, "Antiproliferative activity of G-rich oligonucleotides correlates with protein binding," *Journal of Biological Chemistry*, vol. 274, no. 37, pp. 26369–26377, 1999.

[37] A. R. Haeusler, C. J. Donnelly, G. Periz, E. A. J. Simko, P. G. Shaw, M.-s. Kim, N. J. Maragakis, J. C. Troncoso, A. Pandey, R. Sattler, J. D. Rothstein, and J. Wang, "C9orf72 nucleotide repeat structures initiate molecular cascades of disease," *Nature*, vol. 507, no. 7491, pp. 195–200, 2014.

[38] C. Callebaut, J. Blanco, N. Benkirane, B. Krust, E. Jacotot, G. Guichard, N. Seddiki, J. Svab, E. Dam, S. Muller, J. P. Briand, and A. G. Hovanessian, "Identification of V3 loop-binding proteins as potential receptors implicated in the binding of HIV particles to $CD4^+$ cells," *Journal of Biological Chemistry*, vol. 273, no. 34, pp. 21988–21997, 1998.

[39] A. Castaño and M. S. Maurer, "Aptamers as targeted therapeutics: current potential and challenges," *Nature Reviews Drug Discovery*, vol. 20, no. 2, pp. 163–178, 2015.

[40] P. J. Bates, E. M. Reyes-Reyes, M. T. Malik, E. M. Murphy, M. G. O'Toole, and J. O. Trent, "G-quadruplex oligonucleotide AS1411 as a cancer-targeting agent: Uses and mechanisms," *Biochimica et Biophysica Acta (BBA) - General Subjects*, vol. 1861, pp. 1414–1428, may 2017.

[41] A. C. Girvan, Y. Teng, L. K. Casson, S. D. Thomas, S. Jüliger, M. W. Ball, J. B. Klein, J. W. M. Pierce, S. S. Barve, and P. J. Bates, "AGRO100 inhibits activation of nuclear factor-κB (NF-κB) by forming a complex with NF-κB essential modulator (NEMO) and nucleolin," *Molecular Cancer Therapeutics*, vol. 5, no. 7, pp. 1790–1799, 2006.

[42] A. T. Phan, V. Kuryavyi, J.-B. Ma, A. Faure, M.-L. Andreola, and D. J. Patel, "An interlocked dimeric parallel-stranded DNA quadruplex: A potent inhibitor of HIV-1 integrase," *Proceedings of the National Academy of Sciences*, vol. 102, no. 3, pp. 634–639, 2005.

[43] J. R. Wyatt, T. A. Vickers, J. L. Roberson, R. W. Buckheit, T. Klimkait, E. DeBaets, P. W. Davis, B. Rayner, J. L. Imbach, and D. J. Ecker, "Combinatorially selected guanosine-quartet structure is a potent inhibitor of human immunodeficiency virus envelope-mediated cell fusion," *Proceedings of the National Academy of Sciences*, vol. 91, no. 4, pp. 1356–60, 1994.

[44] D. J. Schneider, J. Feigon, Z. Hostomsky, and L. Gold, "High-affinity ssDNA inhibitors of the reverse transcriptase of type 1 human immunodeficiency virus," *Biochemistry*, vol. 34, no. 29, pp. 9599–9610, 1995.

[45] K. T. Shum and J. A. Tanner, "Differential inhibitory activities and stabilisation of DNA aptamers against the SARS coronavirus helicase," *ChemBioChem*, vol. 9, no. 18, pp. 3037–3045, 2008.

[46] L. A. Jones, L. E. Clancy, W. D. Rawlinson, and P. A. White, "High-affinity aptamers to subtype 3a hepatitis C virus polymerase display genotypic specificity," *Antimicrobial Agents and Chemotherapy*, vol. 50, no. 9, pp. 3019–3027, 2006.

[47] L. Bock, L. Griffin, J. Latham, E. Vermaas, and J. Toole, "Selection of single-stranded DNA molecules that bind and inhibit human thrombin," *Nature*, vol. 355, pp. 564–566, 1992.

[48] K. Padmanabhan, K. P. Padmanabhan, J. D. Ferrara, J. E. Sadler, and A. Tulinsky, "The structure of α-thrombin inhibited by a 15-mer single-stranded DNA aptamer," *Journal of Biological Chemistry*, vol. 268, no. 24, pp. 17651–17654, 1993.

[49] P. Schultze, R. Macaya, and J. Feigon, "Three-dimensional solution structure of the thrombin-binding DNA aptamer d(GGTTGGTGTGGTTGG)," *Journal of Molecular Biology*, vol. 235, no. 5, pp. 1532–1547, 1994.

[50] A. S. Wolberg and R. A. Campbell, "Thrombin generation, fibrin clot formation and hemostasis," *Transfusion and Apheresis Sciences*, vol. 38, no. 1, pp. 15–23, 2008.

[51] E. Zavyalova, N. Samoylenkova, A. Revishchin, A. Turashev, I. Gordeychuk, A. Golovin, A. Kopylov, and G. Pavlova, "The evaluation of pharmacodynamics and pharmacokinetics of anti-thrombin DNA aptamer RA-36," *Frontiers in Pharmacology*, vol. 8:922, pp. 1–12, 2017.

[52] R. Dolot, C. H. Lam, M. Sierant, Q. Zhao, F. W. Liu, B. Nawrot, M. Egli, and X. Yang, "Crystal structures of thrombin in complex with chemically modified thrombin DNA aptamers reveal the origins of enhanced affinity," *Nucleic Acids Research*, vol. 46, no. 9, pp. 4819–4830, 2018.

[53] X.-a. Mao and W. H. Gmeiner, "NMR study of the folding-unfolding mechanism for the thrombin-binding DNA aptamer d(GGTTGGTGTGGTTGG)," *Biophysical chemistry*, vol. 113, no. 2, pp. 155–60, 2005.

[54] A. Pasternak, F. J. Hernandez, L. M. Rasmussen, B. Vester, and J. Wengel, "Improved thrombin binding aptamer by incorporation of a single unlocked nucleic acid monomer," *Nucleic Acids Research*, vol. 39, no. 3, pp. 1155–1164, 2011.

[55] A. Aviñó, C. Fabrega, M. Tintoré, and R. Eritja, "Thrombin binding aptamer, more than a simple aptamer: Chemically modified derivatives and biomedical applications," *Current Pharmaceutical Design*, vol. 18, no. 14, pp. 2036–2047, 2012.

[56] P. Hazel, G. N. Parkinson, and S. Neidle, "Predictive modelling of topology and loop variations in dimeric DNA quadruplex structures," *Nucleic Acids Research*, vol. 34, no. 7, pp. 2117–2127, 2006.

[57] J. Zhou, A. Bourdoncle, F. Rosu, V. Gabelica, and J. L. Mergny, "Tri-G-quadruplex: Controlled assembly of a G-quadruplex structure from three G-rich strands," *Angewandte Chemie - International Edition*, vol. 51, no. 44, pp. 11002–11005, 2012.

[58] F. Aboul-ela, A. I. H. Murchie, and D. M. Lilley, "NMR study of parallel-stranded tetraplex formation by the hexadeoxynucleotide d($TG_4T$)," *Nature*, vol. 360, pp. 280–282, 1992.

[59] P. Hazel and J. Huppert, "Loop-length-dependent folding of G-quadruplexes," *Journal of the American Chemical*

*Society*, vol. 126, no. 5, pp. 29–39, 2004.

[60] N. Smargiasso, F. Rosu, W. Hsia, P. Colson, E. S. Baker, M. T. Bowers, E. De Pauw, and V. Gabelica, "G-quadruplex DNA assemblies: Loop length, cation identity, and multimer formation," *Journal of the American Chemical Society*, vol. 130, no. 31, pp. 10208–10216, 2008.

[61] X. Cang, J. Šponer, and T. E. Cheatham III, "Insight into G-DNA structural polymorphism and folding from sequence and loop connectivity through free energy analysis," *Journal of the American Chemical Society*, vol. 133, no. 36, pp. 14270–14279, 2011.

[62] V. Mukundan and A. Phan, "Bulges in G-quadruplexes: Broadening the definition of G-quadruplex-forming sequences," *Journal of the American Chemical Society*, vol. 135, no. 13, pp. 5017–5028, 2013.

[63] J. T. Nielsen, K. Arar, and M. Petersen, "Solution structure of a locked nucleic acid modified quadruplex: Introducing the V4 folding topology," *Angewandte Chemie - International Edition*, vol. 48, no. 17, pp. 3099–3103, 2009.

[64] B. Bakalar, B. Heddi, E. Schmitt, Y. Mechulan, and A. T. Phan, "A minimal sequence for left-handed G-quadruplex formation," *Angewandte Chemie International Edition*, vol. 58, no. 8, pp. 2331–2335, 2019.

[65] E. Gavathiotis and M. S. Searle, "Structure of the parallel-stranded DNA quadruplex d(TTAGGGT)$_4$ containing the human telomeric repeat: Evidence for A-tetrad formation from NMR and molecular dynamics simulations," *Organic & Biomolecular Chemistry*, vol. 1, no. 10, pp. 1650–1656, 2003.

[66] P. K. Patel, N. S. Bhavesh, and R. V. Hosur, "NMR observation of a novel C-tetrad in the structure of the SV40 repeat sequence GGGCGG," *Biochemical and biophysical research communications*, vol. 270, no. 3, pp. 967–71, 2000.

[67] M. Webba da Silva, "Association of DNA quadruplexes through G:C:G:C tetrads. Solution structure of d(GCGGTGGAT)," *Biochemistry*, vol. 42, no. 49, pp. 14356–14365, 2003.

[68] A. Kettani, A. Gorin, A. Majumdar, T. Hermann, E. Skripkin, H. Zhao, R. Jones, and D. J. Patel, "A dimeric DNA interface stabilized by stacked A · (G · G · G · G) · A hexads and coordinated monovalent cations," *Journal of molecular biology*, vol. 297, no. 3, pp. 627–644, 2000.

[69] D. Zhang, T. Huang, P. S. Lukeman, and P. J. Paukstelis, "Crystal structure of a DNA/Ba$^{2+}$ G-quadruplex containing a water-mediated C-tetrad," *Nucleic Acids Research*, vol. 42, no. 21, pp. 13422–13429, 2014.

[70] K. W. Lim, Z. J. Khong, and A. T. Phan, "Thermal stability of DNA quadruplex-duplex hybrids," *Biochemistry*, vol. 53, no. 1, pp. 247–257, 2014.

[71] K. W. Lim and A. T. Phan, "Structural basis of DNA quadruplex-duplex junction formation," *Angewandte Chemie - International Edition*, vol. 52, no. 33, pp. 8566–8569, 2013.

[72] M. Webba da Silva, "Geometric formalism for DNA qadruplex folding," *Chemistry - A European Journal*, vol. 13, no. 35, pp. 9738–9745, 2007.

[73] X. Cang, J. Šponer, and T. E. Cheatham III, "Explaining the varied glycosidic conformational, G-tract length and sequence preferences for anti-parallel G-quadruplexes," *Nucleic Acids Research*, vol. 39, no. 10, pp. 4499–4512, 2011.

[74] J. Li, J. J. Correia, L. Wang, J. O. Trent, and J. B. Chaires, "Not so crystal clear: The structure of the human telomere G-quadruplex in solution differs from that present in a crystal," *Nucleic Acids Research*, vol. 33, no. 14, pp. 4649–59, 2005.

[75] T. Fujii, P. Podbevšek, J. Plavec, and N. Sugimoto, "Effects of metal ions and cosolutes on G-quadruplex topology," *Journal of Inorganic Biochemistry*, vol. 166, pp. 190–198, 2016.

[76] Y. Liu, W. Lan, C. Wang, and C. Cao, "A putative G-quadruplex structure in the proximal promoter of VEGFR-2 has implications for drug design to inhibit tumor angiogenesis," *Journal of Biological Chemistry*, vol. 293, no. 23, pp. 8947–8955, 2018.

[77] J. Dai, M. Carver, C. Punchihewa, R. A. Jones, and D. Yang, "Structure of the Hybrid-2 type intramolecular human telomeric G-quadruplex in K$^+$ solution: Insights into structure polymorphism of the human telomeric sequence," *Nucleic Acids Research*, vol. 35, no. 15, pp. 4927–40, 2007.

[78] K. N. Luu, A. T. Phan, V. Kuryavyi, L. Lacroix, and D. J. Patel, "Structure of the human telomere in K$^+$ solution: an intramolecular (3 + 1) G-quadruplex scaffold," *Journal of the American Chemical Society*, vol. 128, no. 30, pp. 9963–70, 2006.

[79] B. Heddi and A. T. Phan, "Structure of human telomeric DNA in crowded solution," *Journal of the American Chemical Society*, vol. 133, no. 25, pp. 9824–9833, 2011.

[80] R. Hänsel, F. Löhr, S. Foldynová-Trantírková, E. Bamberg, L. Trantírek, and V. Dötsch, "The parallel G-quadruplex

structure of vertebrate telomeric repeat sequences is not the preferred folding topology under physiological conditions," *Nucleic Acids Research*, vol. 39, no. 13, pp. 5768–5775, 2011.

[81] J. Šponer, G. Bussi, P. Stadlbauer, P. Kührová, P. Banáš, B. Islam, S. Haider, S. Neidle, and M. Otyepka, "Folding of guanine quadruplex molecules–funnel-like mechanism or kinetic partitioning? An overview from MD simulation studies," *Biochimica et Biophysica Acta (BBA) - General Subjects*, vol. 1861, no. 5, pp. 1246–1263, 2017.

[82] S. S. Tavale and H. M. Sobell, "Crystal and molecular structure of 8-bromoguanosine and 8-bromoadenosine, two purine nucleosides in the syn conformation," *Journal of Molecular Biology*, vol. 48, no. 1, pp. 109–123, 1970.

[83] J. P. Gaut, G. C. Yeh, H. D. Tran, J. Byun, J. P. Henderson, G. M. Richter, M.-L. Brennan, A. J. Lusis, A. Belaaouaj, R. S. Hotchkiss, and J. W. Heinecke, "Neutrophils employ the myeloperoxidase system to generate antimicrobial brominating and chlorinating oxidants during sepsis," *Proceedings of the National Academy of Sciences*, vol. 98, no. 21, pp. 11961–11966, 2002.

[84] T. Asahi, H. Kondo, M. Masuda, H. Nishino, Y. Aratani, Y. Naito, T. Yoshikawa, S. Hisaka, Y. Kato, and T. Osawa, "Chemical and immunochemical detection of 8-halogenated deoxyguanosines at early stage inflammation," *Journal of Biological Chemistry*, vol. 285, no. 12, pp. 9282–9291, 2010.

[85] K. Shinmura, H. Kato, M. Goto, H. Tao, Y. Inoue, S. Nakamura, H. Yoshida, E. Tsuzaki, and H. Sugimura, "Mutation Spectrum Induced by 8-Bromoguanine, a Base Damaged by Reactive Brominating Species, in Human Cells," *Oxidative Medicine and Cellular Longevity*, vol. 2017 Article ID 7308501, pp. 1–11, 2017.

[86] J. Brčić and J. Plavec, "Solution structure of a DNA quadruplex containing ALS and FTD related GGGGCC repeat stabilized by 8-bromodeoxyguanosine substitution," *Nucleic Acids Research*, vol. 43, no. 17, pp. 8590–8600, 2015.

[87] A. T. Phan, V. Kuryavyi, K. N. Luu, and D. J. Patel, "Structure of two intramolecular G-quadruplexes formed by natural human telomere sequences in $K^+$ solution," *Nucleic Acids Research*, vol. 35, no. 19, pp. 6517–25, 2007.

[88] A. Matsugami, Y. Xu, Y. Noguchi, H. Sugiyama, and M. Katahira, "Structure of a human telomeric DNA sequence stabilized by 8-bromoguanosine substitutions, as determined by NMR in a $K^+$ solution," *FEBS Journal*, vol. 274, no. 14, pp. 3545–3556, 2007.

[89] V. Esposito, A. Randazzo, G. Piccialli, L. Petraccone, C. Giancola, and L. Mayol, "Effects of an 8-bromodeoxyguanosine incorporation on the parallel quadruplex structure $[d(TGGGT)]_4$," *Organic & biomolecular chemistry*, vol. 2, no. 3, pp. 313–318, 2004.

[90] D. Gudanis, L. Popenda, K. Szpotkowski, R. Kierzek, and Z. Gdaniec, "Structural characterization of a dimer of RNA duplexes composed of 8-bromoguanosine modified CGG trinucleotide repeats: a novel architecture of RNA quadruplexes," *Nucleic Acids Research*, vol. 44, no. 5, pp. 2409–2416, 2016.

# Author Contributions

**Publication I** **Observation of a Dynamic G-Tetrad Flip in Intramolecular G-Quadruplexes**

Karg, B., Haase, L., Funke, A., Dickerhoff, J. & Weisz, K. *Biochemistry*, 2016, 55, 6949–6955.

KW and JD initiated the project. KW and BK designed and BK performed the experiments. LH measured the spectra of MYC-4. AF performed the ligand binding studies. KW and BK wrote the manuscript that was read and edited by all authors.

**Publication II** **Loop Length Affects Syn-Anti Conformational Rearrangements in Parallel G-Quadruplexes**

Karg, B. & Weisz, K. *Chemistry - A European Journal* 2018, 24, 10246–10252.

KW initiated the project. KW and BK designed and BK performed the experiments. BK and KW wrote the manuscript.

**Publication III** **Duplex-Guided Refolding into Novel G-Quadruplex (3+1) Hybrid Conformations**

Karg, B., Mohr, S. & Weisz, K. *Angewandte Chemie*, 2019, DOI: 10.1002/anie.201905372

KW initiated the project. KW and BK designed and BK performed the experiments. SM performed the calorimetric experiments. BK and KW wrote the manuscript that was read and edited by all authors.

**Publication IV** **Manipulating DNA G-Quadruplex Structures Through Guanosine Analogs**

Haase, L., Karg, B. & Weisz, K., *ChemBioChem* 2018, 20, 985-993.

LH, BK and KW wrote the manuscript that was read and edited by all authors.

Prof. Dr. Klaus Weisz

Beatrice Karg

# Publication I

BIOCHEMISTRY
including biophysical chemistry & molecular biology

Article

pubs.acs.org/biochemistry

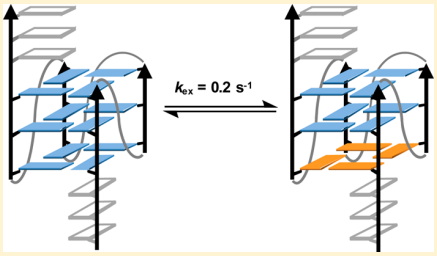
# Observation of a Dynamic G-Tetrad Flip in Intramolecular G-Quadruplexes

Beatrice Karg, Linn Haase, Andrea Funke, Jonathan Dickerhoff, and Klaus Weisz*

Institute of Biochemistry, Ernst-Moritz-Arndt University Greifswald, Felix-Hausdorff-Strasse 4, D-17487 Greifswald, Germany

Supporting Information

**ABSTRACT:** A *MYC* sequence forming an intramolecular G-quadruplex with a parallel topology was modified by the incorporation of 8-bromoguanosine ($^{Br}$G) analogues in one of its outer G-tetrads. The propensity of the $^{Br}$G analogues to adopt a *syn* glycosidic torsion angle results in an exceptional monomolecular quadruplex conformation featuring a complete flip of one tetrad while keeping a parallel orientation of all G-tracts as shown by circular dichroism and nuclear magnetic resonance spectroscopic studies. When substituting three of the four G-tetrad residues with $^{Br}$G analogues, two coexisting quadruplex conformational isomers with an all-*syn* and all-*anti* outer G-quartet are approximately equally populated in solution. A dynamic interconversion of the two quadruplexes with an exchange rate ($k_{ex}$) of 0.2 $s^{-1}$ is demonstrated through the observation of exchange crosspeaks in rotating frame Overhauser effect spectroscopy and nuclear Overhauser effect spectroscopy experiments at 50 °C. The kinetic properties suggest disruption of the corresponding outer G-tetrad but not of the whole quadruplex core during the tetrad flip. Conformational *syn-anti* isomers with homopolar and heteropolar stacking interactions are nearly isoenergetic with a transition enthalpy of 18.2 kJ/mol in favor of the all-*syn* isomer.




Guanine-rich sequences can assemble into stable quadruplex (G4) structures, primarily driven by the propensity of guanine (G) bases to self-associate and to form a planar arrangement with four G bases held together by a cyclic array of Hoogsteen hydrogen bonds.[1] At least two stacked G-tetrads make up the core in these alternative G4 secondary structures. Monovalent cations like sodium or potassium are coordinated to the guanine carbonyl oxygens within a central channel, being of critical importance for additional stabilization. Because of their existence in many regulatory regions of the human genome and their proposed biological function, quadruplex structures have become an attractive target for novel anticancer strategies.[2] Also, being promising scaffolds for various aptamer and sensor systems, synthetic quadruplexes have found increasing use in biotechnological applications.[3,4]

If G-rich tracts are located on a single strand, folding into a quadruplex is associated with intervening sequences that might form different types of loop regions. Thus, propeller loops connect two neighboring parallel G-tracts, whereas lateral or diagonal loops join two antiparallel G-tracts located adjacent or opposite each other. The combination of different loop arrangements results in a high variability of quadruplex topologies, often close in stability and sensitive to the composition of loop and overhang sequences but also to environmental conditions like type of cations, crystal packing, or the presence of crowding agents.[5] Whereas an increasing amount of data has yielded empirical relationships between loop type and intervening sequences, the superposition of forces that drive a particular fold is far from being understood. As a consequence, a reliable prediction or rational design of quadruplex structures remains unsatisfactory as yet.

To form a regular array of Hoogsteen hydrogen bonds, glycosidic torsion angles of G nucleotides within the quadruplex core are closely connected to the relative orientation of the four strands. In a parallel quadruplex with all four G-tracts running in the same direction, all G residues typically feature *anti* conformations. Reverting the polarity of one strand will inevitably induce glycosidic conformational changes to *syn* for some G nucleotides to maintain a stable hydrogen-bonded arrangement. Such a dependency has been exploited in the past to either stabilize or refold quadruplexes by replacing guanosines at suitable positions with G analogues favoring either *syn* or *anti* glycosidic torsion angles.[6,7] Employing such a strategy, a complete flip of a tetrad with all four guanine nucleotides changing from *anti* to *syn* in a conserved strand arrangement was reported upon the incorporation of 8-methyl- or 8-bromoguanosine analogues into critical positions of parallel four-stranded model quadruplexes lacking any loop regions.[8−10] Through this reversal of tetrad polarity as defined by the donor−acceptor direction of the cyclic hydrogen bond array, homopolar stacking interactions are replaced by a heteropolar tetrad stacking with accompanying changes in spectroscopic behavior, in particular apparent in circular dichroism.[11] Very recently, a similar conformational switch has also been observed in more relevant monomolecular quadruplexes with connecting loop regions. To induce such a transition, either *anti*-favoring 2′-fluoro-2′-deoxyguanosine analogues were incorporated into a hybrid-

**Received:** September 9, 2016
**Revised:** November 6, 2016
**Published:** November 15, 2016





DOI: 10.1021/acs.biochem.6b00925
*Biochemistry* 2016, 55, 6949−6955

type quadruplex or two foreign nucleobases were introduced into a single G-tetrad of the parent G4.[12,13]

This paper reports on the incorporation of *syn*-affine 8-bromodeoxyguanosine analogues into specific positions of the MYC sequence. The latter derives from the promoter region of the MYC oncogene and forms a well-defined, stable, and robust quadruplex structure in the presence of monovalent cations.[14] We aimed to expand the concept of concerted conformational changes within a G-tetrad to the parallel topology of the MYC quadruplex to further elaborate on the general requirements and properties of tetrad flipping. Indeed, as shown below, a tetrad polarity reversal with the formation of an unusual all-*syn* G-quartet may be induced in an 8-bromodeoxyguanosine-modified MYC quadruplex without a more extensive refolding of the monomolecular structure. Interestingly, native and flipped tetrad conformations may coexist in solution as has previously been suggested for a nonmodified tetramolecular quadruplex.[15] The energy barrier for interconversion of these conformational isomers allows for a dynamic equilibrium as shown by nuclear magnetic resonance (NMR) exchange experiments, revealing the dynamic nature of tetrad flipping in an intramolecular quadruplex for the first time.

## MATERIALS AND METHODS

**Materials and Sample Preparation.** Unmodified and high-performance liquid chromatography-purified 8-bromo-dG-modified DNA oligonucleotides were purchased from TIB MOLBIOL (Berlin, Germany). Before use, oligonucleotides were precipitated with ethanol, and the concentrations were determined spectrophotometrically by measuring the absorbance at 260 nm. Samples were obtained by dissolving the corresponding oligonucleotides in a low-salt buffer with 10 mM potassium phosphate (pH 7.0) for NMR experiments or in high-salt buffer with 20 mM potassium phosphate and 100 mM KCl (pH 7.0) for CD experiments. Prior to measurements, the samples were annealed by being heated to 90 °C and then slowly cooled to room temperature. Final concentrations of oligonucleotides were 5 $\mu$M for the CD experiments and between 0.11 and 0.52 mM for the NMR experiments.

**Circular Dichroism.** CD spectra were recorded with a Jasco J-810 spectropolarimeter at 20 °C (Jasco, Tokyo, Japan). The spectra were recorded with a bandwidth of 1 nm, a scanning speed of 50 nm/min, and five accumulations. All spectra were blank corrected. CD melting curves were recorded at a wavelength of 295 nm with two data points per degree Celsius in 10 mm quartz cuvettes. Heating and cooling rates of 0.25 °C/min were employed. Melting temperatures were determined by the maximum of the first derivative for the heating curve.

**Differential Scanning Calorimetry (DSC).** DSC measurements were performed on a VP-DSC instrument (Malvern Instruments) using 50 $\mu$M oligonucleotide in a buffer with 20 mM potassium phosphate and 100 mM KCl (pH 7.0). The sample was heated from 20 to 130 °C with a scan rate of 1 °C/min. A buffer versus buffer scan was subtracted from the sample scan, and the melting temperature ($T_m$) corresponding to the maximum of the DSC peak was obtained after a linear baseline subtraction. Melting temperatures with standard deviations are averages of five experiments.

**NMR Experiments.** All NMR spectra were recorded on a Bruker Avance 600 MHz spectrometer equipped with an inverse $^{1}H/^{13}C/^{15}N/^{31}P$ quadruple resonance cryoprobehead and z-field gradients. Data were processed using Topspin 3.1 and analyzed with CcpNmr Analysis.[16] Proton chemical shifts were referenced relative to the temperature-dependent $H_2O$ or HDO chemical shift. For the proton one-dimensional (1D) and two-dimensional (2D) NOE measurements in a 90% $H_2O$/10% $D_2O$ solvent, a WATERGATE with w5 element was employed for solvent suppression. Typically, NOESY experiments in a 90% $H_2O$/10% $D_2O$ solvent were performed between 10 and 50 °C with a mixing time of 300 ms and a spectral width of 10 kHz; 2K × 1K data points with 32 transients per $t_1$ increment and a recycle delay of 2 s were collected in $t_2$ and $t_1$. Prior to Fourier transformation, data were zero-filled to give a 4K × 2K matrix and both dimensions were apodized by squared sine bell window functions. For EASY ROESY experiments,[17] a mixing time of 80 ms was employed with the angle $\theta$ for the tilted rotating frame fixed to 50°.

DQF-COSY experiments were recorded with 24 transients per $t_1$ increment and 4K × 900 data points that were zero-filled to give a final 4K × 2K data matrix. Phase-sensitive $^{1}H-^{13}C$ HSQC experiments optimized for a $^{1}J(C,H)$ of 160 Hz were conducted with a 3-9-19 solvent suppression scheme in a 90% $H_2O$/10% $D_2O$ solvent employing a spectral width of 7.5 kHz in the indirect $^{13}C$ dimension, 128 scans at each of 360 $t_1$ increments, and a relaxation delay of 1.5 s between scans. $^{13}C$ chemical shifts were referenced relative to TMS by using the indirect referencing method.

## RESULTS AND DISCUSSION

Previous substitution experiments on tetramolecular quadruplexes with 8-methyl- or 8-bromo-dG analogues have suggested a higher likelihood of substitutions at the 5′-terminal tetrad to enforce altered glycosidic torsion angles.[8−10] On the basis of these studies, we focused on modifications within the terminal tetrad following the 5′-end (5′-tetrad) for inducing putative structural transitions on the intramolecular *MYC* quadruplex. The unmodified *MYC* sequence employed for these studies is known to form a well-defined quadruplex with parallel G-tracts connected by three propeller loops and overhangs at both its 5′- and 3′-end in a potassium-containing buffer (Figure 1a).[14] To probe the impact of G analogues on the quadruplex conformation, three *syn*-favoring 8-bromo-2′-deoxyriboguanosines ($^{Br}G$) were initially incorporated at positions 4, 13, and 17 of *MYC* to yield the modified sequence *MYC*-3 (Table 1). This trisubstitution pattern keeps the H8 proton of residue G8 as a

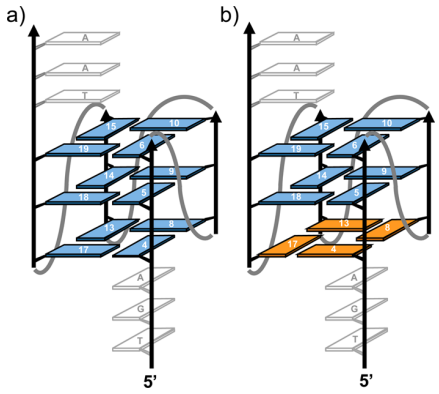



**Figure 1.** (a) Parallel topology of the native *MYC* G-quadruplex with overhang sequences and guanine residue numbers in the G-tetrads indicated. (b) *MYC* quadruplex with a polarity switch of the 5′-tetrad. The *anti* and *syn* guanine nucleotides are colored blue and orange, respectively.



DOI: 10.1021/acs.biochem.6b00925
*Biochemistry* 2016, 55, 6949−6955

valuable NMR probe for following conformational changes of the modified outer G-quartet (*vide infra*).

**Table 1. Sequences and CD Melting Temperatures of Unmodified and 8-$^{Br}$G-Modified *MYC* Quadruplexes[a]**

| | sequence (5′ → 3′)[b] | $T_m$ (°C)[c] |
|---|---|---|
| *MYC* | TGAGGGTGGGTAGGGTGGGTAA | 90.8 ± 0.2[d] |
| *MYC-3* | TGAXGGTGGGTAXGGTXGGTAA | 69.7 ± 0.4 |
| *MYC-4* | TGAXGGTXGGTAXGGTXGGTAA | 70.7 ± 0.7 |

[a]Conditions: 5 μM quadruplex in 20 mM potassium phosphate and 100 mM KCl (pH 7.0). [b]X = 8-bromo-dG; G residues in tetrads are underlined. [c]Average value with the standard deviation from three independent measurements. [d]Determined by DSC measurements because of the high thermal stability in $K^+$ buffer; averaged value from five experiments.

As expected for a parallel all-*anti* topology, the CD spectrum of the native *MYC* exhibits negative and positive amplitudes at 244 and 264 nm, respectively (Figure 2). Structural changes are

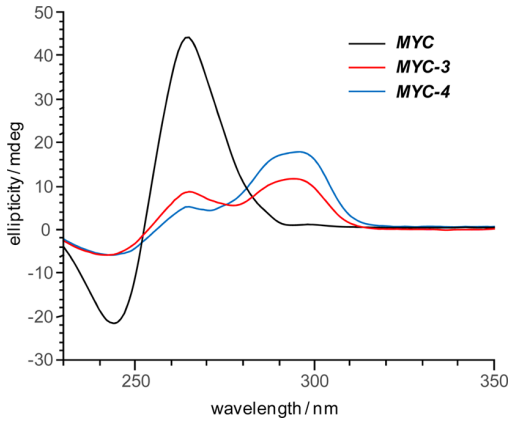



**Figure 2.** CD spectra of unmodified *MYC*, *MYC-3*, and *MYC-4* sequences (5 μM) at 20 °C in 20 mM potassium phosphate and 100 mM KCl (pH 7).

suggested by the corresponding CD signature of the $^{Br}$G-modified *MYC-3*. Thus, amplitudes at 244 and 264 nm are noticeably reduced, and a new strong band at 295 nm reminiscent of a hybrid-type quadruplex with homopolar as well as heteropolar G-tetrad stacking appears. Melting experiments indicate that *MYC-3* forms a thermally stable structure with a melting transition ($T_m$) centered at 70 °C. However, compared to that of the native structure, there is a noticeable decrease in $T_m$ by 20 °C upon incorporation of the $^{Br}$G analogues (Table 1).

**NMR Spectroscopic Analysis of $^{Br}$G-Modified Quadruplexes.** To check for structural homogeneity, 1D $^1$H NMR experiments were performed. As shown in Figure 3a, a total of 12 sharp and well-resolved guanine imino resonances are observed between 10.5 and 12.0 ppm in the NMR spectrum of unmodified *MYC*. These findings confirm the formation of a single well-defined quadruplex with a core consisting of three stacked G-tetrads. On the other hand, the imino proton spectral region of *MYC-3* reveals an increased number of G imino resonances that are incompatible with a single monomolecular species (Figure 3b). In fact, a total of 24 partially overlapped imino signals rather suggest the formation of a quadruplex with higher molecularity

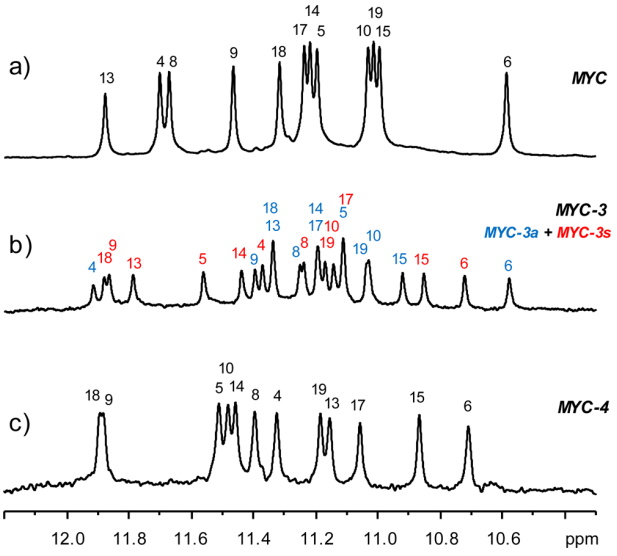



**Figure 3.** Imino proton spectral region of (a) *MYC*, (b) *MYC-3*, and (c) *MYC-4* at 30 °C in 10 mM potassium phosphate (pH 7). Peak assignments for the G-tract guanines are indicated.

or the coexistence of two quadruplex structures in approximately equal amounts.

Strong support for the latter comes from a spectral comparison of *MYC-3* with unmodified *MYC*. Conspicuously, many characteristic crosspeak patterns observed in 2D NOE, DQF-COSY, and $^1$H−$^{13}$C HSQC spectra of *MYC* are also identified in corresponding *MYC-3* spectral regions. The spectral superposition shows only small shift deviations, especially for those residues that are not directly involved in the modified 5′-tetrad, and clearly indicates a *MYC-3* quadruplex with a conserved *MYC*-like topology (*MYC-3a*) but also reveals additional resonances that can be attributed to another species. Several starting points for the $^1$H and $^{13}$C resonance assignments of *MYC-3a* are provided by their close correspondence to the already assigned resonances of native *MYC* (Figure S1). Thus, despite the presence of two sets of signals, almost all of the base as well as H1′/H2′/H2″ sugar resonances are easily assigned from standard 2D NMR experiments because of their good spectral resolution (see the Supporting Information). Sequential and intranucleotide sugar−H8 NOE contacts along three G-tracts comprise only two G bases in case of $^{Br}$G substitutions. However, uninterrupted connectivities can be traced for the G8-G9-G10 tract bearing no modification and confirm an all-*anti* glycosidic conformation for all G residues of the *MYC-3a* quadruplex core (Figure S2).[18]

At elevated temperatures (>40 °C), 2D NOE spectra of the trisubstituted *MYC-3* exhibit increasingly strong crosspeaks identified as exchange crosspeaks through corresponding ROESY experiments. Part of an EASY ROESY spectrum[17] acquired at 50 °C on *MYC-3* is shown in Figure 4. Apparently, the spectrum at this temperature is dominated by exchange crosspeaks identified by their same sign as found for the diagonal peaks, and no additional contacts through dipolar coupling are observed under these conditions. Because the ROE crosspeaks connect corresponding protons of the two coexisting structures, they also allow for the straightforward identification of resonances in the second *MYC-3* species (*MYC-3s*) and can be used as convenient starting points for tracing additional NOE connectivities in the NOESY spectra.



DOI: 10.1021/acs.biochem.6b00925
*Biochemistry* 2016, 55, 6949−6955

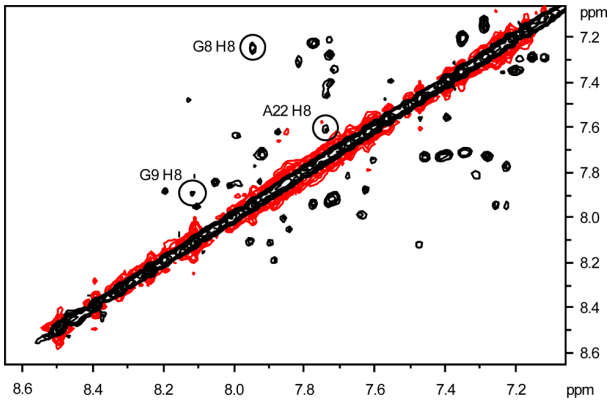



**Figure 4.** Portion of an EASY ROESY spectrum of *MYC-3* at 50 °C (80 ms mixing time) showing exchange crosspeaks of H8/H6 base protons. H8 crosspeaks of G8, G9, and A22 are highlighted.

Especially for residues located in loops and in the 3′-tetrad, H6/H8−C6/C8 correlations as observed in a $^{1}$H−$^{13}$C HSQC spectrum show only small chemical shift differences in *MYC-3a* and *MYC-3s*. In contrast, residue G8 with an upfield-shifted H8 is accompanied by a significant C8 downfield shift of 3.4 ppm in *MYC-3s* (Figure 5). In general, changing an *anti* conformation to

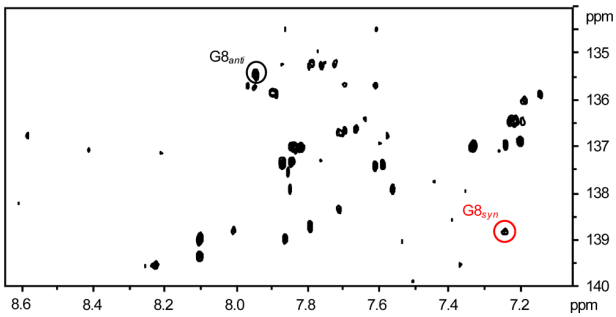



**Figure 5.** Portion of a $^{1}$H−$^{13}$C HSQC spectrum at 40 °C showing H8/H6−C8/C6 correlations of *MYC-3*. The G8 H8−C8 crosspeak of the two coexisting species is marked by the black and red circle.

a *syn* conformation will induce downfield shifts of guanine C8 by 3−6 ppm.[12,19,20] This C8 deshielding effect was shown to be quite insensitive to other torsion angles and constitutes a reliable marker of an *anti* → *syn* interconversion for G8 in *MYC-3s*. Note that H8−C8 correlations are missing for the other residues of the 5′-tetrad because of their 8-bromine substituent.

Resonance assignments for *MYC-3s* are additionally confirmed and complemented by standard assignment strategies based on NOESY and DQF-COSY spectra (see the Supporting Information). Whereas continuous H1′−H8 NOE connectivities between G residues of the central tetrad and 3′-tetrad demonstrate their all-*anti* conformation, the single strong intranucleotide H8−H1′ NOE contact observed for G8 again confirms its *syn* configuration (Figure S2).[18] Consequently, with all four G-tracts in a parallel orientation, the other three modified $^{Br}$G residues in the 5′-tetrad must also adopt *syn* glycosidic torsion angles associated with a complete G-tetrad flip.

Chemical shifts of sugar H2′ protons were shown to constitute good markers for *anti* and *syn* nucleotides in nucleic acid structures.[19,21] Although there is no direct information about glycosidic torsion angles for the three $^{Br}$G analogues in the absence of H8, a clear trend toward more deshielded H2′ protons as observed in *MYC-3s* when compared to *MYC* or *MYC-3a* clearly corroborates a corresponding switch from *anti* to *syn* conformations. Thus, whereas a mean chemical shift difference ($\Delta\delta$) of only +0.09 ppm is observed between H2′ protons in the 5′-tetrad of *MYC-3a* and unmodified *MYC*, a significant downfield shift of +0.28 ppm when averaged over all H2′ resonances of the 5′-tetrad is detected when comparing *MYC-3s* with *MYC-3a*. Note that major contributions of the bromo substituent to H2′ chemical shift changes can mostly be excluded in the latter case.

These results show that the incorporation of three $^{Br}$G analogues into the 5′-tetrad of the intramolecular *MYC* quadruplex induces a flip associated with a polarity reversal of the outer G-quartet despite putative restraints exerted by connecting loops (see Figure 1b), yet two conformational isomers with and without *anti* → *syn* transitions coexist in solution in approximately equal amounts. To test the impact of a fully $^{Br}$G modifed 5′-outer tetrad on the type and population of formed species, we also studied a *MYC* sequence tetrasubstituted at positions 4, 8, 13, and 17 (*MYC-4*) (see Figure 1). As observed for the trisubstituted *MYC-3*, the CD spectrum of *MYC-4* exhibits a negative amplitude at 244 nm and two moderately intense positive amplitudes at 264 and 295 nm (Figure 2). However, when compared to that of *MYC-3*, the CD spectrum of *MYC-4* shows some noticeable changes in intensity of its two positive bands. Thus, with an increase at 295 nm and a decrease at 264 nm, it appears to further shift toward the spectrum of antiparallel quadruplexes with enhanced heteropolar stacking interactions.

Melting temperatures of *MYC-4* and *MYC-3* were found to be almost identical under the conditions employed with a thermal destabilization of the tetrasubstituted quadruplex in comparison to that of native *MYC* (Table 1). On the other hand, the imino proton spectral region in a one-dimensional $^{1}$H NMR spectrum of *MYC-4* shows a total of only 12 guanine imino resonances partially shifted compared to the parent *MYC* (Figure 3c). These findings again indicate the formation of a single intramolecular *MYC-4* quadruplex composed of three stacked G-tetrads.

For a more detailed structural characterization, 2D NOE spectra were subsequently recorded for *MYC-4*. In contrast to *MYC-3*, all G residues within the 5′-tetrad have been replaced with $^{Br}$G analogues and interrupt sugar−H8 contacts along all four G-tracts. On the other hand, inspection of the 2D NOE spectra reveals a close match to corresponding NOEs as found for the *MYC-3s* quadruplex (Figure S4). This striking similarity strongly suggests equal conformations for both quadruplexes and allows the unambiguous assignment through the existing network of NOE connectivities (Figures S5 and S6). In line with an *anti* → *syn* conformational transition of the outer tetrad, H2′ protons within the $^{Br}$G-substituted G-quartet experience noticeable deshielding effects as already observed for the *MYC-3s* glycosidic isomer.

Changes in H6/H8 proton chemical shifts of *MYC-3a*, *MYC-3s*, and *MYC-4* when compared to chemical shifts of parent *MYC* are plotted in Figure 6. From the observed minor differences, a close conformational relationship is reasserted for *MYC* and *MYC-3a*. On the other hand, both *MYC-3s* and *MYC-4* exhibit significant but nearly identical deviations in line with their common conformational features. Thus, in addition to the upfield-shifted G8 H8 not present in *MYC-4*, G residues in the central G-quartet flanking the inverted all-*syn* 5′-tetrad of *MYC-3s* and *MYC-4* are more noticeably shifted. In contrast, protons of loop residues and of the 3′-tetrad are less affected, in agreement



DOI: 10.1021/acs.biochem.6b00925
*Biochemistry* 2016, 55, 6949−6955

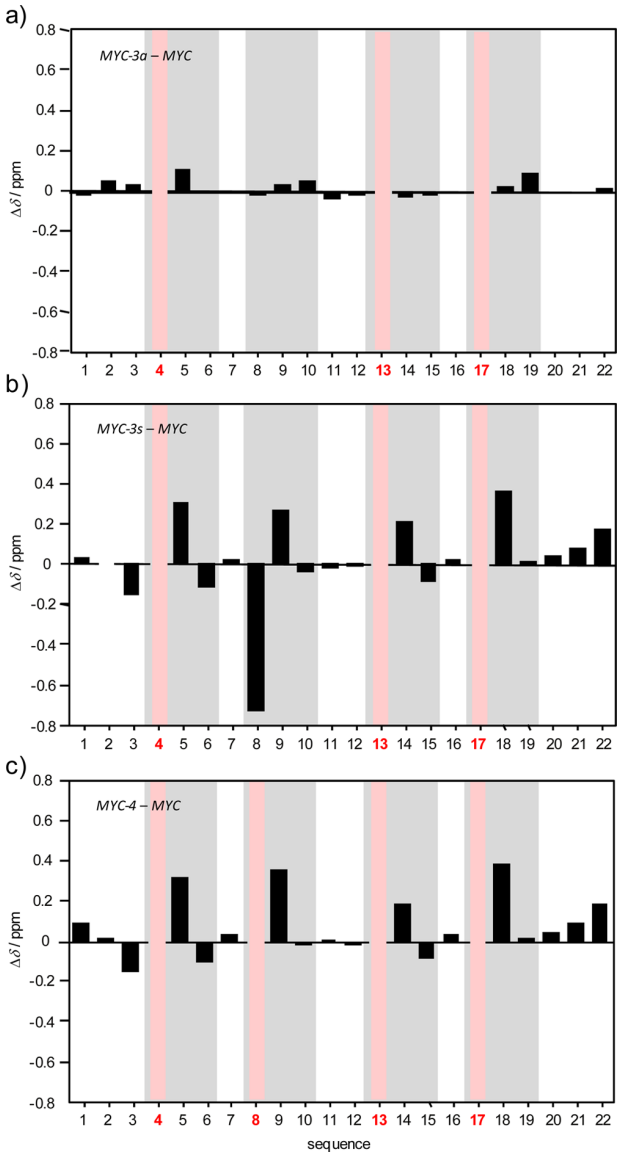



**Figure 6.** Comparison of H6/H8 chemical shifts of modified and native *MYC* quadruplex. The bar graphs show differences between (a) *MYC-3a*, (b) *MYC-3s*, or (c) *MYC-4* and unmodified *MYC* for each G4 residue at 30 °C. Positions with missing data due to the incorporation of 8-$^{Br}$G analogues are colored red.

with a conserved overall fold of the parallel-stranded quadruplexes.

**Kinetics of the Tetrad Polarity Switch.** The presence of two sets of signals in the *MYC-3* NMR spectra demonstrates the coexistence of *syn-anti* glycosidic isomers with different 5′-tetrad polarity. A dynamic interconversion of the two structures becomes evident via the observation of exchange crosspeaks in NOESY and ROESY spectra (Figure 4). It should be mentioned that noticeable exchange crosspeaks start to develop only at temperatures above 40 °C, and contacts solely due to dipolar coupling are detected at lower temperatures. Clearly, the observation of slowly exchanging individual resonances of the two species already places an upper limit on the exchange rate ($k_{ex}$) of approximately <10 $s^{-1}$.

To determine the dynamics of the tetrad polarity switch, we recorded NOESY spectra at 50 °C with eight different mixing times, selecting the G8 H8 resonance as a probe for the kinetic analysis (Figure 7a). A fit of the resulting build-up data of

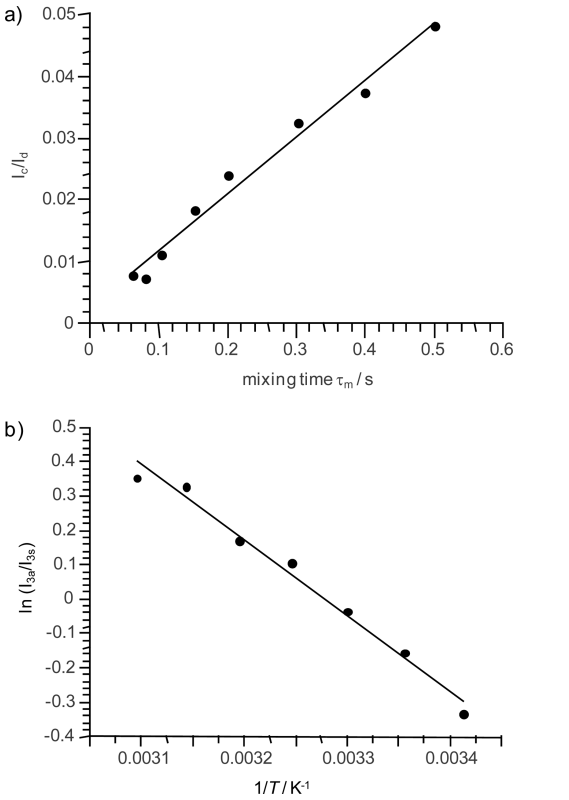



**Figure 7.** (a) Plot of the intensity ratio for exchange crosspeaks and corresponding diagonal peaks as a function of mixing time $\tau_m$ and best fit line. Integrals were extracted for G8 H8 in 2D NOE spectra at 50 °C. (b) van't Hoff plot constructed from the ratio of signal intensity of the A12 H2 resonance in *MYC-3a* and *MYC-3s* as a function of temperature and best fit line.

exchange peak intensities yields an exchange rate ($k_{ex}$) of 0.2 $s^{-1}$ for interconversion of *MYC-3s* and *MYC-3a* at 50 °C. Similar exchange rates of 0.2−0.3 $s^{-1}$ were obtained from H8 crosspeak intensities of G9 and A22 located in the central tetrad and the 3′-overhang, respectively. Because the tetrad flip involves four concerted *anti* → *syn* glycosidic torsion angle transitions, hydrogen bonds of the 5′-outer tetrad must be broken, but additional disruption of the central tetrad or even a complete dissociation of the whole quadruplex core cannot be excluded. However, with an enthalpy of formation of ∼80 kJ/mol per G-tetrad as estimated by previous calorimetric studies of various quadruplexes,[22,23] full dissociation of the three-tetrad quadruplex seems unlikely given the relatively fast kinetics of interconversion. Accordingly, the complete refolding associated with a topological transition between quadruplexes formed by a human telomeric sequence under molecular crowding conditions was found to proceed in ∼1 h with corresponding activation enthalpies determined to be 136 kJ/mol.[24] Unfortunately, exchange rates at different temperatures to yield activation energies are not accessible experimentally in this case because exchange peaks are hardly observed at ≤40 °C due to long lifetimes of exchanging species with respect to the quadruplex



DOI: 10.1021/acs.biochem.6b00925
*Biochemistry* 2016, 55, 6949−6955

relaxation time. Also, dynamic studies at higher temperatures are hampered by the onset of quadruplex melting at ≥60 °C.

**Relative Stability of *MYC* Conformational Isomers.** Knowing relative stabilities of homopolar and heteropolar G-tetrad arrangements would allow us to acquire deeper insight into forces that govern individual quadruplex folds. Unfortunately, an experimental evaluation of such energy differences in intramolecular G4 structures must fail in the case of different topologies associated with a complex superposition of various contributions to energy. However, the G4 conformational isomers *MYC-3a* and *MYC-3s* differ only in their glycosidic pattern and associated tetrad polarity, offering the unique opportunity to evaluate relative energies free of different strand alignments, loop arrangements, and numbers or types of modifications. Because of its good spectral resolution, the A12 H2 resonance was selected for a quantification of enthalpy changes associated with the dynamic transition between the two exchanging isomers. Having determined populations of the two species between 10 and 50 °C, we subjected the temperature-dependent equilibrium constants to a van't Hoff analysis based on a two-state process to yield a temperature-independent enthalpy change upon the *MYC-3a* to *MYC-3s* transition ($\Delta H^{\circ}$) of 18.2 kJ/mol (Figure 7b).

Because two neighboring guanosine residues in an *anti* conformation are expected to differ in stacking interactions when compared to a 5′-*syn-anti* step, stacking energies are intimately connected with guanosine glycosidic bond angles along the G-tracts. Despite their different conformational pattern, differences in enthalpy and free energy of the two trisubstituted G4 isomers are small and in line with a single melting event observed for the *MYC-3* mixture. These results are also in general accord with theoretical calculations that have been employed in trying to evaluate relative energies of different tetrad arrangements in quadruplexes.[25−27] Thus, a preference for *syn-anti* steps has been reported in molecular mechanics-based computations on simplified two-quartet models.[25] However, more recent quantum chemical approaches have suggested that *anti-anti* GpG constitutes the most favored dinucleotide step closely followed by *syn-anti* arrangements.[26]

Free 5′-GMP is known to be exceptional among the natural pyrimidine and purine nucleotides in quite easily adopting a *syn* glycosidic torsion angle.[28] On the other hand, experimental evidence points to a stronger propensity for the $^{Br}$G analogue with its sterically demanding bromo substituent to adopt *syn* conformations.[29] Clearly, it is a favored $^{Br}$G *syn* glycosidic torsion angle that is expected to induce a tetrad polarity switch with an unusual all-*syn* G-quartet. The substitution of 8-bromodeoxyguanosine in G4 structures at a matching position with a *syn* conformation has often been exploited for the stabilization of conformers, reducing heterogeneity and increasing their thermal stability.[30−32] However, although all $^{Br}$G residues in *MYC-3s* and *MYC-4* are in a favored *syn* conformation, thermal destabilization of these quadruplexes is observed upon comparison to unmodified *MYC* (see Table 1). Obviously, changing *anti*-G to unfavorable *anti*-$^{Br}$G in *MYC-3a* or to *syn*-$^{Br}$G in *MYC-3s* is invariably accompanied by an energetic penalty. Unfavorable contributions in *MYC-3s* may involve poorer base stacking interactions, steric effects of the bulky bromine substituent, and noncanonical backbone conformations as imposed by the *syn* glycosidic torsion angle.

## CONCLUSIONS

*syn-anti* conformational changes upon incorporation of guanosine analogues with sugar or base modifications into a G-quadruplex have frequently been reported in the past. Because of their different propensities to adopt a *syn* or *anti* glycosidic torsion angle, modified G nucleotides have been able to either stabilize or completely refold native quadruplex topologies. However, only recently was a complete G-tetrad flip observed through suitable modifications of physiologically relevant monomolecular quadruplexes without affecting their overall topology. By enforcing concerted *anti* → *syn* transitions through the incorporation of 8-bromoguanosine analogues in this study, we were able to allow a complete switch of a G-tetrad polarity even in case of the intramolecular *MYC* quadruplex with its very robust conformation. To the best of our knowledge, the resulting structure features a unique conformation in representing a unimolecular parallel-stranded quadruplex with overhang residues and an all-*syn* G-quartet. In addition to its importance for a better understanding of forces that govern quadruplex folding and topology, the ability to fine-tune the G4 conformation may expand the quadruplex conformational space and may lead to new structural motifs for various practical applications. Thus, on the basis of differential binding affinities for outer stacking ligands, a tetrad polarity switch may potentially be useful for the future design of novel catalytic or aptameric quadruplex scaffolds that rely on selective substrate or ligand binding.

For a $^{Br}$G-trisubstituted *MYC* quadruplex, only a partial conformational transition was observed and approximately equal amounts of two stable conformational isomers of different tetrad polarity but with the same overall fold coexist in solution. Interestingly, a dynamic interconversion of the two tetrad polarity isomers could be observed for the first time on the NMR time scale and attests to a relatively low energy barrier for such a G-tetrad flip. Although the detailed mechanism of interconversion is unknown, it may be suggested that thermal breathing motions will disrupt the outer G-tetrad to allow for a concerted *anti* → *syn* change of all four G residues. Consequently, such a transition is expected not to occur only on the modified quadruplex but may be a more general phenomenon of G4 structures. On the other hand, given large differences in the free energy of corresponding conformational isomers bearing no adequate modifications, the dynamics of interconversion will remain unnoticed in case of only one observable species.

## ASSOCIATED CONTENT

### Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.biochem.6b00925.

NMR resonance assignment strategy and kinetic analysis, $^{1}$H and C6/C8 $^{13}$C chemical shifts of *MYC* quadruplexes (Tables S1−S4), superposition of 2D NOE spectral regions of *MYC*, *MYC-3*, and *MYC-4* (Figures S1 and S4), and 2D NOE spectral regions of *MYC-3* (Figures S2 and S3) and *MYC-4* (Figures S5 and S6) (PDF)

## AUTHOR INFORMATION

### Corresponding Author

*E-mail: weisz@uni-greifswald.de. Fax: (+49) 3834 86-4427. Phone: (+49) 3834 86-4426.



DOI: 10.1021/acs.biochem.6b00925
*Biochemistry* 2016, 55, 6949−6955

**ORCID**

Klaus Weisz: 0000-0002-2736-6606

**Funding**

B.K. and A.F. were supported by a State Postgraduate Scholarship.

**Notes**

The authors declare no competing financial interest.

## ■ ABBREVIATIONS

DSC, differential scanning calorimetry; NOE, nuclear Overhauser effect; ROESY, rotating frame Overhauser effect spectroscopy; HSQC, heteronuclear single-quantum correlation; DQF-COSY, double-quantum-filtered correlated spectroscopy.

## ■ REFERENCES

(1) Burge, S., Parkinson, G. N., Hazel, P., Todd, A. K., and Neidle, S. (2006) Quadruplex DNA: sequence, topology and structure. *Nucleic Acids Res. 34*, 5402−5415.

(2) Monchaud, D., and Teulade-Fichou, M.-P. (2008) A hitchhiker's guide to G-quadruplex ligands. *Org. Biomol. Chem. 6*, 627−636.

(3) Neo, J. L., Kamaladasan, K., and Uttamchandani, M. (2012) G-Quadruplex based probes for visual detection and sensing. *Curr. Pharm. Des. 18*, 2048−2057.

(4) Ge, B., Huang, Y. C., Sen, D., and Yu, H. Z. (2010) A robust electronic switch made of immobilized duplex/quadruplex DNA. *Angew. Chem., Int. Ed. 49*, 9965−9967.

(5) Phan, A. T., Kuryavyi, V., Luu, K. N., and Patel, D. J. (2006) in *Quadruplex Nucleic Acids* (Balasubramanian, S., and Neidle, S., Eds.) Chapter 3, Royal Society of Chemistry, London.

(6) Martín-Pintado, N., Yahyaee-Anzahaee, M., Deleavey, G. F., Portella, G., Orozco, M., Damha, M. J., and González, C. (2013) Dramatic effect of furanose C2′ substitution on structure and stability: directing the folding of the human telomeric quadruplex with a single fluorine atom. *J. Am. Chem. Soc. 135*, 5344−5347.

(7) Li, Z., Lech, C. J., and Phan, A. T. (2014) Sugar-modified G-quadruplexes: effects of LNA-, 2′F-RNA- and 2′F-ANA-guanosine chemistries on G-quadruplex structure and stability. *Nucleic Acids Res. 42*, 4068−4079.

(8) Esposito, V., Randazzo, A., Piccialli, G., Petraccone, L., Giancola, C., and Mayol, L. (2004) Effects of an 8-bromodeoxyguanosine incorporation on the parallel quadruplex structure $[d(TGGGT)]_4$. *Org. Biomol. Chem. 2*, 313−318.

(9) Virgilio, A., Esposito, V., Randazzo, A., Mayol, L., and Galeone, A. (2005) 8-Methyl-2′-deoxyguanosine incorporation into parallel DNA quadruplex structures. *Nucleic Acids Res. 33*, 6188−6195.

(10) Thao Tran, P. L., Virgilio, A., Esposito, V., Citarella, G., Mergny, J.-L., and Galeone, A. (2011) Effects of 8-methylguanine on structure, stability and kinetics of formation of tetramolecular quadruplexes. *Biochimie 93*, 399−408.

(11) Masiero, S., Trotta, R., Pieraccini, S., De Tito, S., Perone, R., Randazzo, A., and Spada, G. P. (2010) A non-empirical chromophoric interpretation of CD spectra of DNA G-quadruplex structures. *Org. Biomol. Chem. 8*, 2683−2692.

(12) Dickerhoff, J., and Weisz, K. (2015) Flipping a G-tetrad in a unimolecular quadruplex without affecting its global fold. *Angew. Chem., Int. Ed. 54*, 5588−5591.

(13) Cheong, V. V., Lech, C. J., Heddi, B., and Phan, A. T. (2016) Inverting the G-tetrad polarity of a G-quadruplex by using xanthine and 8-oxoguanine. *Angew. Chem., Int. Ed. 55*, 160−163.

(14) Ambrus, A., Chen, D., Dai, J., Jones, R. A., and Yang, D. (2005) Solution structure of the biologically relevant G-quadruplex element in the human c-MYC promoter. Implications for G-quadruplex stabilization. *Biochemistry 44*, 2048−2058.

(15) Šket, P., Virgilio, A., Esposito, V., Galeone, A., and Plavec, J. (2012) Strand directionality affects cation binding and movement within tetramolecular G-quadruplexes. *Nucleic Acids Res. 40*, 11047−11057.

(16) Vranken, W. F., Boucher, W., Stevens, T. J., Fogh, R. H., Pajon, A., Llinas, M., Ulrich, E. L., Markley, J. L., Ionides, J., and Laue, E. D. (2005) The CCPN data model for NMR spectroscopy: development of a software pipeline. *Proteins: Struct., Funct., Genet. 59*, 687−696.

(17) Thiele, C. M., Petzold, K., and Schleucher, J. (2009) EASY ROESY: reliable cross-peak integration in adiabatic symmetrized ROESY. *Chem. - Eur. J. 15*, 585−588.

(18) Feigon, J., Koshlap, K. M., and Smith, F. W. (1995) $^1$H NMR spectroscopy of DNA triplexes and quadruplexes. In *Methods in Enzymology* (James, T. L., Ed.) Vol. *261*, pp 225−256, Academic Press, San Diego.

(19) Fonville, J. M., Swart, M., Vokáčová, Z., Sychrovský, V., Šponer, J. E., Šponer, J., Hilbers, C. W., Bickelhaupt, F. M., and Wijmenga, S. S. (2012) Chemical shifts in nucleic acids studied by density functional theory calculations and comparison with experiment. *Chem. - Eur. J. 18*, 12372−12387.

(20) Greene, K. L., Wang, Y., and Live, D. (1995) Influence of the glycosidic torsion angle on $^{13}$C and $^{15}$N shifts in guanosine nucleotides: investigations of G-tetrad models with alternating *syn* and *anti* bases. *J. Biomol. NMR 5*, 333−338.

(21) Wijmenga, S. S., and van Buuren, B. N. M. (1998) The use of NMR methods for conformational studies of nucleic acids. *Prog. Nucl. Magn. Reson. Spectrosc. 32*, 287−387.

(22) Jin, R., Gaffney, B. L., Wang, C., Jones, R. A., and Breslauer, K. J. (1992) Thermodynamics and structure of a DNA tetraplex: A spectroscopic and calorimetric study of the tetramolecular complexes of $d(TG_3T)$ and $d(TG_3T_2G_3T)$. *Proc. Natl. Acad. Sci. U. S. A. 89*, 8832−8836.

(23) Petraccone, L., Erra, E., Esposito, V., Randazzo, A., Mayol, L., Nasti, L., Barone, G., and Giancola, C. (2004) Stability and structure of telomeric DNA sequences forming quadruplexes containing four G-tetrads with different topological arrangements. *Biochemistry 43*, 4877−4884.

(24) Heddi, B., and Phan, A. T. (2011) Structure of human telomeric DNA in crowded solution. *J. Am. Chem. Soc. 133*, 9824−9833.

(25) Cang, X., Šponer, J., and Cheatham, T. E., III (2011) Explaining the varied glycosidic conformational, G-tract length and sequence preferences for anti-parallel G-quadruplexes. *Nucleic Acids Res. 39*, 4499−4512.

(26) Šponer, J., Mládek, A., Špačková, N., Cang, X., Cheatham, T. E., III, and Grimme, S. (2013) Relative stability of different DNA guanine quadruplex stem topologies derived using large-scale quantum-chemical computations. *J. Am. Chem. Soc. 135*, 9785−9796.

(27) Islam, B., Stadlbauer, P., Neidle, S., Haider, S., and Šponer, J. (2016) Can we execute reliable MM-PBSA free energy computations of relative stabilities of different guanine quadruplex folds? *J. Phys. Chem. B 120*, 2899−2912.

(28) Son, T.-D., Guschlbauer, W., and Guéron, M. (1972) Flexibility and conformations of guanosine monophosphates by the Overhauser effect. *J. Am. Chem. Soc. 94*, 7903−7911.

(29) Tavale, S. S., and Sobell, H. M. (1970) Crystal and molecular structure of 8-bromoguanosine and 8-bromoadenosine, two purine nucleosides in the *syn* conformation. *J. Mol. Biol. 48*, 109−123.

(30) Dias, E., Battiste, J. L., and Williamson, J. R. (1994) Chemical probe for glycosidic conformation in telomeric DNAs. *J. Am. Chem. Soc. 116*, 4479−4480.

(31) Lim, K. W., Ng, V. C. M., Martín-Pintado, N., Heddi, B., and Phan, A. T. (2013) Structure of the human telomere in $Na^+$ solution: an antiparallel (2 + 2) G-quadruplex scaffold reveals additional diversity. *Nucleic Acids Res. 41*, 10556−10562.

(32) Brčić, J., and Plavec, J. (2015) Solution structure of a DNA quadruplex containing ALS and FTD related GGGGCC repeat stabilized by 8-bromodeoxyguanosine substitution. *Nucleic Acids Res. 43*, 8590−8600.



DOI: 10.1021/acs.biochem.6b00925
*Biochemistry* 2016, 55, 6949−6955

# Supporting Information

## Observation of a Dynamic G-Tetrad Flip in Intramolecular G-Quadruplexes

Beatrice Karg, Linn Haase, Andrea Funke, Jonathan Dickerhoff, and Klaus Weisz*

Institute of Biochemistry, Ernst-Moritz-Arndt University Greifswald, Felix-Hausdorff-Str. 4, D-17487 Greifswald, Germany

*Corresponding author: weisz@uni-greifswald.de

## NMR resonance assignment strategies

***MYC-3* quadruplexes.** Despite a more complex appearance due to two coexisting species, 2D NOE spectra of *MYC-3* allowed to trace separate continuous NOE connectivities. In case of more severe peak overlap, connectivities were independently followed in spectra acquired at various temperatures between 10 and 50 °C to benefit from different temperature dependent shifts for signal separation. Thus, two independent sets of four uninterrupted intranucleotide and sequential H8-H1' contacts could be detected along the G-tracts with partial extension into the 3'-overhang (Figure S2). Whereas three NOE walks for each species only include two H8 base protons of guanines in *anti* conformation, two sets of a longer NOE run comprising another guanine H8 must be attributed to the G8-G9-G10 tract bearing no 8-Br modification at its 5'-end.

Exploiting the close similarities in most chemical shifts to parent *MYC*, resonances could be specifically assigned to *MYC-3a* (see also Figure S1). The continuous H8-H1' NOE connectivities along G8-G9-G10 indicates an all-*anti* G-tract in line with a conserved conformation (Figure S2). On the other hand, a very strong H8-H1' intranucleotide crosspeak found for the second more upfield shifted G8 H8 strongly points to a *syn* glycosidic torsion angle for *MYC-3s* and further confirms results from C8 chemical shifts as extracted from $^{1}$H-$^{13}$C HSQC experiments. Based on parallel G-tracts, a *syn* conformation for G8 must be associated with corresponding transitions of the $^{Br}$G analogs to result in a complete flip of the 5'-quartet.

All labile G H1 imino protons are connected to iminos in adjacent tetrads. Assignments are based on the observation, that in an all-*anti* guanosine arrangement imino-imino NOEs preferably run along individual G-tracts whereas in case of a *syn-anti* step shorter contacts only occur between imino protons positioned in neighboring tracts (Figure S3). With G H8 resonances fully assigned, the identification of imino protons is mostly completed by observed imino-H8 NOE contacts in the central and 3'-tetrads with their conserved polarity. A single G13 H1 to G8 H8 contact observed for *MYC-3s* matches a modified 5'-outer tetrad with opposite polarity (Figure S3). Additional NOE contacts between G4 H1 and A3 protons of the 5'-overhang are again in line with the resonance assignments for *MYC-3s*.

***MYC-4* quadruplex.** The assignment of *MYC-4* resonances is mostly based on the analysis of 2D NOE spectra and additionally supported by the close similarity to *MYC-3s* (see Figure S4). Starting at the 3'-terminus, uninterrupted 3'→5' intranucleotide and sequential H6/H8-H1' NOE

contacts allow assignment of all H6/H8 and H1' protons along the overhang sequence and the following run of guanosines (Figure S5a). The continuous walk ends with $^{Br}$G at position 17 lacking a corresponding H8 proton. Interrupted by the Br-modified guanines of the 5'-tetrad, corresponding walks are also observed along the other G-tracts. Likewise, H6/H8-H2' and H6/H8-H3' intranucleotide and sequential connectivities allow the sequential assignment of H2' and H3' protons and confirm previous assignments (not shown). The glycosidic torsion angles for all guanine nucleotides within the central and 3'-tetrad are in the *anti* range as clearly indicated by H8-H1' intra- and interresidual NOEs of similar intensity.[1] In line with two stacked all-*anti* G-quartets bearing no modifications, four sequential H8-H8 NOE contacts are observed between neighboring guanines within the four G-runs and continue into the 3'-overhang for the fourth G-tract (Figure S5b).

The unambiguous alignment of G-tracts benefits from spectral comparisons with the related quadruplexes. Additional support comes from NOE crosspeaks observed between nonexchangeable base and sugar protons of loop residue A12 and residues G10 and G14 within the second and third G-tract (see Figure S5). Given an *anti* conformation for all non-modified G residues, the assignment of guanine imino protons is based on four uninterrupted networks with two interresidual H1-H1 NOE contacts as expected for three stacked G-tetrads together with weak H1-H8 connectivities observed within and between the central and 3'-tetrad and between $^{Br}$G4 H1 and A3 base protons in the 5'-overhang (Figure S6). Lacking H8, assignments of the $^{Br}$G imino protons of the 5'-tetrad are based on their H1-H1 contacts to central guanines of the neighboring G-tract as expected for $^{Br}$G nucleotides in a *syn* glycosidic conformation (Figure S6b, see also Figure S3a).

**References**

(1) Feigon, J., Koshlap, K. M., and Smith, F. W. (1995) $^{1}$H NMR spectroscopy of DNA triplexes and quadruplexes. *Methods Enzymol*. (James, T. L., Ed.) *261*, 225-256.

## Kinetic analysis

Intensities of NOE crosspeaks and corresponding diagonal peaks $I_c$ and $I_d$ are given by

$$I_c = ¼ [(\exp(-\tau_m/T_1)][1-\exp(-k_{ex}\cdot\tau_m)] \quad (1)$$

$$I_d = ¼ [(\exp(-\tau_m/T_1)][1+\exp(-k_{ex}\cdot\tau_m)] \quad (2)$$

with $\tau_m$, $T_1$, and $k_{ex}$ being the mixing time, the spin-lattice relaxation time, and the exchange rate constant, respectively. Consequently

$$I_c/I_d = [1-\exp(-k_{ex}\cdot\tau_m)]/[1+\exp(-k_{ex}\cdot\tau_m)] \quad (3)$$

At the limits of $k_{ex}\cdot\tau_m << 1$, $I_c/I_d$ shows a linear dependence on $\tau_m$

$$I_c/I_d \simeq ½(k_{ex}\cdot\tau_m) \quad (4)$$

**Table S1.** $^{1}$H and $^{13}$C chemical shifts $\delta$ (in ppm) of protons and C6/C8 in *MYC*[a]

| | imino | H6/H8 | H2/Me | H1‘ | H2‘1 | H2‘2 | C6/C8 |
|---|---|---|---|---|---|---|---|
| T1 | n.d. | 7.20 | 1.64 | 5.78 | 1.59 | 2.04 | 136.84 |
| G2 | n.d. | 7.62 | --- | 5.62 | 2.38 | 2.46 | 136.63 |
| A3 | --- | 7.99 | 7.76 | 5.87 | 2.44 | 2.59 | 138.74 |
| G4 | 11.71 | 7.99 | --- | 6.04 | 2.74 | 2.99 | 135.28 |
| G5 | 11.20 | 7.70 | --- | 6.12 | 2.59 | 2.89 | 134.63 |
| G6 | 10.60 | 7.73 | --- | 6.39 | 2.59 | 2.76 | 135.07 |
| T7 | n.d. | 7.85 | 1.97 | 6.50 | 2.45 | 2.65 | 137.20 |
| G8 | 11.68 | 7.97 | --- | 6.12 | 2.45 | 2.89 | 135.20 |
| G9 | 11.48 | 7.88 | --- | 6.12 | 2.63 | 2.83 | 135.54 |
| G10 | 11.04 | 7.84 | --- | 6.41 | 2.58 | 2.74 | 135.10 |
| T11 | n.d. | 7.63 | 1.92 | 6.21 | 2.20 | 2.44 | 137.32 |
| A12 | --- | 8.51 | 8.35 | 6.67 | 2.95 | 3.07 | 140.76 |
| G13 | 11.89 | 8.09 | --- | 6.15 | 2.61 | 2.95 | 135.73 |
| G14 | 11.23 | 7.78 | --- | 6.18 | 2.68 | 2.95 | 135.00 |
| G15 | 11.01 | 7.78 | --- | 6.43 | 2.60 | 2.71 | 135.00 |
| T16 | n.d. | 7.85 | 1.97 | 6.50 | 2.45 | 2.65 | 137.20 |
| G17 | 11.24 | 7.88 | --- | 5.98 | 2.33 | 2.78 | 135.06 |
| G18 | 11.33 | 7.88 | --- | 6.02 | 2.65 | 2.71 | 135.42 |
| G19 | 11.02 | 7.59 | --- | 6.15 | 2.52 | 2.79 | 134.24 |
| T20 | n.d. | 7.13 | 1.48 | 5.89 | 1.90 | 2.36 | 135.61 |
| A21 | --- | 7.76 | 7.10 | 5.75 | 1.99 | 2.44 | 138.38 |
| A22 | --- | 7.50 | 7.37 | 5.60 | 2.19 | 2.36 | 137.38 |

[a]At 30 °C in 90% $H_2O$/10% $D_2O$, 10 mM potassium phosphate buffer, pH 7.0; n.d. = not determined.

**Table S2.** $^{1}$H and $^{13}$C chemical shifts $\delta$ (in ppm) of protons and C6/C8 in *MYC-3a*[a]

| | imino | H6/H8 | H2/Me | H1‘ | H2‘1 | H2‘2 | C6/C8 |
|---|---|---|---|---|---|---|---|
| T1 | n.d. | 7.19 | 1.64 | 5.71 | 1.69 | 2.13 | 136.70 |
| G2 | n.d. | 7.67 | --- | 5.79 | 2.17 | 2.43 | 136.44 |
| A3 | --- | 8.02 | 7.87 | 6.16 | 2.49 | 2.77 | 138.69 |
| $^{Br}$G4 | 11.93 | --- | --- | 6.21 | 2.89 | 3.29 | n.d. |
| G5 | 11.13 | 7.81 | --- | 6.13 | 2.63 | 2.91 | 135.10 |
| G6 | 10.59 | 7.73 | --- | 6.38 | 2.57 | 2.74 | 134.99 |
| T7 | n.d. | 7.85 | 1.98 | 6.50 | 2.45 | 2.66 | 137.33 |
| G8 | 11.26 | 7.95 | --- | 6.14 | 2.50 | 2.86 | 135.29 |
| G9 | 11.41 | 7.91 | --- | 6.23 | 2.66 | 2.84 | 135.72 |
| G10 | 11.04 | 7.89 | --- | 6.46 | 2.59 | 2.73 | 135.10 |
| T11 | n.d. | 7.59 | 1.90 | 6.19 | 2.19 | 2.41 | 137.26 |
| A12 | --- | 8.49 | 8.40 | 6.70 | 3.00 | 3.00 | 140.58 |
| $^{Br}$G13 | 11.35 | --- | --- | 6.26 | 2.65 | 2.94 | n.d. |
| G14 | 11.21 | 7.75 | --- | 6.20 | 2.63 | 2.92 | n.d. |
| G15 | 10.93 | 7.76 | --- | 6.42 | 2.56 | 2.69 | 135.05 |
| T16 | n.d. | 7.85 | 1.98 | 6.50 | 2.45 | 2.66 | 137.33 |
| $^{Br}$G17 | 11.20 | --- | --- | 6.02 | 2.49 | 2.83 | n.d. |
| G18 | 11.35 | 7.90 | --- | 6.09 | 2.72 | 2.72 | n.d. |
| G19 | 11.05 | 7.68 | --- | 6.18 | 2.53 | 2.79 | 134.30 |
| T20 | n.d. | 7.13 | 1.53 | 5.87 | 1.87 | 2.33 | 135.67 |
| A21 | --- | 7.76 | n.d. | 5.72 | 2.02 | 2.40 | 138.45 |
| A22 | --- | 7.51 | n.d. | 5.64 | 2.21 | 2.33 | 137.56 |

[a] At 30 °C in 90% $H_2O$/10% $D_2O$, 10 mM potassium phosphate buffer, pH 7.0; n.d. = not determined.

**Table S3.** $^{1}$H and $^{13}$C chemical shifts $\delta$ (in ppm) of protons and C6/C8 in *MYC-3s*[a]

| | imino | H6/H8 | H2/Me | H1‘ | H2‘1 | H2‘2 | C6/C8 |
|---|---|---|---|---|---|---|---|
| T1 | n.d. | 7.23 | 1.66 | 5.85 | 1.67 | 2.05 | 136.76 |
| G2 | n.d. | n.d. | --- | n.d. | n.d. | n.d. | n.d. |
| A3 | --- | 7.84 | n.d. | 6.04 | 2.00 | 2.55 | 137.77 |
| $^{Br}$G4 | 11.38 | --- | --- | 6.19 | 3.05 | 3.93 | n.d. |
| G5 | 11.57 | 8.01 | --- | 5.96 | 2.44 | 2.88 | 135.70 |
| G6 | 10.73 | 7.62 | --- | 6.40 | 2.59 | 2.68 | 135.59 |
| T7 | n.d. | 7.87 | 1.98 | 6.50 | 2.40 | 2.76 | 137.33 |
| G8 | 11.25 | 7.24 | --- | 5.91 | 2.70 | 3.00 | 138.72 |
| G9 | 11.87 | 8.15 | --- | 5.90 | 2.66 | 2.75 | n.d. |
| G10 | 11.15 | 7.80 | --- | 6.48 | 2.66 | 2.71 | 135.10 |
| T11 | n.d. | 7.61 | 1.93 | 6.19 | 2.17 | 2.44 | 137.26 |
| A12 | --- | 8.50 | 8.31 | 6.66 | 2.92 | 3.06 | 140.58 |
| $^{Br}$G13 | 11.80 | --- | --- | 6.14 | 3.06 | 3.49 | n.d. |
| G14 | 11.45 | 7.99 | --- | 5.98 | 2.49 | 2.92 | 135.80 |
| G15 | 10.86 | 7.70 | --- | 6.48 | 2.45 | 2.67 | 135.41 |
| T16 | n.d. | 7.87 | 1.98 | 6.50 | 2.40 | 2.76 | 137.33 |
| $^{Br}$G17 | 11.13 | --- | --- | 5.92 | 2.68 | 2.80 | n.d. |
| G18 | 11.89 | 8.24 | --- | 5.75 | 2.63 | 2.63 | n.d. |
| G19 | 11.19 | 7.60 | --- | 6.19 | 2.49 | 2.75 | 134.31 |
| T20 | n.d. | 7.17 | 1.60 | 5.84 | 1.80 | 2.25 | 135.80 |
| A21 | --- | 7.84 | n.d. | 5.79 | 2.12 | 2.33 | 138.75 |
| A22 | --- | 7.67 | n.d. | 5.80 | 2.30 | 2.30 | 138.04 |

[a] At 30 °C in 90% $H_2O$/10% $D_2O$, 10 mM potassium phosphate buffer, pH 7.0; n.d. = not determined.

**Table S4.** $^{1}$H chemical shifts $\delta$ (in ppm) of protons in *MYC-4*[a]

| | imino | H6/H8 | H2/Me | H1‘ | H2‘1 | H2‘2 | H3‘ |
|---|---|---|---|---|---|---|---|
| T1 | n.d. | 7.29 | 1.67 | 5.85 | 1.79 | 2.12 | 4.45 |
| G2 | n.d. | 7.63 | --- | 5.85 | 2.43 | 2.54 | n.d. |
| A3 | --- | 7.85 | 7.55 | 6.05 | 1.96 | 2.53 | 4.68 |
| $^{Br}$G4 | 11.33 | --- | --- | 6.20 | 3.05 | 3.97 | 4.92 |
| G5 | 11.51 | 8.01 | --- | 5.97 | 2.44 | 2.90 | 5.00 |
| G6 | 10.71 | 7.64 | --- | 6.42 | 2.61 | 2.66 | 5.12 |
| T7 | n.d. | 7.88 | 1.98 | 6.50 | 2.39 | 2.77 | 5.09 |
| $^{Br}$G8 | 11.40 | --- | --- | 5.97 | 2.81 | 2.95 | 5.05 |
| G9 | 11.88 | 8.23 | --- | 5.85 | 2.60 | 2.78 | 5.05 |
| G10 | 11.48 | 7.83 | --- | 6.42 | 2.60 | 2.66 | 5.08 |
| T11 | n.d. | 7.63 | 1.94 | 6.21 | 2.20 | 2.45 | 5.09 |
| A12 | --- | 8.50 | 8.31 | 6.65 | 2.95 | 3.07 | 5.25 |
| $^{Br}$G13 | 11.16 | --- | --- | 6.16 | 3.09 | 3.56 | 5.05 |
| G14 | 11.46 | 7.96 | --- | 5.98 | 2.47 | 2.92 | 5.00 |
| G15 | 10.87 | 7.71 | --- | 6.48 | 2.68 | 2.68 | 5.13 |
| T16 | n.d. | 7.88 | 1.98 | 6.50 | 2.39 | 2.77 | 5.10 |
| $^{Br}$G17 | 11.06 | --- | --- | 5.95 | 2.73 | 2.82 | 5.05 |
| G18 | 11.89 | 8.26 | --- | 5.74 | 2.62 | 2.65 | 5.05 |
| G19 | 11.19 | 7.60 | --- | 6.19 | 2.48 | 2.76 | 4.93 |
| T20 | n.d. | 7.17 | 1.60 | 5.84 | 1.80 | 2.25 | 4.72 |
| A21 | --- | 7.85 | 7.37 | 5.80 | 2.12 | 2.34 | 4.61 |
| A22 | --- | 7.68 | 7.49 | 5.80 | 2.30 | 2.30 | 4.39 |

[a]At 30 °C in 90% $H_2O$/10% $D_2O$, 10 mM potassium phosphate buffer, pH 7.0; n.d. = not determined.

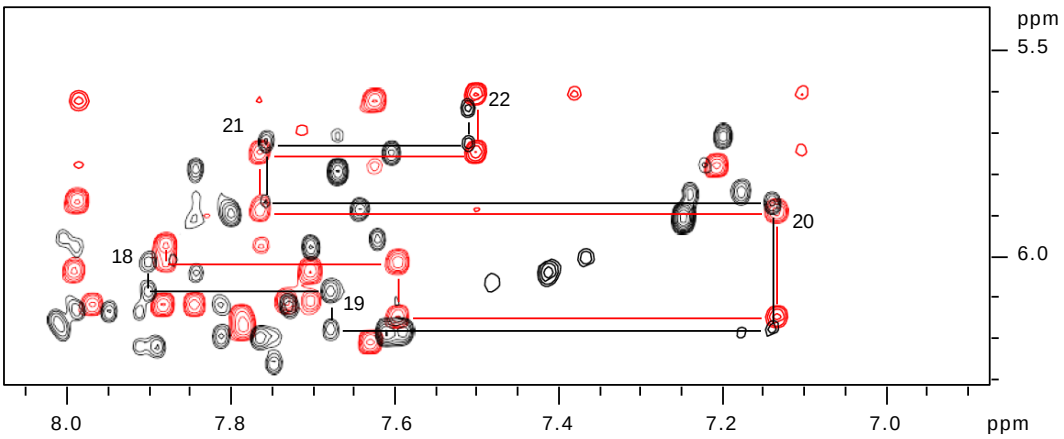



**Figure S1.** Superposition of H8/H6($\omega_2$)–H1'($\omega_1$) spectral regions from 2D NOE spectra of *MYC* (red) and *MYC-3* (black) acquired at 30 °C with $\tau_m$ = 300 ms. The same general pattern of uninterrupted NOE walks starting from the 3'-terminal A22 and ending for *MYC-3* at modified G17 are exemplarily indicated by red and black lines for the two quadruplexes; intranucleotide crosspeaks are labeled by residue numbers.

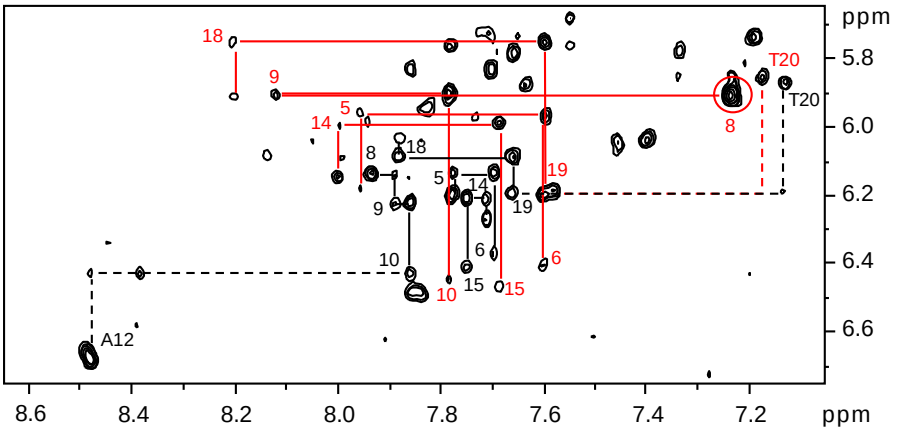



**Figure S2.** 2D NOE spectrum of *MYC-3* acquired at 40 °C with $\tau_m$ = 350 ms. H8/H6($\omega_2$)–H1'($\omega_1$) region tracing continuous NOE connectivities along the G-tracts of the coexisting 5'-*syn* (*MYC-3s*, red) and all-*anti* species (*MYC-3a*, black); connectivities to overhang and loop nucleotides are followed by dotted lines; intranucleotide crosspeaks are indicated by residue number with the corresponding G8 contact of *MYC-3s* marked by a circle.

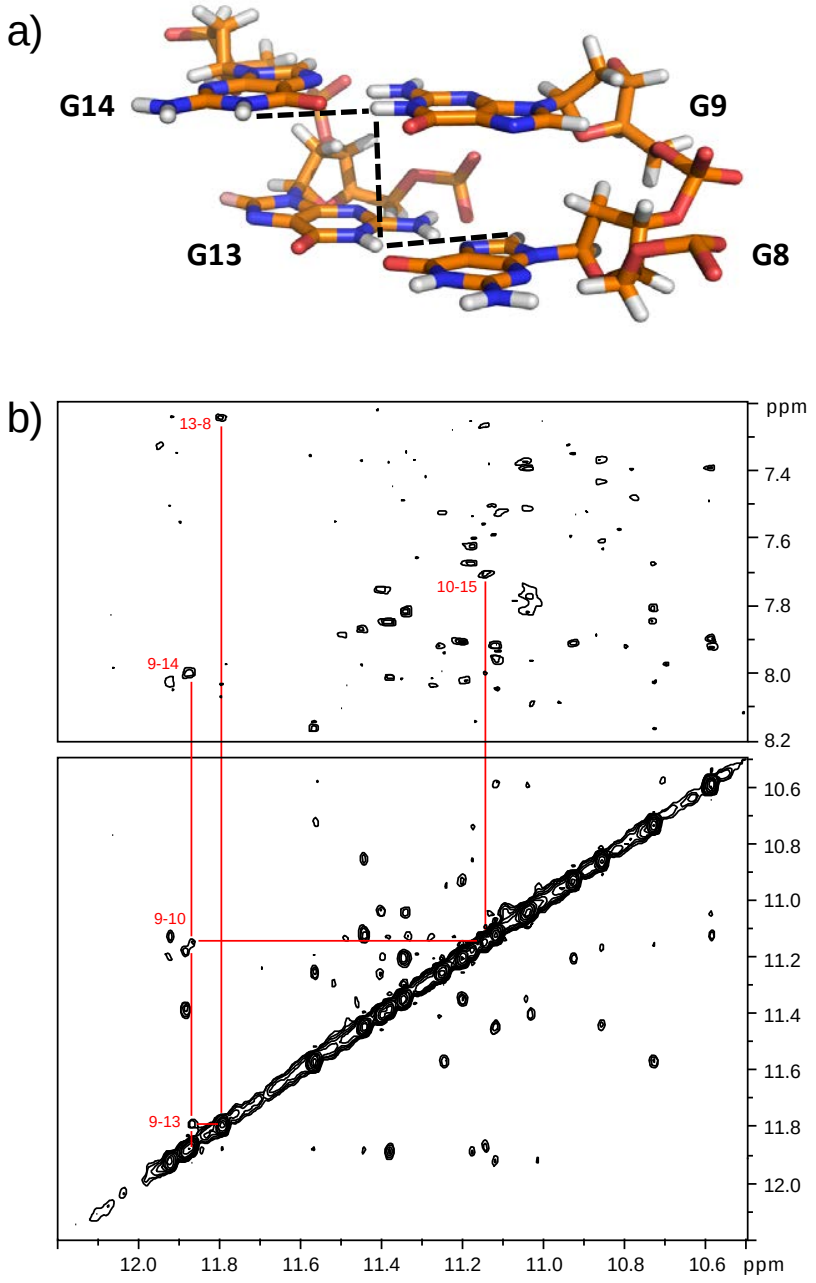



**Figure S3.** a) Structural model of a *MYC-3s* fragment with G8(*syn*)-G9(*anti*) and G13(*syn*)-G14(*anti*) dinucleotide steps; short internucleotide imino-imino and imino-H8 distances are indicated by broken lines; b) 2D NOE spectral regions showing imino-imino (bottom) and imino-H8 (top) NOE contacts for *MYC-3*. Sequential imino-imino connectivities are exemplarily indicated for G9 H1 and followed to intratetrad H1-H8 contacts for the three connected H1 protons; $T$ = 30 °C, $\tau_m$ = 600 ms.

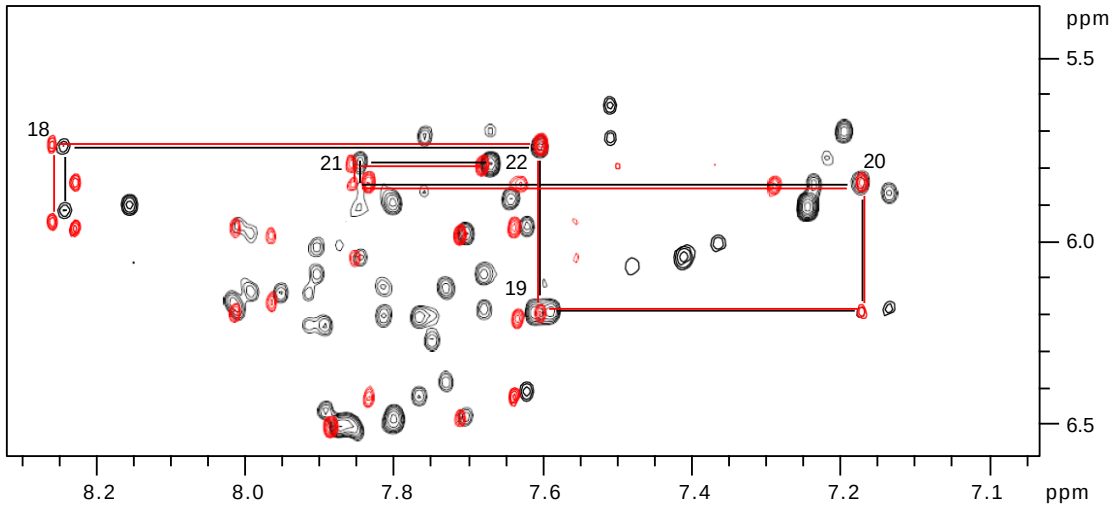



**Figure S4.** Superposition of H8/H6($\omega_2$)–H1'($\omega_1$) spectral regions from 2D NOE spectra of *MYC-4* (red) and *MYC-3* (black) acquired at 30 °C with $\tau_m$ = 300 ms. The same pattern of uninterrupted NOE walks starting from the 3'-terminal A22 and ending at modified G17 are exemplarily indicated by red and black lines for the two quadruplexes; intranucleotide crosspeaks are labeled by residue numbers.

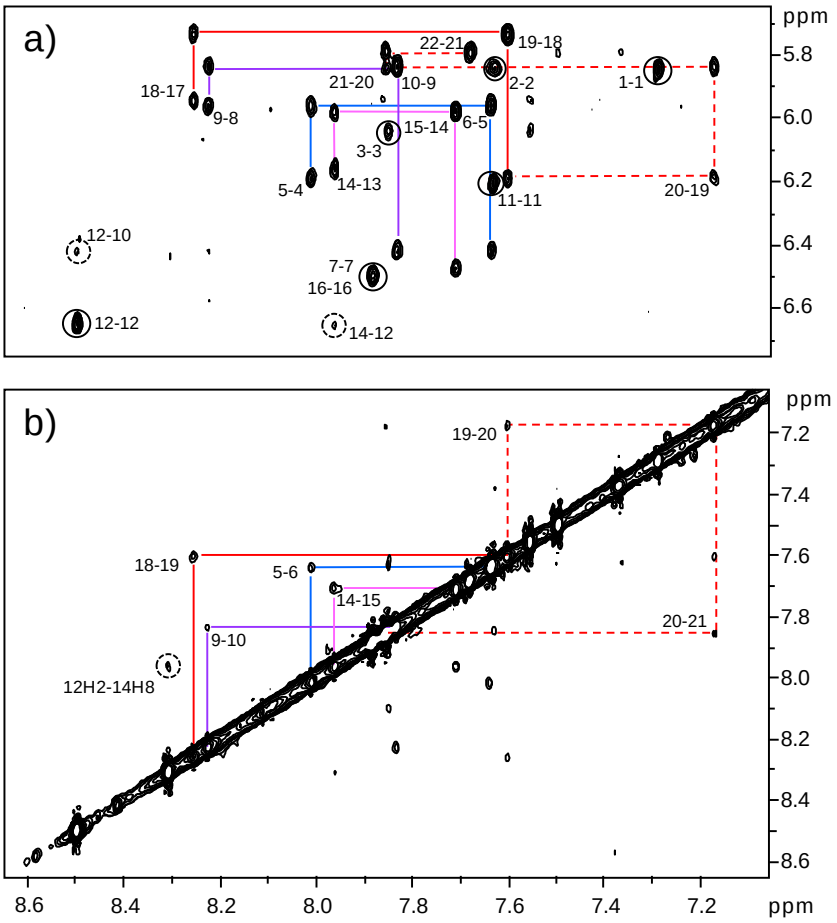



**Figure S5.** Portions of a 2D NOE spectrum of *MYC-4* acquired at 30 °C with $\tau_m$ = 300 ms. a) H8/H6($\omega_2$)–H1'($\omega_1$) region tracing continuous NOE connectivities along each of the four G-tracts and b) H8/H6($\omega_2$)–H8/H6($\omega_1$) region with sequential NOE contacts involving the unmodified G-tetrads. Connectivities extending into the 3'-overhang are traced by dotted lines; sequential crosspeaks are marked by residue number in $\omega_2$ and $\omega_1$ and non-sequential NOE contacts involving overhang and loop nucleotides are marked by circles.

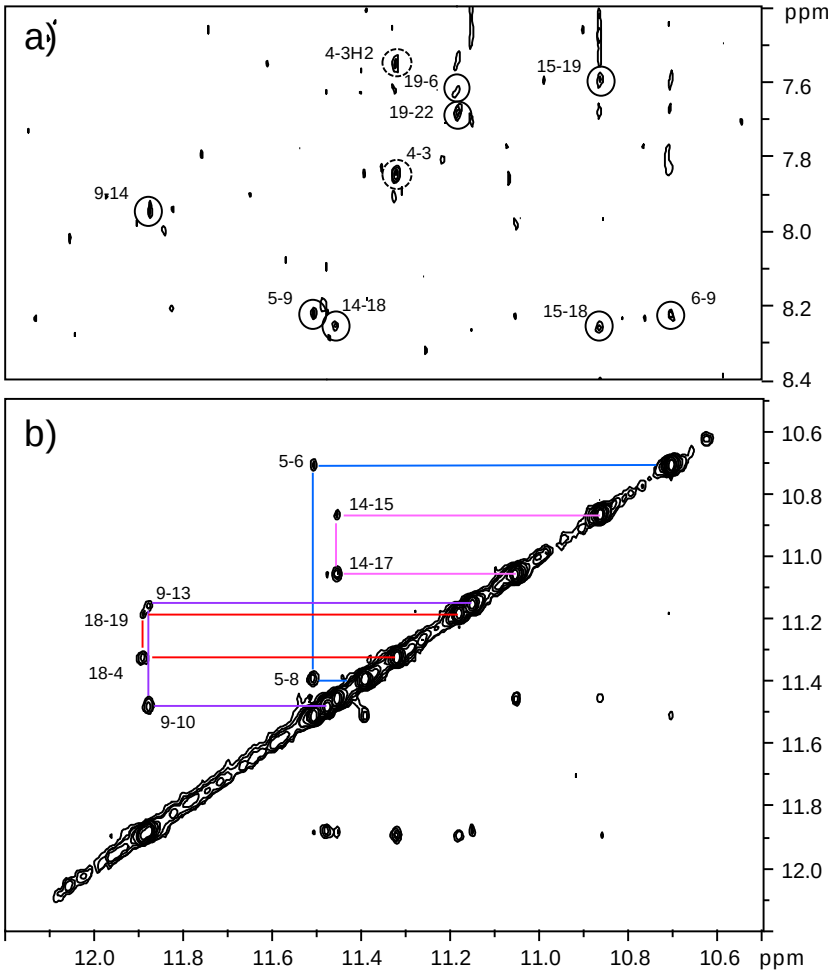



**Figure S6**. Portions of a 2D NOE spectrum of *MYC-4* acquired at 30 °C with $\tau_m$ = 300 ms. a) Spectral region showing H1($\omega_2$)–H8($\omega_1$) NOE contacts; two contacts between $^{Br}$G4 H1 and A3 H2 as well as A3 H8 are marked by dotted circles. b) Spectral region tracing four imino-imino sequential NOE walks along the stacked G-tetrads; crosspeaks are marked by residue number in $\omega_2$ and $\omega_1$.

# Publication II

ChemPubSoc Europe

DOI: 10.1002/chem.201801851

CHEMISTRY A European Journal Full Paper

Biochemistry | *Hot Paper*

# Loop Length Affects *Syn–Anti* Conformational Rearrangements in Parallel G-Quadruplexes

Beatrice Karg and Klaus Weisz*[a]

**Abstract:** A G-quadruplex forming sequence from the *MYC* promoter region was modified with *syn*-favoring 8-bromo-2′-deoxyguanosine residues. Depending on the number and position of modifications in the intramolecular parallel G-quadruplex, substitutions with the bromoguanosine analogue at the 5′-tetrad induce conformational rearrangements with concerted all-*anti* to all-*syn* transitions for all residues of the modified G-quartet. No unfavorable steric interactions of the C8-substituents in the medium grooves are apparent in the high-resolution structure as determined for a tetrasubstituted *MYC* quadruplex that exclusively forms the all-*syn* isomer. In contrast, considerable steric clashes with 5′-phosphate oxygen atoms for those analogues that follow a less flexible 1-nucleotide loop in the native all-*anti* conformation seem to constitute the major driving force for the tetrad inversion and allow for the rational design of appropriately substituted sequences. Correlations found between the population of species subjected to a tetrad flip and melting temperatures indicate that more effective conformational transitions are compromised by lower thermal stabilities of the modified parallel quadruplexes.

## Introduction

Guanosine (G)-rich DNA and RNA sequences can fold into tetrastranded quadruplexes (G4) composed of a core of stacked G-quartets that are further stabilized by the inner coordination of cations like sodium or potassium. Their formation in vivo and their role in various physiological processes has been convincingly demonstrated in recent years.[1,2] Depending on the particular sequence and environmental conditions, G4 structures can form a variety of different topologies.[3–5] Thus, the molecularity, type of coordinated cation, cosolutes, and pH may all contribute to the individual conformation within a parallel, antiparallel, or hybrid-type topological family.[6–9] Each global fold is associated with a particular quadruplex core since guanosine torsion angles are closely connected to the relative orientation of the four G-tracts. Parallel structures, favored under molecular crowding conditions and formed by most G-rich sequences in oncogenic promoters, fold into G4s with all core guanosines adopting an *anti* glycosidic torsion angle. In contrast, antiparallel structures demand both *syn* and *anti* conformations and generally feature tetrads with different polarity of the cyclic hydrogen bond network associated with heteropolar stacking interactions.[10] However, quadruplex folding from genomic sequences may involve various kinetically controlled intermediate states to finally result in coexisting conformations of comparable thermodynamic stability.

For the stabilization of particular topologies, the incorporation of guanosine analogues has widely been employed in the past.[11–14] In particular, the *syn*-favoring 8-bromoguanosine (8-BrG) has often been used in matching *syn* positions as the stabilizing G substitute of a specific G4 fold.[15–19] Previous studies have even shown faster folding kinetics for site-specific 8-BrG modifications.[20] The bromo substituent at C8 is positioned in the grooves of the quadruplex and does not interfere with the Hoogsteen hydrogen bonds of a tetrad. However, even its incorporation at matching sites results in position-dependent G4 (de)stabilization ascribed to different steric hindrances of the 8-substituent protruding into narrow, medium, or wide grooves.[21,22]

Few attempts have been reported in substituting G residues for 8-BrG at non-matching *anti* positions to deliberately redirect folding into other topologies. An all-*anti* to an all-*syn* configurational transition of all four analogues within a tetrad while conserving a parallel topology was shown for loop-free tetramolecular model quadruplexes substituted with 8-bromo- or 8-methylguanosine.[23–27] Such polarity inversions of a single tetrad were recently also demonstrated in intramolecular [3+1] hybrid-type quadruplexes upon the incorporation of sugar-modified guanosine or a combination of xanthine and 8-oxoguanine bases.[28,29] Likewise, upon G to 8-BrG substitutions in a G-quartet a tetrad flip was very recently demonstrated to also occur in the very stable parallel quadruplex derived from the *MYC* promoter region (Figure 1).[30] In the latter studies, a trisubstituted quadruplex was also shown to form about equally populated all-*anti* and all-*syn* 5′-tetrad conformational isomers being in dynamic exchange. Interestingly, an all-*syn* G-

[a] *B. Karg, Prof. Dr. K. Weisz*
*Institute of Biochemistry*
*Ernst-Moritz-Arndt University Greifswald*
*Felix-Hausdorff-Str. 4, 17487 Greifswald (Germany)*
*E-mail: weisz@uni-greifswald.de*

*Supporting information and the ORCID identification number(s) for the author(s) of this article can be found under: https://doi.org/10.1002/chem.201801851.*

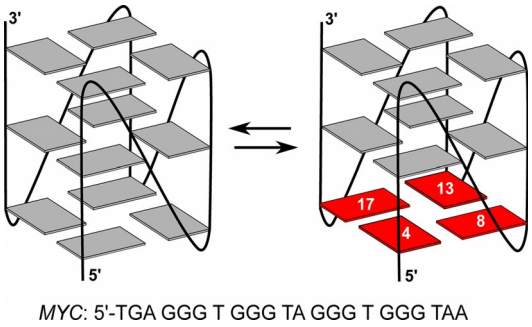



**Figure 1.** Schematic representation and sequence of the parallel *MYC* quadruplex in its native form (left) and with an inverted 5′-tetrad (right); *syn* and *anti* G residues are shown by red and grey rectangles, respectively.

quartet seems to be disfavored in unmodified quadruplex structures although few exceptions have been reported.[31,32]

Because G4 structures are not only targets for medicinal interventions but are also promising scaffolds for various aptamer and sensor systems,[33,34] the design and development of novel conformational variants through directed substitutions are of prime interest for G4-based bio- and nanotechnological applications. On the other hand, *syn*-favoring 8-BrG analogues may serve as excellent mimics for major DNA lesions introduced by carcinogenic arylamines and heterocyclic amines. These aromatic mutagens are known to be enzymatically converted into aryl radicals, reacting with the guanine base to yield 8-aryl-guanine derivatives that closely resemble bromo analogues in their propensity to adopt a *syn* conformation.[35,36] Likewise, DNA lesions like 8-oxo-guanine or 8-methyl-guanine form in vivo by oxidative damage or alkylation and are suggested to contribute to carcinogenic effects.[37,38]

We herein report on the impact of systematic nonmatching 8-BrG substitutions in the *MYC* quadruplex. Through a comparison of the parent structure with the derived high-resolution structure of an 8-BrG modified quadruplex with flipped tetrad, position-dependent conformational rearrangements can be rationalized. These are shown for G analogues within the 5′-tetrad of *MYC* to primarily depend on the type of 5′-adjacent propeller loops. The results presented allow for a more reliable prediction and thus rational design of modified G4s but also for a better appreciation of the potential effects exerted by lesions of C8-substituted guanines within G-rich quadruplex forming sequences.

## Results and Discussion

### Conformational transitions upon 8-BrG substitutions

Based on a sequence of the *MYC* promoter region that was previously shown to fold into a well-defined parallel quadruplex with three propeller loops and flanking sequences,[39] various positions within the 5′- and 3′-tetrads were selected for the incorporation of 8-BrG analogues. All modified sequences are compiled together with their UV melting temperature in Table 1 (see also the Supporting Information, Figure S1). The central tetrad was excluded from any modification due to its strong stacking between the two outer tetrads and its considerable resistance towards conformational rearrangements.[24,25] Quadruplexes with a modified 3′-tetrad include trisubstituted *MYC*(6,10,15) and *MYC*(6,15,19) as well as *MYC*(6,10,15,19) with all four guanines at positions 6, 10, 15, and 19 replaced by 8-BrG. NMR imino proton spectral regions and signatures in the CD spectra of all three sequences are in line with the formation of a major quadruplex featuring a conserved parallel topology with exclusive homopolar stacking interactions and all-*anti* G core residues (Supporting Information, Figure S2). However, decreased CD amplitudes and additional NMR resonances for nonexchangeable protons as well as a time-dependent decay of the imino proton signal intensity, in particular for tetrasubstituted *MYC*(6,10,15,19), suggest variable amounts of coexisting unfolded species and a low thermodynamic stability for all 3′-modified G4 analogues. The latter is also evidenced by their significantly reduced melting temperature with $|\Delta T_m| > 40\,^{\circ}C$ (Table 1). Such destabilizing effects have already been noted for symmetric tetramolecular quadruplexes fully substituted with 8-BrG or 8-MeG analogues at their 3′-tetrad.[23–25] However, 8-BrG substitutions of the present intramolecular *MYC* quadruplex do not enforce a 3′-tetrad inversion to a measurable extent in contrast to reports on tetramolecular quadruplexes featuring a C8-modified all-*syn* 3′-G quartet.[25,26]

In the following, we focused on the impact of 8-BrG substitutions at various positions within the 5′-tetrad of *MYC*. A corresponding 4,13,17-trisubstituted quadruplex *MYC*(4,13,17) has recently been shown to undergo a 5′-tetrad polarity switch with a dynamic interconversion of the two coexisting G4 conformational isomers with an all-*syn* and all-*anti* outer G-quartet. Replacing all four G residues by 8-BrG analogues resulted in a complete flip of the 5′-tetrad and the formation of a single quadruplex with a parallel topology and heteropolar stacking interactions.[30]

**Table 1.** *MYC*-derived sequences and their $T_m$ values.

| Name | Sequence 5′→3′[a] | $T_m$ [°C][b] |
|---|---|---|
| *MYC* | TGA GGG T GGG TA GGG T GGG TAA | 90.8 ± 0.2[c] |
| *MYC*(6,15,19) | TGA GGX T GGG TA GGX T GGX TAA | 47.9 ± 1.1 |
| *MYC*(6,10,15) | TGA GGX T GGX TA GGX T GGG TAA | 45.5 ± 0.4 |
| *MYC*(6,10,15,19) | TGA GGX T GGX TA GGX T GGX TAA | 41.5 ± 1.1 |
| *MYC*(4,13) | TGA XGG T GGG TA XGG T GGG TAA | 83.5 ± 1.9 |
| *MYC*(8,17) | TGA GGG T XGG TA GGG T XGG TAA | 67.0 ± 1.3 |
| *MYC*(4,8,13) | TGA XGG T XGG TA XGG T GGG TAA | 72.4 ± 0.6 |
| *MYC*(4,13,17) | TGA XGG T GGG TA XGG T XGG TAA | 70.4 ± 0.8 |
| *MYC*(8,13,17) | TGA GGG T XGG TA XGG T XGG TAA | 66.3 ± 0.9 |
| *MYC*(4,8,17) | TGA XGG T XGG TA GGG T XGG TAA | 69.5 ± 0.6 |
| *MYC*(4,8,13,17) | TGA XGG T XGG TA XGG T XGG TAA | 70.0 ± 1.0 |
| *MYC*-21 | TGA GGG TA GGG T GGG T GGG TAA | 85.6 ± 0.6 |
| *MYC*-21(4,9) | TGA XGG TA XGG T GGG T GGG TAA | 77.6 ± 1.0 |
| *MYC*-21(9,17) | TGA GGG TA XGG T GGG T XGG TAA | 73.0 ± 0.5 |
| *MYC*-21(4,13) | TGA XGG TA GGG T XGG T GGG TAA | 70.5 ± 0.6 |
| *MYC*-21(13,17) | TGA GGG TA GGG T XGG T XGG TAA | 66.9 ± 0.6 |

[a] X = 8-Bromo-2′-deoxyguanosine (8-BrG). [b] Determined by UV melting experiments in 20 mM potassium phosphate, 100 mM KCl, pH 7.0. [c] Taken from ref. [30]; determined by differential scanning calorimetry (DSC).

Initially, all four possible combinations with three substitution sites at the 5′-tetrad of *MYC* were evaluated (Table 1). CD spectra of native and substituted *MYC* sequences reveal transitions from the type I signature of the parallel *MYC* quadruplex as a result of hydrogen-bond networks of the same polarity to type II spectra characterized by the simultaneous occurrence of both homo- and heteropolar tetrad stacking as typically observed for [3+1]-hybrid structures (Supporting Information, Figure S3 a). Distinct attenuations of the amplitude at 265 nm and varied intensities of an additional positive CD band at 295 nm suggest 5′-tetrad polarity inversions to variable extents and the exclusive formation of an inverted species for the tetrasubstituted *MYC*(4,8,13,17).[30]

The observation of 12 well-resolved imino resonances in the $^{1}$H NMR spectra of native *MYC* and *MYC*(4,8,13,17) confirms the formation of a single well-defined quadruplex structure with 3 tetrads of Hoogsteen-bound guanines. In contrast, two sets with a total of 24 imino signals corroborate the coexistence of two quadruplex species in slow exchange for *MYC*(4,8,13), *MYC*(4,13,17), and *MYC*(8,13,17) (Figure 2). *MYC*(4,8,17) shows a strongly predominant species with a second low-populated species still noticeable. Standard strategies including 2D NOE experiments were used for the full or partial assignment of imino/H8/H6/H2 base and H1′/H2′/H2″ sugar proton resonances for major and minor G4 species. Assignments of the 8-BrG modified quadruplexes were strongly supported by a mutual comparison of spectral patterns starting with already assigned native *MYC* exclusively composed of G residues in an *anti* configuration and *MYC*(4,8,13,17) featuring a flipped all-*syn* 5′-tetrad (Supporting Information, Figure S4).[30] This also allowed for the unambiguous identification of an all-*anti* or all-*syn* G-quartet. It should also be noted, that despite missing diagnostic guanine H8 NOEs in the case of an 8-bromo substituent, information on the glycosidic torsion angle for unmodified guanine residues within the same tetrad fix the configuration of 8-BrG in the case of a parallel quadruplex.

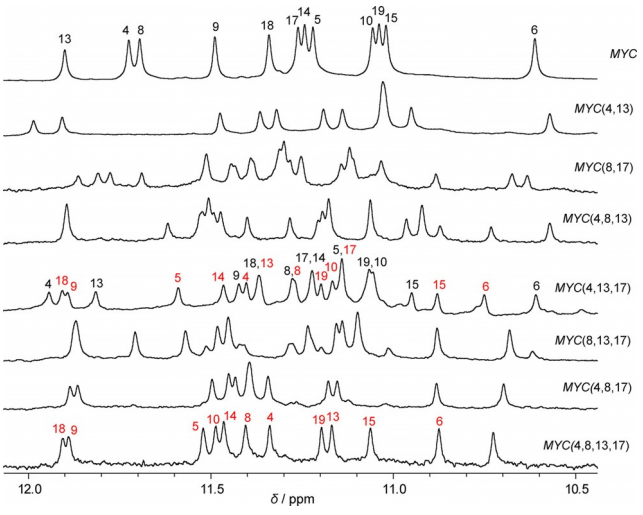



**Figure 2.** Imino proton spectral region of native and 8-BrG substituted *MYC* quadruplexes at 30 °C in 10 mM potassium phosphate buffer, pH 7; peak assignments for the G-tract guanines are indicated for the native *MYC* quadruplex, for the tetrasubstituted *MYC*(4,8,13,17) with flipped 5′-tetrad, and for trisubstituted *MYC*(4,13,17) forming two coexisting species with conserved and flipped 5′-tetrad; residue numbers in black and red refer to quadruplexes with an all-*anti* and all-*syn* 5′-tetrad, respectively.

NMR data clearly confirm the exclusive formation of parallel quadruplexes with either all-*anti* or inverted all-*syn* 5′-tetrads. Relative populations of coexisting species were determined by integration of the well-resolved A12 H2 peak to yield reliable values for each sequence (Supporting Information, Figure S3 b). Similar results were also obtained by an inspection of the relative CD amplitudes at 265 and 295 nm. However, the absence of a sharp isoelliptic point for the superimposed CD spectra indicates an intrinsic effect of the 8-bromo substituent on the electronic transition dipole moment and hence on the CD amplitude, introducing significant uncertainties.

As can be seen in Figure 3, the efficiency of a 5′-tetrad flip strongly depends on the number and positioning of modifications. Looking at the four trisubstituted quadruplexes, *MYC*(4,8,17) almost exclusively undergoes a polarity reversal of its outer tetrad. Apparently, an additional substitution at position 13 is dispensable for enforcing a concerted *anti*→*syn* conversion. On the other hand, lacking a modification at position 17 in *MYC*(4,8,13) will only result in a ≈40% tetrad reversal. Populations of modified quadruplexes with flipped tetrad indicate additive effects and substitution sites can easily be ordered for their growing contributions to a 5′-tetrad flip according to 13 < 4 < 8 < 17.

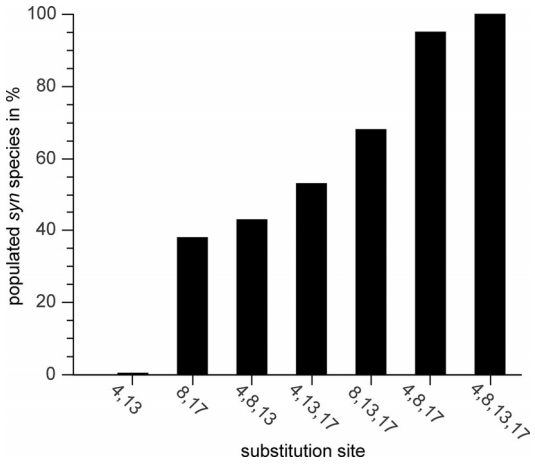



**Figure 3.** Population of quadruplex structures with an inverted all-*syn* 5′-tetrad for different 8-BrG substituted *MYC* analogues.

It is conspicuous that the two most efficient sites for an 8-BrG-induced 5′-tetrad inversion, namely position 8 and 17, follow a rather rigid single-nucleotide (1-nt) propeller loop. In contrast, less sensitive positions 4 and 13 are preceded by a 3-nt overhang and a more flexible 2-nt propeller loop, respectively. This suggests steric contributions to be effective in the substitution-enforced *anti*→*syn* transitions. To further test the position-dependent propensities for a tetrad flip, two additional *MYC* derivatives with only two modifications at position 4/13 and 8/17 were subjected to corresponding CD and NMR spectral analyses (Figure 2 and Figure S3 in the Supporting Informa-

tion). In line with additive effects, disubstituted quadruplexes have a lower tendency compared to the trisubstituted quadruplexes for a tetrad inversion. However, whereas only a conserved all-*anti* species is detected for *MYC*(4,13), *MYC*(8,17) shows a considerable population of about 40% of *syn* conformational isomers, again emphasizing the larger impact of position 8 and 17 in *MYC* for a 5′-tetrad polarity inversion.

### NMR structure of tetrabrominated *MYC*(4,8,13,17)

To gain more insight into the conformational changes associated with a 5′-tetrad flip and potential steric interactions of an 8-bromo substituent, the three-dimensional structure of *MYC*(4,8,13,17) was determined through restrained molecular dynamics simulations in explicit water based on NMR-derived distances and torsion angles (Supporting Information, Table S1). Final RMSD values for the ten final structures amount to 0.58 and 1.32 Å for the G-core and the overall quadruplex in line with a well-defined conformation of the modified G4 (Figure 4a). Although glycosidic torsion angles for all G residues of the 5′-tetrad have changed to a *syn* configuration, the overall fold is largely conserved for the 8-bromo-substituted G4 when compared to the previously reported structure of the parent quadruplex (Figure 4b).[39] Some deviations are observed for the loop regions and in particular for the 5′-overhang sequence adjacent to the inverted tetrad. All core residues including the 8-BrG analogues adopt a sugar pucker in the *south* domain as demonstrated experimentally by the analysis of sugar proton scalar couplings (not shown). The flipped 5′-tetrad manifests itself in a noticeable change of stacking interactions with the central tetrad. While stacking interactions through partial overlap of the five- and six-membered rings of guanine bases are observed in the native quadruplex with its *anti–anti* steps, there is an extensive overlap of the five-membered rings of purines along the *syn-anti* steps of the flipped structure as generally found for antiparallel and hybrid-type quadruplexes (Figure 4c).

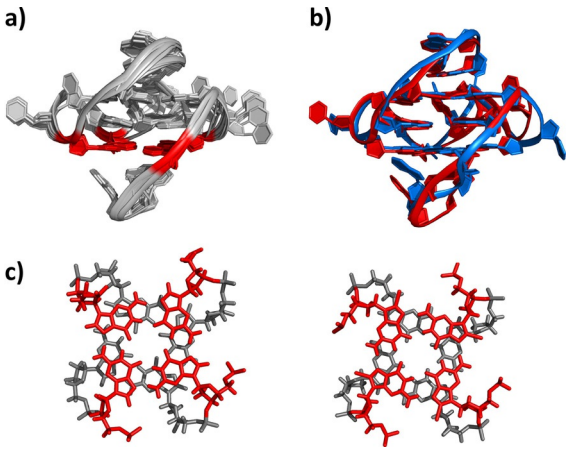



**Figure 4.** a) Superposition of the 10 final structures of *MYC*(4,8,13,17) with the 5′-tetrad shown in red. b) Superimposed representative structures of unmodified *MYC* (blue, pdb 1XAV) and tetrasubstituted *MYC*(4,8,13,17) (red). c) Stacking of the 5′-tetrad (red) on the central tetrad (grey) in *MYC* (left) and *MYC*(4,8,13,17) (right).

A base flip about the glycosidic bond must be accompanied by translational motions to maintain stabilizing stacking interactions and concerted motions along the sugar-phosphate backbone enable the compensation for major structural aberrations. Generally, $\alpha$, $\beta$, and $\gamma$ dihedral angles in helical polynucleotides preferably adopt *gauche*− ($g^-$), *trans* (t), and *gauche*+ ($g^+$) conformations, respectively.[40,41] However, 5′-neighboring short propeller loops are expected to affect these torsion angles as observed in G residues of the *MYC*(4,8,13,17) 5′-tetrad (Supporting Information, Figure S5). In line with the known influence of *syn*/*anti* orientations on $\gamma$, the inverted all-*syn* tetrad in *MYC*(4,8,13,17) tends to increasingly populate t and $g^-$ conformations for $\gamma$ when compared to corresponding positions in the parent *MYC*. It should be noted, however, that a more detailed evaluation and comparison of backbone parameters is limited by a notorious difficulty in their accurate determination. This is not only attributed to lacking NMR data but also to their sensitivity towards the force field employed during molecular dynamics calculations.[42]

With these limitations in mind we wanted to assess potential steric interactions of the 8-bromine substituent in the modified quadruplex with its inverted 5′-tetrad as well as in the conserved all-*anti* quadruplex structure. Due to the unavailability of the latter, H8 of guanines in the 5′-tetrad of the reported *MYC* high-resolution structure (pdb 1XAV) were replaced by bromine while keeping the native conformation. Notably, significant steric clashes between the bromine and 5′-phosphate oxygen atoms with interatomic distances much smaller than the sum of their van der Waals radii become apparent for G positions following a propeller loop (Figure 5). On the other hand, *syn* 8-BrG residues in the flipped tetrad of *MYC*(4,8,13,17) reside in a medium groove and experience no obvious unfavorable steric interactions. These observations suggest that tight 1-nt loops severely restrict appropriate rearrangements to attenuate bromine clashes in the *MYC* quadruplex, strongly enhancing the tendency of a base inversion at position 8 and 17 to enforce a concerted tetrad flip. In contrast, position 4 but also position 13 with a preceding 2-nt loop are less restrained and may allow conformational adjustments for a better accommodation of the bromine atom. It should be mentioned, that corresponding steric clashes are partially also expected for 8-BrG residues when incorporated at the 3′-tetrad

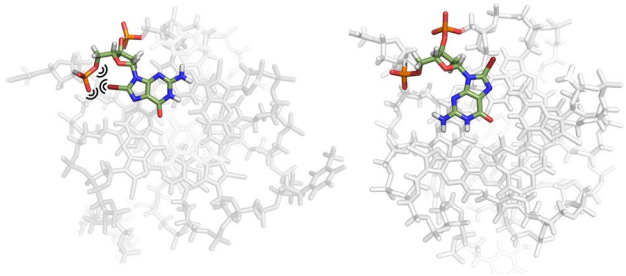


**Figure 5.** Orientation and steric interactions of the bromine substituent in 8-BrG at position 17 in the all-*anti MYC* (left) and in *MYC*(4,8,13,17) with an all-*syn* 5′-tetrad (right). The all-*anti MYC* quadruplex was generated by an H8→Br substitution in the native structure (pdb 1XAV).

of the native *MYC* quadruplex. However, low thermal stabilities of the corresponding modified quadruplexes associated with a resistance towards a 3′-tetrad inversion also point to serious energetic penalties upon the formation of an all-*syn* 3′-tetrad (see Table 1 and Figure S2 in the Supporting Information). Unfavorable interactions in such a 3′-inverted quadruplex may largely be attributed to the four *anti-syn* steps between the central and 3′-outer tetrad that have been shown to be significantly disfavored when compared to *anti–anti* or *syn–anti* steps.[43]

Assuming base-pair opening to be the rate-limiting step in the guanine imino exchange with water, enhanced flexibility can be expected to also impact solvent exchange rates. Deuterium exchange experiments demonstrate significantly higher exchange rates for the 5′- and 3′-outer tetrads when compared to the central tetrad of *MYC* but also show noticeably fast and slow exchange processes for residue G13 and G17, respectively (Supporting Information, Figure S6 a). In order to quantitatively determine exchange rates for imino protons at the *MYC* 5′-tetrad, ROESY spectra were acquired and the imino–$H_2O$ exchange crosspeaks analyzed as a function of mixing time as reported previously (Supporting Information, Figure S6 b).[30] Whereas a $k_{ex} = 1.3\ s^{-1}$ was obtained for the G13 imino proton at 50 °C, exchange of the G17 imino was found to be too slow to be determined under the experimental conditions employed. Increasing exchange rates follow the order G17 < G4 ≤ G8 < G13, indicating that loop flexibilities partially also determine the ease of base-pair opening. A relatively slow exchange rate found at G4 may be a consequence of additional screening effects by a stacked 5′-overhang.

### 8-BrG substitutions on a *MYC* variant

In seeking to further validate the proposed molecular basis of position-dependent effects on the conformational transitions, we reshuffled loop sequences in the *MYC* quadruplex. Thus, the 1-nt and 2-nt loops following the first and second G-tract were exchanged to yield the *MYC*-21 sequence. Initial NMR as well as CD spectra confirmed a conserved and well-defined all-*anti* quadruplex for *MYC*-21 (Supporting Information, Figure S7). In the following, *MYC*-21 was disubstituted with 8-BrG at positions 4, 9, 13, or 17 of the 5′-tetrad. Note, that the two 1-nt loops now precede positions 13 and 17 whereas position 9 is linked to the more flexible 2-nt loop in the newly designed G4. As in the case of modified *MYC* all four variants of *MYC*-21 gave NMR spectra only composed of a native and a 5′-inverted species with populations determined by relative intensities of corresponding proton resonances (Figure 6 top). In line with the distribution and length of loops, efficiencies of a 5′-tetrad inversion decrease in the order *MYC*-21(13,17) > *MYC*-21(4,13) > *MYC*-21(9,17) > *MYC*-21(4,9). This translates to an increasing propensity of 8-BrG to adopt a *syn* configuration in a positional order 9 < 4 < 17 < 13. Thus, *MYC*-21(4,9) only forms an all-*anti* conformer like its counterpart *MYC*(4,13). On the other hand, *MYC*-21(13,17) with substitutions at positions 13 and 17 closely follows *MYC*(8,17) in exhibiting a significant population of ≈40% quadruplex with an all-*syn* 5′-tetrad.

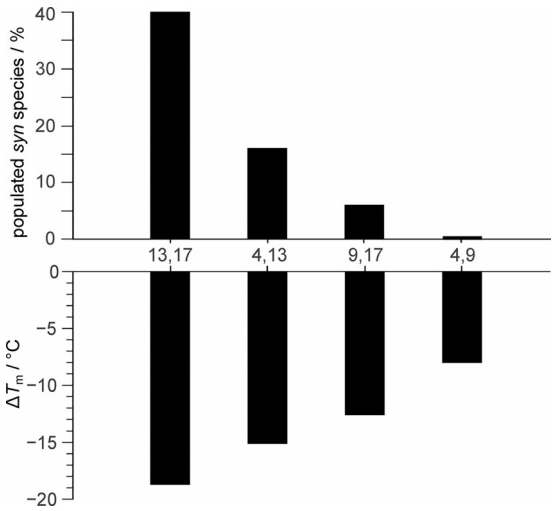



**Figure 6.** Population of quadruplex structures with an inverted all-*syn* 5′-tetrad (top) and changes in melting temperature $\Delta T_m$ (bottom) for different 8-BrG substituted *MYC*-21 analogues; substitution sites are indicated.

These results may serve as a proof of concept in the prediction of bromine-induced conformational changes required for a rational design of 8-BrG modified parallel quadruplexes with propeller loops.

Substituted quadruplexes are destabilized by the 8-BrG analogue when incorporated into the all-*anti* parallel topology (Table 1). Yet, with melting temperatures > 65 °C all quadruplexes modified at their 5′-tetrad are thermally stable. Conspicuously, thermal stabilities show strong correlations with the extent of a 5′-tetrad flip. Because 8-BrG is expected to have an intrinsic effect on the quadruplex stability, only quadruplexes with the same sequence and number of substitution sites may strictly be used for a direct comparison. Following increasing populations of an all-*syn* tetrad, melting temperatures $T_m$ gradually decrease from *MYC*-21(4,9) > *MYC*-21(9,17) > *MYC*-21(4,13) > *MYC*-21(13,17) (Figure 6 bottom). Likewise, corresponding correlations exist for the di- as well as the trisubstituted *MYC* sequences except for *MYC*(4,8,17) with its almost exclusive formation of an all-*syn* tetrad and an associated $T_m$ higher than expected. Assuming that the *syn* conformational isomers being devoid of bromine steric clashes are energetically similar, differences in melting temperatures directly reflect differences in the relative energies of the *anti* isomers with their position-dependent steric interactions. In fact, an endothermic all-*syn* → all-*anti* transition with an enthalpy change of $\Delta H^\circ \approx 20\ kJ\,mol^{-1}$ has been determined for *MYC*(4,13,17) and suggests higher populations of the all-*anti* isomers at the respective melting temperature.[30] As a consequence, driving a tetrad polarity inversion through appropriate substitutions at nonmatching positions may be compromised by a decrease in thermal stability.

## Conclusion

The site-specific incorporation of modified guanosine analogues not only constitutes an attractive strategy for stabilizing

a particular quadruplex fold but also for the targeted refolding into a new G4 conformation. This can be achieved by sugar or base-modified G residues that either favor *syn* or *anti* glycosidic torsion angles and are incorporated at matching or non-matching positions. In contrast to most sugar modifications, C8-substituted guanine nucleotides like 8-BrG or 8-MeG favor a *syn* configuration and are appropriate tools when trying to redirect folding of all-*anti* parallel quadruplexes that are almost exclusively formed from native G-rich DNA promoter and RNA sequences in vitro.

Substituting G for 8-BrG at non-matching *anti* positions will almost invariably destabilize the quadruplex structure, often resulting in a complete disruption of the native fold and the formation of various coexisting species. However, the observation of concerted *anti*→*syn* transitions with a tetrad polarity inversion even for the kinetically and thermodynamically robust *MYC* quadruplex demonstrates the power of 8-BrG for a directed conformational change when substituted at appropriate positions. Through a detailed assessment of the impact of various substitution patterns on conformation and stability in combination with quadruplex high-resolution structures, position-dependent interactions that are expected to drive such conformational changes were identified. The information gained will enable a more rational design of 8-BrG modified sequences for novel structural motifs. These may be useful for various technological applications that rely on a G4 platform. On the other hand, a detailed account of interactions exerted by 8-substituted guanine analogues may also provide hints at potential effects of carcinogen-induced DNA lesions within G4-forming sequences and of the roles played by quadruplexes within the cellular environment.

## Experimental Section

### Materials and sample preparation

Unmodified and HPLC-purified 8-BrG modified DNA oligonucleotides were purchased from TIB MOLBIOL (Berlin, Germany). Before use, oligonucleotides were ethanol precipitated and the concentrations were determined spectrophotometrically by measuring their absorbance $A_{260}$ at 90 °C in salt-free $H_2O$ using an extinction coefficient for all oligonucleotides of $\varepsilon_{260} = 258\,930\ L\,mol^{-1}\,cm^{-1}$. Samples for optical measurements were prepared by dissolving the corresponding oligonucleotides in a buffer with 20 mM potassium phosphate, 100 mM KCl, pH 7.0. To increase sensitivity without affecting the quadruplex conformation, a low salt buffer with 10 mM potassium phosphate, pH 7.0, was used for the NMR experiments. Prior to measurements, the buffered samples were annealed by heating to 90 °C followed by slow cooling to room temperature within approximately 6 hours. Final concentrations of oligonucleotides were 5 μM for the CD and UV experiments and between 0.11 and 0.52 mM for the NMR experiments.

### Circular dichroism

CD spectra were acquired with a Jasco J-810 spectropolarimeter at 20 °C (Jasco, Tokyo, Japan). Spectra of oligonucleotides were recorded with a bandwidth of 1 nm, a scanning speed of 50 nm min$^{-1}$ and 5 accumulations in 10 mm quartz cuvettes. All spectra were blank corrected.

### UV melting

Melting curves were recorded in triplicate between 10 and 90 °C on a Cary 100 spectrophotometer equipped with a Peltier temperature control unit (Varian, Darmstadt, Germany). Quartz cuvettes of 10 mm path length were used and data points were measured in 0.5 °C intervals with heating rates of 0.2 °C min$^{-1}$. Melting temperatures $T_m$ were determined by the maximum in a first derivative plot of the absorbance recorded at 295 nm as a function of temperature.

### NMR spectroscopy

All NMR spectra were acquired on a Bruker Avance 600 MHz spectrometer equipped with an inverse $^{1}H/^{13}C/^{15}N/^{31}P$ quadruple resonance cryoprobehead and *z*-field gradients. Data were processed using Topspin 3.1 and analyzed with CcpNmr Analysis.[44] Proton chemical shifts were referenced relative to the temperature-dependent $H_2O$ or HDO chemical shift. For the proton 1D and 2D NOE measurements in 90% $H_2O$/10% $D_2O$ a WATERGATE with w5 element was employed for solvent suppression. Typically, NOESY experiments in 90% $H_2O$/10% $D_2O$ were performed between 10 and 50 °C with mixing times from 80 to 300 ms and a spectral width of 10 kHz. 2048×1024 data points with 16–64 transients per t1 increment and a recycle delay of 2 s were collected in t2 and t1. Prior to Fourier transformation data were zero-filled to give a 4096×2048 matrix and both dimensions were apodized by squared sine bell window functions. For EASY-ROESY experiments mixing times from 40 to 120 ms were employed with the angle $\theta$ for the tilted rotating frame fixed to 50°. Kinetic analysis of ROESY crosspeaks has been performed as previously described.[30]

### Structure refinement

The RED software was used to calculate the partial atomic charges for the modified 8-BrG nucleotides.[45] Ten lowest-energy starting models from 300 extended structures were generated with a distance geometry simulated annealing protocol in Xplor-NIH 2.39.[46] Subsequently, restrained simulated annealing was performed with the Amber15 software in implicit water using the parmbsc0 force field including the $\chi_{OL4}$, $\varepsilon\zeta_{OL1}$, and $\beta_{OL1}$ corrections.[42,47–49] The structures were equilibrated for 1 ns with explicit solvent and shortly minimized in vacuum. Details of the calculation process have been reported previously.[50] Structural parameters were determined with the 3DNA software package.[51]

### Accession Codes

Atomic coordinates of the *MYC*(4,8,13,17) quadruplex have been deposited in the Protein Data Bank (accession code 6ERL).

## Conflict of interest

The authors declare no conflict of interest.

**Keywords:** C8-substituted guanine · circular dichroism · glycosidic torsion angle · G-quadruplex · NMR spectroscopy

[1] V. S. Chambers, G. Marsico, J. M. Boutell, M. Di Antonio, G. P. Smith, S. Balasubramanian, *Nat. Biotechnol.* **2015**, *33*, 877–881.

[2] 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*, 1267–1272.

[3] P. Hazel, J. Huppert, S. Balasubramanian, S. Neidle, *J. Am. Chem. Soc.* **2004**, *126*, 16405 – 16415.
[4] S. Burge, G. N. Parkinson, P. Hazel, A. K. Todd, S. Neidle, *Nucleic Acids Res.* **2006**, *34*, 5402 – 5415.
[5] M. Webba da Silva, *Chem. Eur. J.* **2007**, *13*, 9738 – 9745.
[6] J. Dai, M. Carver, C. Punchihewa, R. A. Jones, D. Yang, *Nucleic Acids Res.* **2007**, *35*, 4927 – 4940.
[7] N. Smargiasso, F. Rosu, W. Hsia, P. Colson, E. S. Baker, M. T. Bowers, E. De Pauw, V. Gabelica, *J. Am. Chem. Soc.* **2008**, *130*, 10208 – 10216.
[8] B. Heddi, A. T. Phan, *J. Am. Chem. Soc.* **2011**, *133*, 9824 – 9833.
[9] P. Galer, B. Wang, P. Šket, J. Plavec, *Angew. Chem. Int. Ed.* **2016**, *55*, 1993 – 1997; *Angew. Chem.* **2016**, *128*, 2033 – 2037.
[10] A. I. Karsisiotis, C. O'Kane, M. Webba da Silva, *Methods* **2013**, *64*, 28 – 35.
[11] C. J. Lech, Z. Li, B. Heddi, A. T. Phan, *Chem. Commun.* **2012**, *48*, 11425 – 11427.
[12] N. Martín-Pintado, M. Yahyaee-Anzahaee, G. F. Deleavey, G. Portella, M. Orozco, M. J. Damha, C. González, *J. Am. Chem. Soc.* **2013**, *135*, 5344 – 5347.
[13] M. Marušič, R. N. Veedu, J. Wengel, J. Plavec, *Nucleic Acids Res.* **2013**, *41*, 9524 – 9536.
[14] V. Kuryavyi, A. Majumdar, A. Shallop, N. Chernichenko, E. Skripkin, R. Jones, D. J. Patel, *J. Mol. Biol.* **2001**, *310*, 181 – 194.
[15] E. Dias, J. L. Battiste, J. R. Williamson, *J. Am. Chem. Soc.* **1994**, *116*, 4479 – 4480.
[16] J. Brčić, J. Plavec, *Nucleic Acids Res.* **2015**, *43*, 8590 – 8600.
[17] Y. Xu, Y. Noguchi, H. Sugiyama, *Bioorg. Med. Chem.* **2006**, *14*, 5584 – 5591.
[18] A. Matsugami, Y. Xu, Y. Noguchi, H. Sugiyama, M. Katahira, *FEBS J.* **2007**, *274*, 3545 – 3556.
[19] K. W. Lim, S. Amrane, S. Bouaziz, W. Xu, Y. Mu, D. J. Patel, K. N. Luu, A. T. Phan, *J. Am. Chem. Soc.* **2009**, *131*, 4301 – 4309.
[20] J. Gros, F. Rosu, S. Amrane, A. De Cian, V. Gabelica, L. Lacroix, J.-L. Mergny, *Nucleic Acids Res.* **2007**, *35*, 3064 – 3075.
[21] C. J. Lech, J. K. C. Lim, J. M. W. Lim, S. Amrane, B. Heddi, A. T. Phan, *Biophys. J.* **2011**, *101*, 1987 – 1998.
[22] A. I. Karsisiotis, M. Webba da Silva, *Molecules* **2012**, *17*, 13073 – 13086.
[23] V. Esposito, A. Randazzo, G. Piccialli, L. Petraccone, C. Giancola, L. Mayol, *Org. Biomol. Chem.* **2004**, *2*, 313 – 318.
[24] A. Virgilio, V. Esposito, A. Randazzo, L. Mayol, A. Galeone, *Nucleic Acids Res.* **2005**, *33*, 6188 – 6195.
[25] P. L. Thao Tran, A. Virgilio, V. Esposito, G. Citarella, J.-L. Mergny, A. Galeone, *Biochimie* **2011**, *93*, 399 – 408.
[26] A. Virgilio, V. Esposito, G. Citarella, A. Pepe, L. Mayol, A. Galeone, *Nucleic Acids Res.* **2012**, *40*, 461 – 475.
[27] A. Virgilio, V. Esposito, G. Citarella, L. Mayol, A. Galeone, *ChemBioChem* **2012**, *13*, 2219 – 2224.
[28] J. Dickerhoff, K. Weisz, *Angew. Chem. Int. Ed.* **2015**, *54*, 5588 – 5591; *Angew. Chem.* **2015**, *127*, 5680 – 5683.
[29] V. V. Cheong, C. J. Lech, B. Heddi, A. T. Phan, *Angew. Chem. Int. Ed.* **2016**, *55*, 160 – 163; *Angew. Chem.* **2016**, *128*, 168 – 171.
[30] B. Karg, L. Haase, A. Funke, J. Dickerhoff, K. Weisz, *Biochemistry* **2016**, *55*, 6949 – 6955.
[31] X. Tong, W. Lan, X. Zhang, H. Wu, M. Liu, C. Cao, *Nucleic Acids Res.* **2011**, *39*, 6753 – 6763.
[32] P. Šket, A. Virgilio, V. Esposito, A. Galeone, J. Plavec, *Nucleic Acids Res.* **2012**, *40*, 11047 – 11057.
[33] D. Sen, L. C. H. Poon, *Crit. Rev. Biochem. Mol. Biol.* **2011**, *46*, 478 – 492.
[34] J. L. Neo, K. Kamaladasan, M. Uttamchandani, *Curr. Pharm. Des.* **2012**, *18*, 2048 – 2057.
[35] R. A. Manderville, S. D. Wetmore, *Chem. Sci.* **2016**, *7*, 3482 – 3493.
[36] M. Sproviero, A. M. R. Verwey, K. M. Rankin, A. A. Witham, D. V. Soldatov, R. A. Manderville, M. I. Fekry, S. J. Sturla, P. Sharma, S. D. Wetmore, *Nucleic Acids Res.* **2014**, *42*, 13405 – 13421.
[37] A. M. Fleming, Y. Ding, C. J. Burrows, *Proc. Natl. Acad. Sci. USA* **2017**, *114*, 2604 – 2609.
[38] N. Shrivastav, D. Li, J. M. Essigmann, *Carcinogenesis* **2010**, *31*, 59 – 70.
[39] A. Ambrus, D. Chen, J. Dai, R. A. Jones, D. Yang, *Biochemistry* **2005**, *44*, 2048 – 2058.
[40] W. Saenger, *Principles of Nucleic Acid Structure*, Springer, New York, **1984**, chapter 4.
[41] W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell, P. A. Kollman, *J. Am. Chem. Soc.* **1995**, *117*, 5179 – 5197.
[42] M. Krepl, M. Zgarbová, P. Stadlbauer, M. Otyepka, P. Banáš, J. Koča, T. E. Cheatham III, P. Jurečka, J. Šponer, *J. Chem. Theory Comput.* **2012**, *8*, 2506 – 2520.
[43] J. Šponer, A. Mládek, N. Špačková, X. Cang, T. E. Cheatham, S. Grimme, *J. Am. Chem. Soc.* **2013**, *135*, 9785 – 9796.
[44] W. F. Vranken, W. Boucher, T. J. Stevens, R. H. Fogh, A. Pajon, M. Llinas, E. L. Ulrich, J. L. Markley, J. Ionides, E. D. Laue, *Proteins Struct. Funct. Bioinf.* **2005**, *59*, 687 – 696.
[45] E. Vanquelef, S. Simon, G. Marquant, E. Garcia, G. Klimerak, J. C. Delepine, P. Cieplak, F.-Y. Dupradeau, *Nucleic Acids Res.* **2011**, *39*, W511 – W517.
[46] C. D. Schwieters, J. J. Kuszewski, N. Tjandra, G. M. Clore, *J. Magn. Reson.* **2003**, *160*, 65 – 73.
[47] P. Várnai, D. Djuranovic, R. Lavery, B. Hartmann, *Nucleic Acids Res.* **2002**, *30*, 5398 – 5406.
[48] M. Zgarbová, J. Šponer, M. Otyepka, T. E. Cheatham III, R. Galindo-Murillo, P. Jurečka, *J. Chem. Theory Comput.* **2015**, *11*, 5723 – 5736.
[49] M. Zgarbová, F. J. Luque, J. Šponer, T. E. Cheatham III, M. Otyepka, P. Jurečka, *J. Chem. Theory Comput.* **2013**, *9*, 2339 – 2354.
[50] J. Dickerhoff, L. Haase, W. Langel, K. Weisz, *ACS Chem. Biol.* **2017**, *12*, 1308 – 1315.
[51] X.-J. Lu, W. K. Olson, *Nucleic Acids Res.* **2003**, *31*, 5108 – 5121.

Manuscript received: April 15, 2018
Revised manuscript received: May 12, 2018
Accepted manuscript online: May 14, 2018
Version of record online: June 14, 2018

CHEMISTRY
A European Journal

Supporting Information

# Loop Length Affects *Syn–Anti* Conformational Rearrangements in Parallel G-Quadruplexes

Beatrice Karg and Klaus Weisz*[a]

chem_201801851_sm_miscellaneous_information.pdf

## Author Contributions

B.K. Formal analysis: Lead; Investigation: Lead; Validation: Equal; Writing – original draft: Equal.

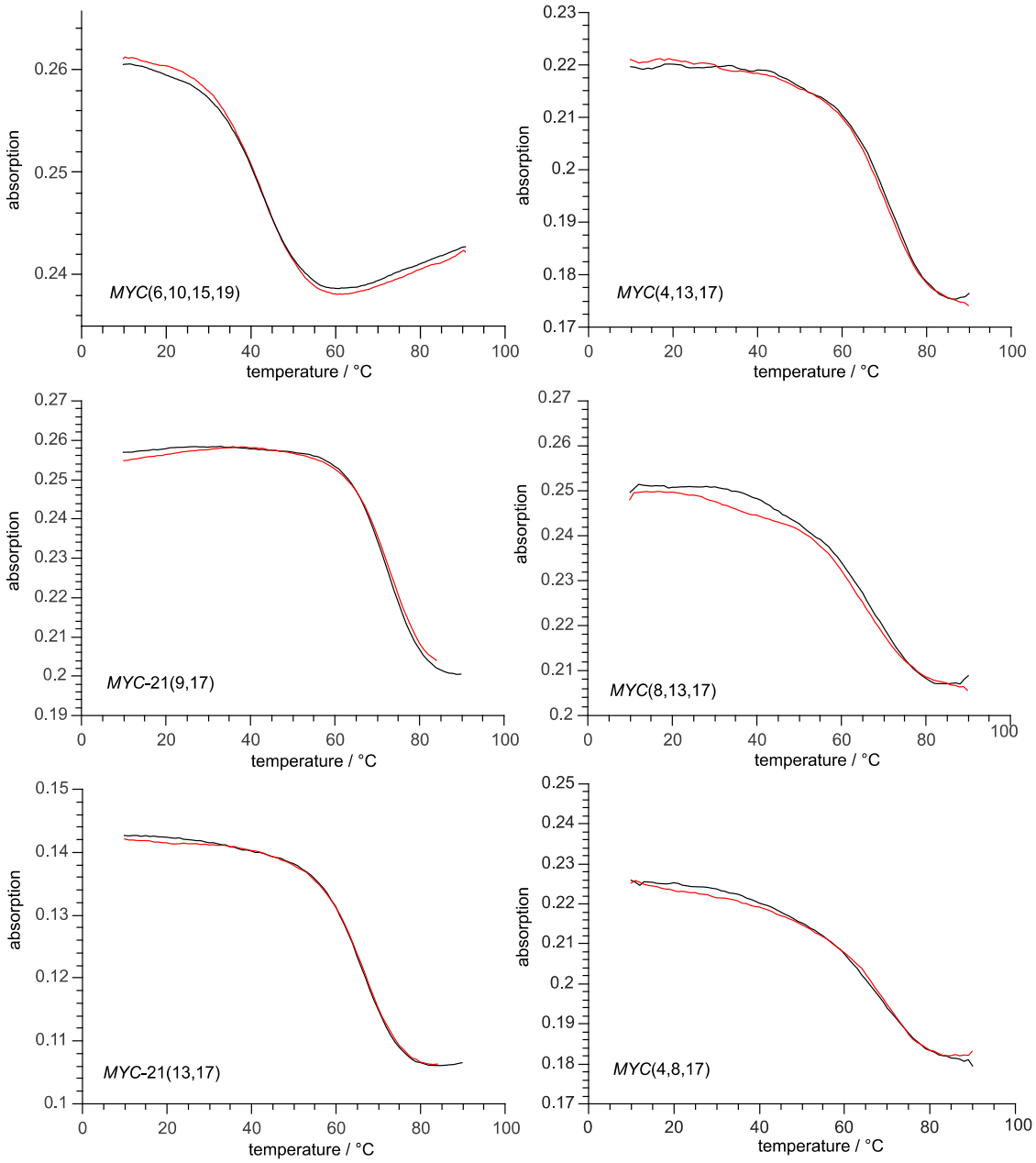



**Figure S1.** Representative UV melting profiles with heating (black) and cooling phase (red) of 8-BrG modified *MYC* quadruplexes. *MYC* species trisubstituted at their 5'-tetrad (right side) tend to show broader transitions with increasing populations of an all-*syn* 5'-tetrad.

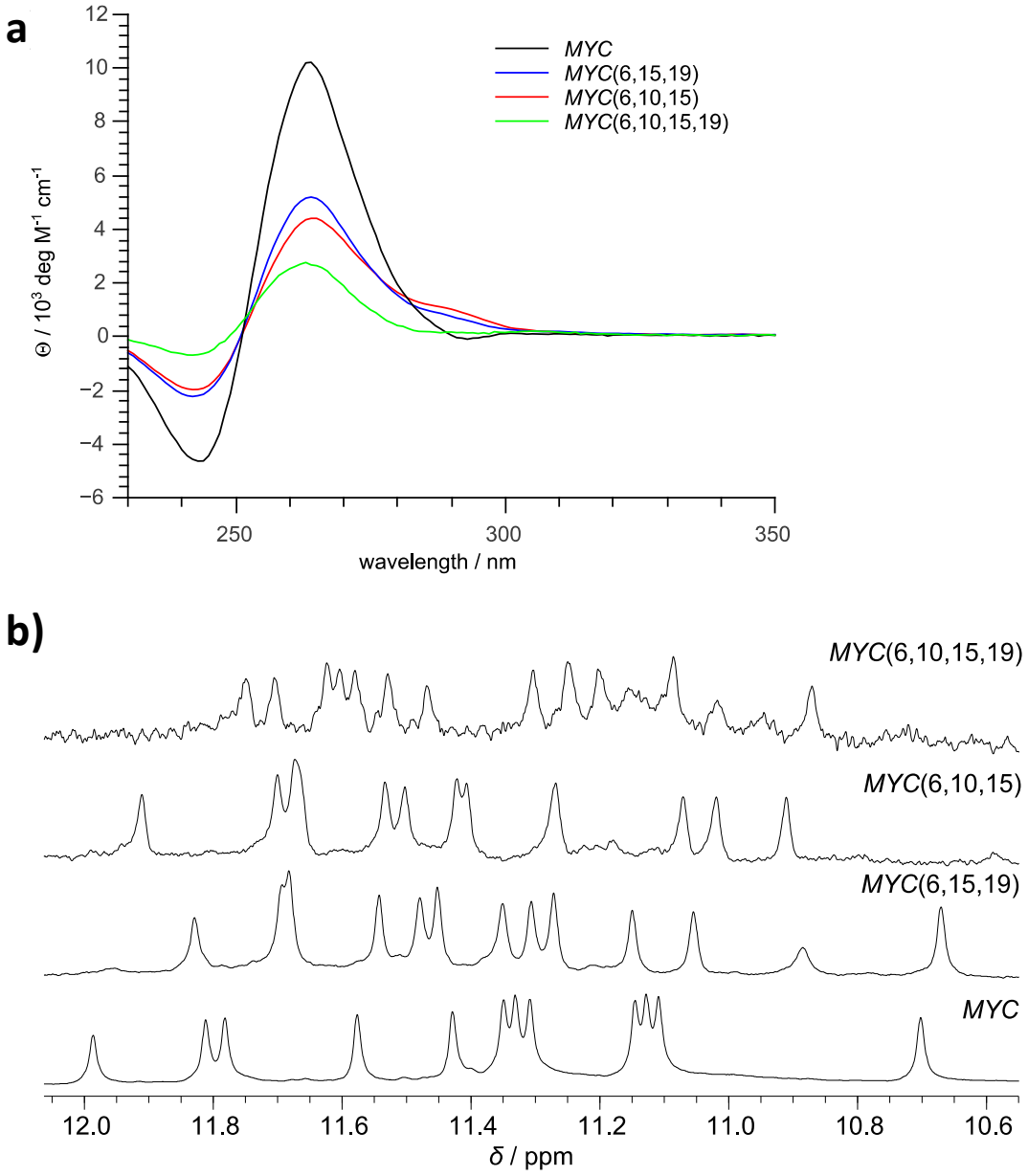



**Figure S2.** a) Superimposed CD spectra with molar ellipticities and b) NMR imino proton spectral region of unmodified *MYC* and of *MYC* substituted at its 3'-tetrad with 8-BrG analogs; numbers in parentheses denote the substitution sites. CD and NMR spectra were acquired at 20 °C and 30 °C, respectively.

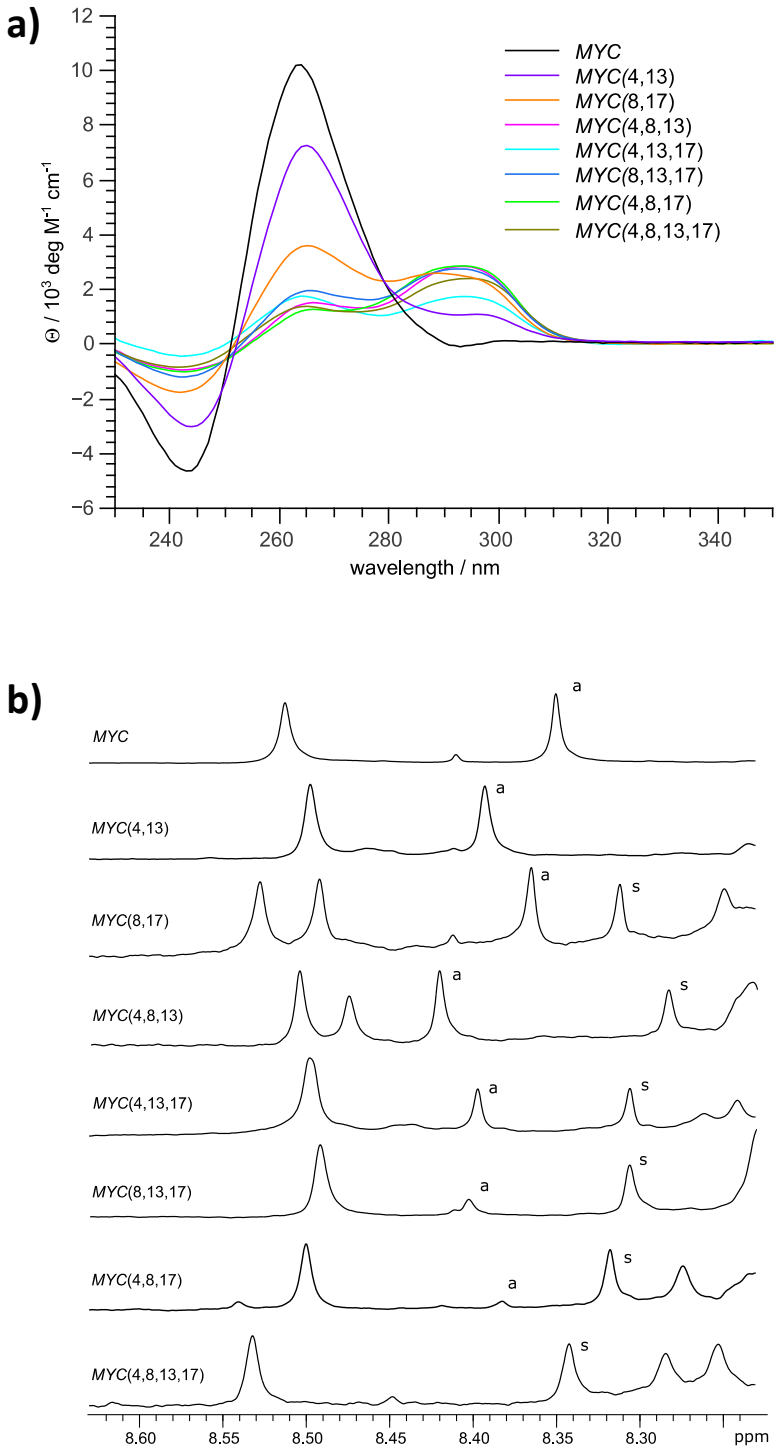



**Figure S3.** a) Superimposed CD spectra with molar ellipticities and b) NMR spectral region with A12 H2 resonances labeled a and s for isomers with an all-*anti* and all-*syn* 5'-tetrad; numbers in parentheses denote substitution sites in the 8-BrG modified *MYC* sequences. CD and NMR spectra were acquired at 20 °C and 30 °C, respectively.

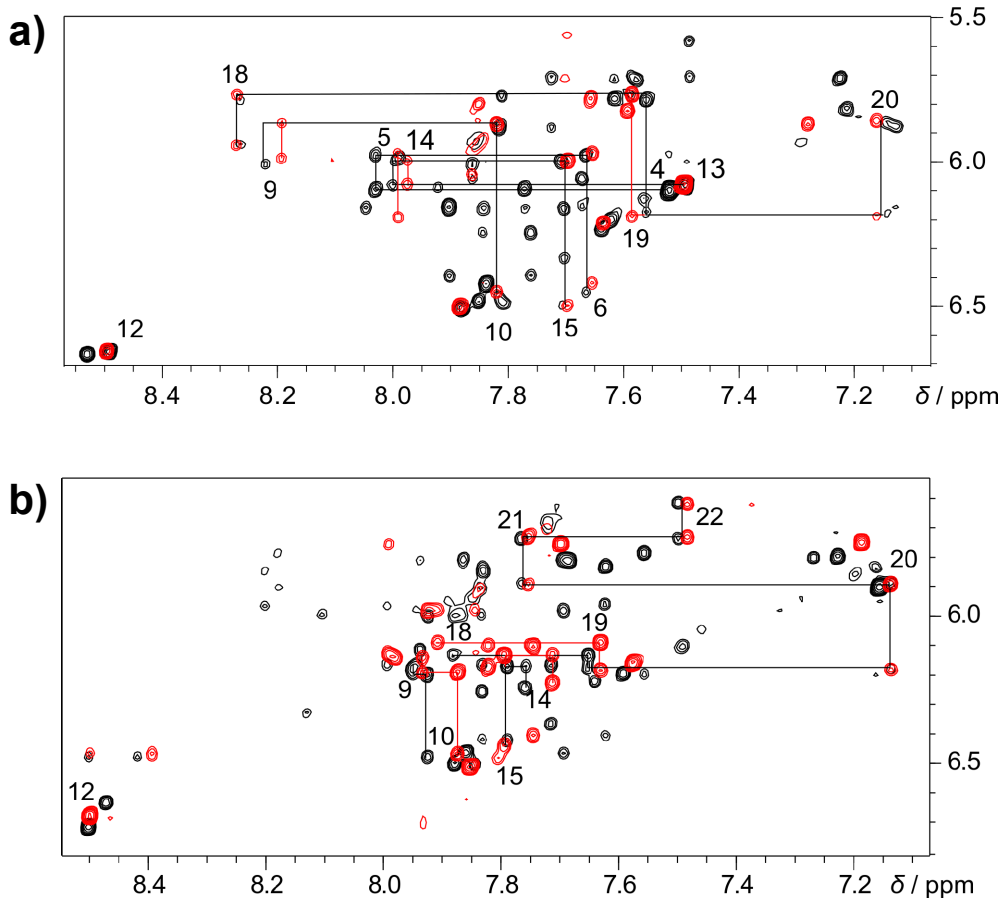



**Figure S4.** Superimposed H8/H6($\omega_2$)–H1'($\omega_1$) 2D NOE spectral regions of a) *MYC*(8,17) (black) and *MYC*(4,8,17) (red) and of b) *MYC*(4,8,13) (black) and *MYC*(4,13) (red) acquired at 30 °C with a mixing time $\tau_m$ = 300 ms (10 mM potassium phosphate, pH 7); NOE connectivities along the G-tracts are traced by vertical and horizontal lines. Red spectra represent spectra of quadruplexes with only a single conformation. *Syn* glycosidic torsion angles for G4 and G13 in *MYC*(8,17) are easily identified in a) through their strong intranucleotide H8-H1' contact.

**Table S1**. NMR restraints and structural statistics for the *MYC*(4,8,13,17) quadruplex

| | |
|---|---|
| **NOE-derived distance restraints** | |
| total | 262 |
| intra-residue | 143 |
| inter-residue | 119 |
| sequential | 91 |
| long-range | 28 |
| **other restraints** | |
| hydrogen bonds | 48 |
| torsion angles | 104 |
| G-quartet planarity | 36 |
| **structural statistics** | |
| pairwise heavy atom RMSD (Å) | |
| G-tetrad core | 0.58 ± 0.14 |
| all residues | 1.32 ± 0.46 |
| violations | |
| mean NOE restraint violations (Å) | 0.0047 ± 0.0004 |
| max. NOE restraint violations (Å) | 0.22 |
| deviations from idealized geometry | |
| bond length (Å) | 0.01 ± 0.00005 |
| bond angle (deg) | 2.28 ± 0.03 |

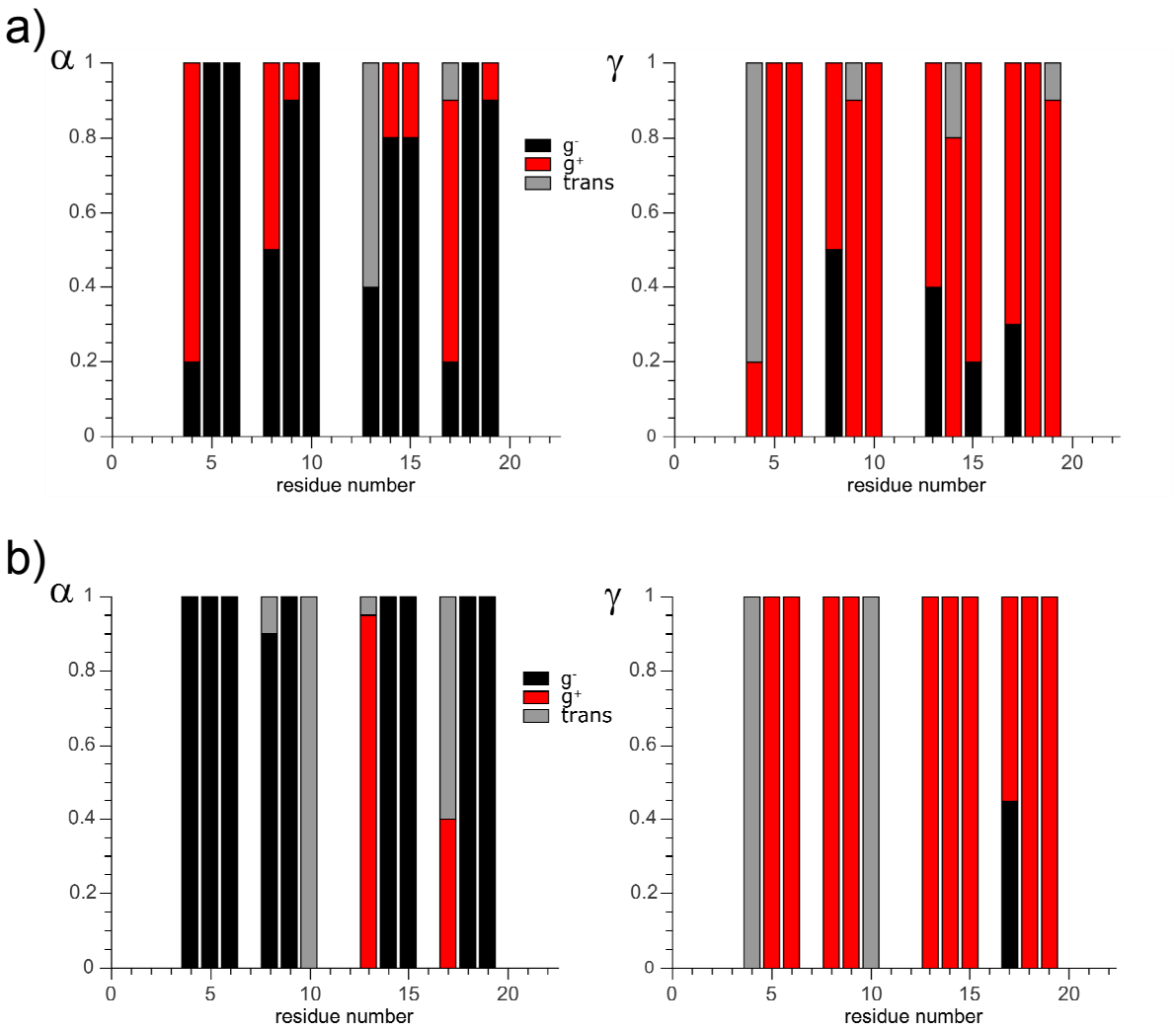



**Figure S5.** Distribution of α and γ conformers in a) *MYC*(4,8,13,17) and b) native *MYC* (pdb 1XAV). Fractional populations of $g^+$ (0-120°), t (120-240°), and $g^-$ (240-360°) for residues within the quadruplex core are based on the analysis of all final lowest-energy conformations.

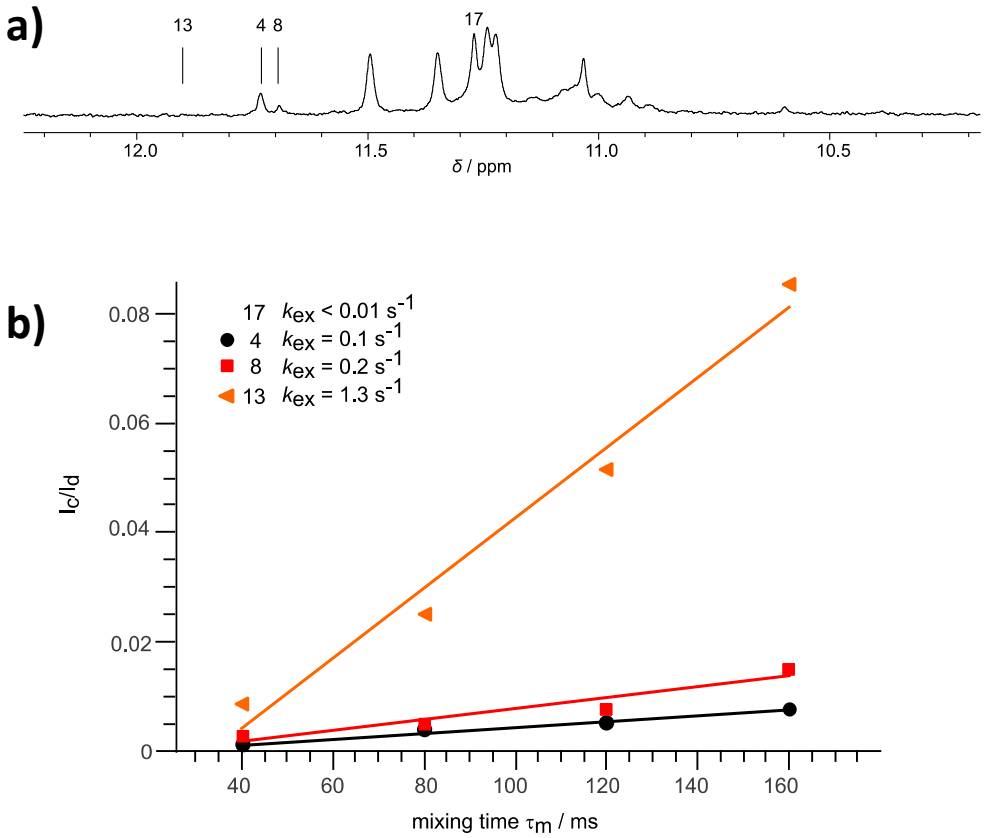



**Figure S6.** a) Differential decay of *MYC* imino proton signal intensities due to deuterium exchange with the solvent; the spectrum was acquired 1 min after the addition of deuterated buffer at 25 °C (10 mM potassium phosphate, pH 7). b) Plot of the intensity ratio of imino-$H_2O$ exchange crosspeaks to diagonal peaks as a function of mixing time $\tau_m$ and least-square linear fits. Intensities were determined from EASY-ROESY spectra at 50 °C.

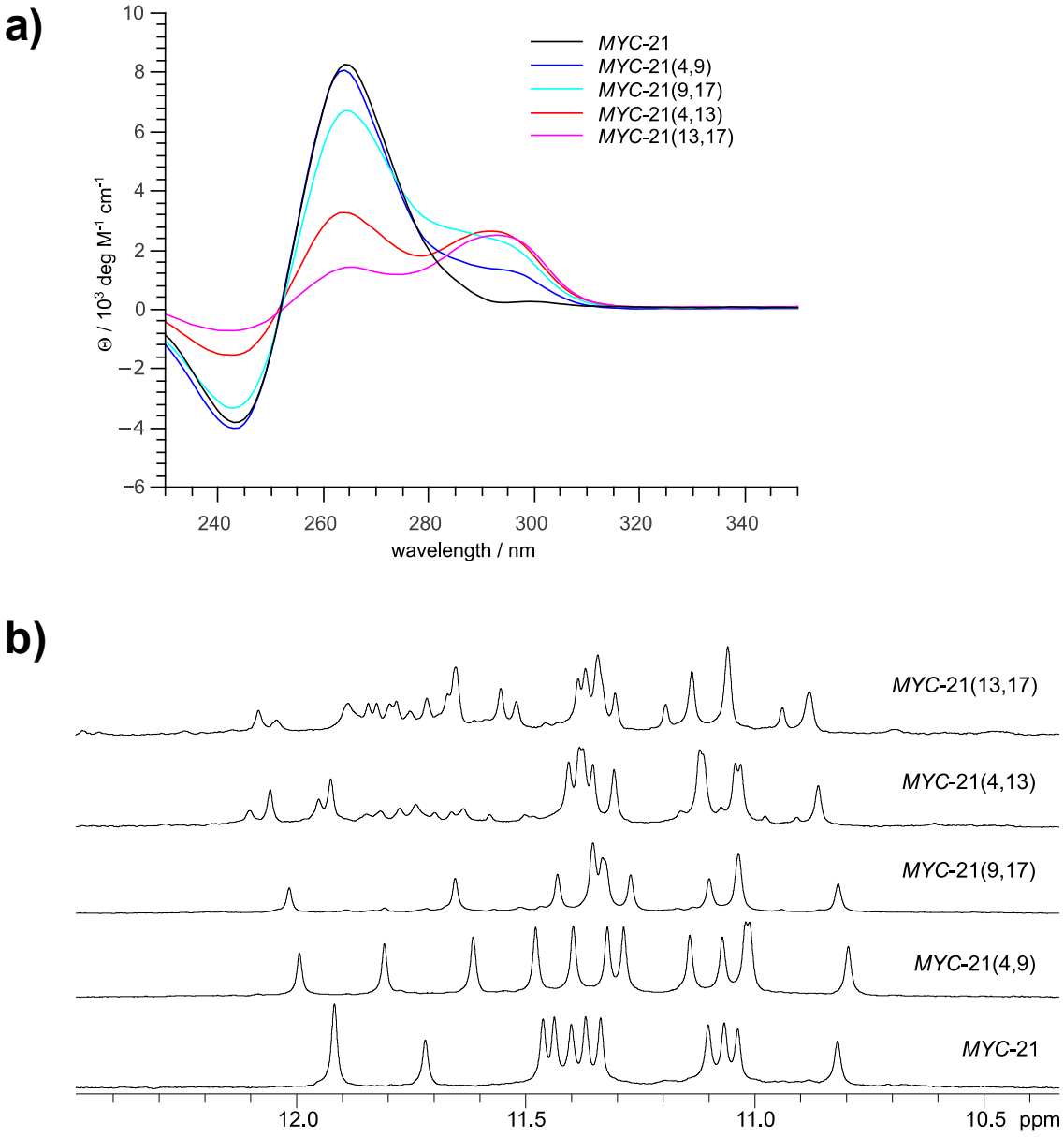



**Figure S7.** a) Superimposed CD spectra with molar ellipticities at 20 °C and b) imino proton spectral region at 30 °C of unmodified *MYC*-21 and 8-BrG substituted *MYC*-21 quadruplexes; numbers in parentheses denote the substitution sites.

# Publication III

G-Quadruplexes

International Edition: DOI: 10.1002/anie.201905372
German Edition: DOI: 10.1002/ange.201905372

# Duplex-Guided Refolding into Novel G-Quadruplex (3+1) Hybrid Conformations

*Beatrice Karg, Swantje Mohr, and Klaus Weisz**

***Abstract:*** *The oligomer d(GCGTG$_3$TCAG$_3$TG$_3$TG$_3$ACGC) with short complementary flanking sequences at the 5′- and 3′-ends was shown to fold into three different DNA G-quadruplex species. In contrast, a corresponding oligomer that lacks base complementarity between the two overhang sequences folds into a single parallel G-quadruplex. The three coexisting quadruplex structures were unambiguously identified and structurally characterized through detailed spectral comparisons with well-defined G-quadruplexes formed upon the deliberate incorporation of syn-favoring 8-bromoguanosine analogues into specific positions of the G-core. Two (3+1) hybrid structures coexist with the parallel fold and feature a novel lateral–propeller–propeller loop architecture that has not yet been confirmed experimentally. Both hybrid quadruplexes adopt the same topology and only differ in their pattern of anti→syn transitions and tetrad stackings.*

G-quadruplexes (G4s) are formed by guanine-rich nucleic acids that fold into a four-stranded structure with a stacked arrangement of two to four guanine tetrads. Their noticeable polymorphism derives from the type of loops adopted by intervening sequences, from the relative orientations of the G-tracts forming the columns of the quadruplex core, and from the pattern of glycosidic conformations along the tracts.[1] Although rules have emerged that correlate sequences to the stability of particular G4 topological features,[2] reliable predictions of G4 conformations are challenging because of our still limited knowledge of the various interactions that determine G4 folding. Deliberately engineering specific folds will therefore not only provide an expanded collection of G-quadruplex scaffolds to be used for the increasing number of technological applications, for example, as aptamers or parts of nanoarchitectures,[3,4] but will also contribute to a better understanding of the forces acting along the G4 folding pathway.[5]

The ssQ oligomer d(GCGTG$_3$TCAG$_3$TG$_3$TG$_3$CTCA) with 5′- and 3′-overhangs and one three-nucleotide (3-nt) as well as two 1-nt intervening sequences between its four G-tracts is based on a truncated variant of the guanine-rich strand of the nuclease hypersensitive element III$_1$ in the *c-MYC* promoter. The latter has been shown to fold into a very robust parallel G-quadruplex with three short propeller-type loops.[6] However, the 3-nt intervening sequence in ssQ is expected to also allow formation of a lateral loop to yield a novel (3+1) hybrid topology with a first lateral followed by two propeller loops (Figure 1). Notably, such a fold has been proposed to occur as a minor species for some G4-forming sequences and very recently has also been suggested based on specific cleavage patterns for modified telomeric quadruplexes.[7–9] Yet, despite the absence of apparent steric restrictions, a corresponding topology has been elusive to date and not yet unambiguously confirmed, let alone structurally characterized in detail.

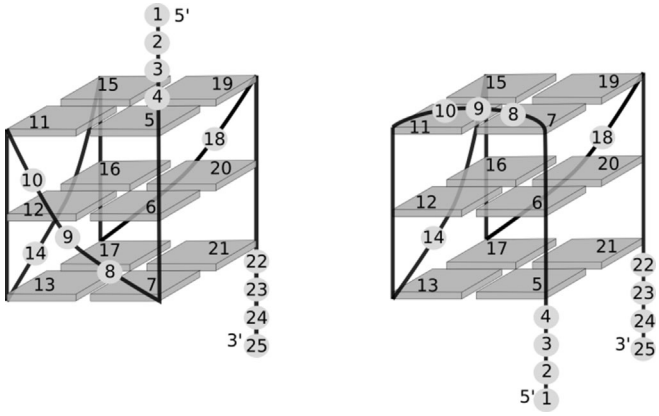


***Figure 1.*** Parallel ssQ quadruplex adopting an all-*anti* G-core (left) and fold of a putative (3+1) hybrid topology with inversion of the 5′-terminal G-tract (right).

In line with its resistance to form a hybrid-type G4 species with a flipped 5′-terminal G-tract, ssQ folds into a well-defined parallel G4 structure as shown by CD and NMR spectral analysis (Figure 2 and Figure S1 in the Supporting Information). In contrast to the parallel structure, the 5′- and 3′-flanking sequences would be positioned at the same side of the quadruplex core in case of a putative strand-inverted species (see Figure 1). We therefore reasoned that base pairing of the two short overhang sequences may provide for an additional driving force for refolding into a hybrid-type structure. In fact, an almost perfect fit for a duplex extension when spanning the quadruplex wide groove allows for Watson–Crick base pairing without any additional spacers.[10,11] This should give extra stabilization when compared to unpaired overhangs or to a conceivable orthogonal duplex–quadruplex arrangement with oppositely located flanking sequences, which would require additional unpaired spacer residues.

[*] B. Karg, S. Mohr, Prof. Dr. K. Weisz
Institut für Biochemie, Universität Greifswald
Felix-Hausdorff-Str. 4, 17487 Greifswald (Germany)
E-mail: weisz@uni-greifswald.de

Supporting information and the ORCID identification number(s) for the author(s) of this article can be found under: https://doi.org/10.1002/anie.201905372.

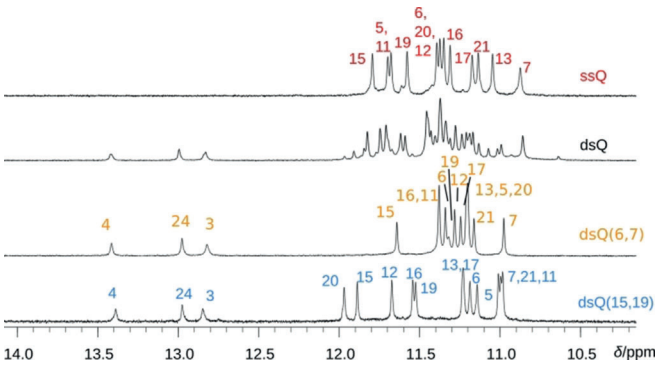



***Figure 2.*** Imino proton spectral region of ssQ and the modified sequences dsQ, dsQ(6,7), and dsQ(15,19) in $K^+$ solution at 30 °C.

To determine whether duplex formation based on complementary flanking sequences will enable the formation of a strand-flipped hybrid-type G-quadruplex, the folding of a sequence dsQ d(GCGTG$_3$TCAG$_3$TG$_3$TG$_3$ACGC) with a modified 3′-overhang was evaluated. Indeed, the presence of one AT and two GC Watson–Crick imino proton signals at low field suggests additional duplex formation between neighboring 5′- and 3′-overhang sequences (Figure 2). However, the Hoogsteen imino proton spectral region between 12.0 and 10.8 ppm clearly shows folding into more than one quadruplex structure. A closer inspection suggests the coexistence of three differently populated quadruplexes. Although overwhelming at first, the dispersion and resolution of signals are encouraging when trying to identify G4 species based on a spectral comparison with appropriate reference spectra. Thus, a native ssQ-like parallel fold was revealed by striking similarities in the position and pattern of correlation peaks in overlays of the corresponding 2D NOE and especially $^1$H–$^{13}$C HSQC spectra of dsQ and ssQ (Figure 3). Except for the different 3′-flanking residues, the latter spectra show nearly superimposable C6/C8–H6/H8 correlations for most ssQ cross-peaks. These include T4 of the 5′-overhang as well as T8, C9, and A10 of the first loop. The corresponding H8 protons of A10 are significantly deshielded, which is typical for solvent exposure and their location within a propeller loop.[6,7]

Assuming duplex formation of additional species to be associated with a strand flip and the formation of a (3+1) hybrid structure, G residues of the first G-tract are expected to undergo *anti*→*syn* conformational transitions. Consequently, in trying to further enforce a strand inversion, two 8-bromoguanosine ($^{Br}$G) analogues favoring a *syn* glycosidic conformation were substituted at position 6 and 7 of the first G-tract to give dsQ(6,7) with the sequence d(GCGTG$^{Br}$G$^{Br}$GTCAG$_3$TG$_3$TG$_3$ACGC). As shown in Figure 2, the imino proton spectral region of dsQ(6,7) indicates its folding into a single well-defined structure with three observable Watson–Crick imino protons of the same chemical shift as seen for the dsQ mixture. Whereas the imino resonance of the terminal base pair is unobservable due to fraying effects and fast solvent exchange, the observation of an AT imino proton next to the G4 core clearly corroborates refolding into a topology with adjacent overhangs.

Unambiguous proton assignments for dsQ(6,7) without $^{15}$N labeling were enabled essentially through inter-residual NOE contacts also involving overhang and intervening sequences as well as through $^1$H–$^{13}$C HSQC and HMBC experiments at natural abundance (for details see the Supporting Information and Figure S3).[12] Analysis of the 2D NMR spectra demonstrated a flip of the first G-tract as evidenced by sequential (G H1)$_n$–(H2′/H2″)$_{n-1}$ NOE contacts for all three residues of the first G-tract, which is in line with an all-*syn* strand,[13,14] and by the NOE contacts of a 3-nt lateral loop with both the preceding and following G-tract.

Extracted distance and torsion angle restraints were used for structure calculations of dsQ(6,7) (for the calculations and structural statistics, see the Supporting Information and Table S2). As seen in Figure 4, the quadruplex wide groove smoothly merges with the duplex minor groove with only small adjustments to the backbone. Thus, the phosphate-to-phosphate distance is reduced following the quadruplex–duplex junction from 19.2 to 18.3 Å. Also, the neighboring AT

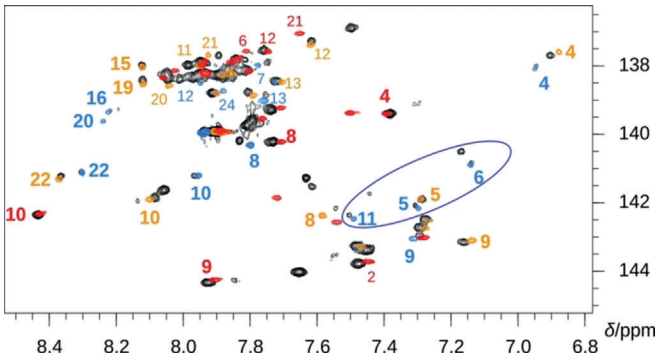



***Figure 3.*** Superimposed $^1$H–$^{13}$C HSQC spectra acquired in $K^+$ solution at 30 °C showing H8/H6–C8/C6 correlations of dsQ (black), ssQ (red), dsQ(6,7) (orange), and dsQ(15,19) (blue). Cross-peaks of dsQ assigned to the particular fold of a sequence analogue based on overlapping signals of nearly identical chemical shifts are labeled by color-coded residue numbers. Some representative cross-peaks are highlighted in bold whereas other cross-peaks have been left unlabeled for clarity; encircled in blue are *syn*-guanosine residues.

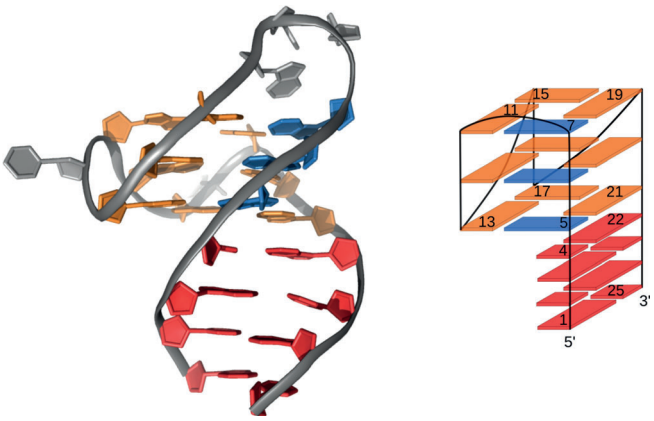



***Figure 4.*** Side view of a representative low-energy solution structure of dsQ(6,7) (left, PDB ID 6R9K) and schematic representation of the hybrid-1R conformation with some of the residues numbered (right). The *anti* and *syn* residues of the G-core are shown in orange and blue, respectively.

base pair efficiently stacks on the outer G-quartet. The novel three-layered (3+1) G-quadruplex can be classified by a +(lpp) looping topology, with one lateral followed by two propeller loops in a clockwise direction.[15] As it is related to the well-known hybrid-1 form by changing lateral into propeller loops and vice versa but also by a formally reversed progression of loops, we term this topology "hybrid-1R". It should be noted that the flipped 5′-terminal G-tract is composed of three *syn* residues, resulting in a (3+1) G-core with solely homopolar stacking interactions in line with its type I CD signature (Figure S1).[16] A10 at the 3′-end of the lateral loop spanning the G4 narrow groove stacks on the lower tetrad whereas T8 at the 5′-end is forced to be more exposed to solvent.

Remarkably, the lateral loop in the hybrid-1R structure starts with an unusual *syn*-G located at the end of the 5′-terminal G-tract.[17] Only in rare cases have *syn* residues situated in a central all-*syn* strand of G-modified quadruplex structures been reported to precede lateral loops.[18,19] However, this distinct conformation does not seem to noticeably affect local geometries, and no major structural peculiarities are apparent. On the other hand, a conspicuous broadening of the G19 imino proton hydrogen-bonded to *syn*-G7 in dsQ(6,7) through fast solvent exchange may be associated with a more flexible outer tetrad (Figure S6).

As for ssQ, overlays of 2D NMR spectra of dsQ and dsQ(6,7) clearly identify the hybrid-1R structure of dsQ(6,7) to coexist in the dsQ quadruplex mixture. Thus $^{1}$H–$^{13}$C HSQC spectra reveal several superimposable cross-peaks of diagnostic value such as T4 and A22 following the 5′- and 3′-end of the G-core and highfield-shifted residues A10 and C9 typical of their location in a lateral loop. Also, as suggested by their highfield and downfield C8 chemical shifts,[18,20,21] some distinct G-core residues such as G15/G19 as well as G5 are found in *anti* and *syn* conformation, respectively.

Noticeable $^{13}$C downfield-shifted G C8–H8 correlations (Figure 3) and in particular strong intra-nucleotide G H8–H1′ NOE intensities (Figure S2) suggest a total of eight *syn*-G residues in the dsQ mixture. This leaves five *syn*-guanosines for a third and still unknown G4 species. It should be noted that the identified hybrid-1R topology contains two consecutive *syn–syn* steps associated with unfavorable stacking interactions.[22,23] With no restrictions imposed by $^{Br}$G substitutions, it is therefore conceivable for the dsQ sequence to fold into the same +(lpp) looping architecture while simultaneously changing the polarity of its outer tetrad adjoining the lateral loop. This would be accompanied by one *syn*→*anti* and three *anti*→*syn* transitions, resulting in five *syn*-Gs that participate in four favorable *syn–anti* steps and only one disfavored *syn–syn* step.

To select for such a putative structure, two $^{Br}$G analogues were substituted in dsQ at positions 15 and 19, both being *anti* for dsQ(6,7) but proposed to adopt a *syn* conformation for a potential tetrad-inverted conformation. The imino proton spectral region of the dsQ(15,19) sequence d(GCGTG$_3$TCAG$_3$T$^{Br}$**G**GGT$^{Br}$**G**GGACGC) exhibits only a single species with three Watson–Crick imino protons of the same chemical shift as found for dsQ(6,7), again confirming additional duplex formation with a flipped G-tract (Figure 2). Standard 2D NMR experiments allow for unambiguous resonance assignments (for details see the Supporting Information and Figure S4). The most notable observations that confirm the proposed hybrid structure include a typical *syn–anti* sequential H8–H1′ NOE rectangular pattern for the G6–G7 step, sequential (G H1)$_n$–(H2′/H2″)$_{n-1}$ NOE contacts for G6 and G5, pointing to their *syn* conformation, and NOE contacts of the lateral loop with neighboring G-tracts.

Structure calculations for dsQ(15,19) employing NMR distance and torsion angle restraints yield a three-dimensional structure termed hybrid-1R′ (Figure 5) with high similarities to the hybrid-1R structure of dsQ(6,7) (for calculations and structural statistics, see the Supporting Information and Table S2). Both forms feature the same +(lpp) topology with rearrangement of the first propeller into a lateral loop as induced by the strand flip (Figure S5). The hybrid-1R′ G-core is built from one *syn–syn–anti* and three *syn–anti–anti* columns. Compared to the hybrid-1R conformation, adenine 10 within the lateral loop is slightly shifted to maintain efficient stacking with the polarity-inverted tetrad. Also, the hybrid-1R′ G4 features homopolar as well as heteropolar stacking with efficient interactions between the five-membered purine rings for the *syn*/*anti* steps, yielding a type II CD spectrum with two positive bands at $\lambda = 265$ nm and 295 nm (Figure S1 and Figure S5C).

With the availability of additional reference spectra for a second strand-inverted quadruplex and in noting that like dsQ(6,7) some correlations are lacking for dsQ(15,19) because of its two $^{Br}$G substitutions, additional *syn–anti* patterns in the NOE spectra as well as *syn*-residues in $^{1}$H–$^{13}$C HSQC spectra of dsQ can easily be traced to a dsQ(15,19)-like hybrid-1R′ structure with several diagnostic cross-peaks located at nearly identical chemical shifts (Figure 3 and Figure S2). Integration of the well resolved T4 H6 resonances in the dsQ mixture allows the ratio of the parallel, hybrid-1R, and hybrid-1R′ forms to be estimated as 2:1:0.4. This was further confirmed by $^{15}$N edited spectra on a dsQ sequence specifically $^{15}$N-labeled at loop position A10 (Figure S7). The

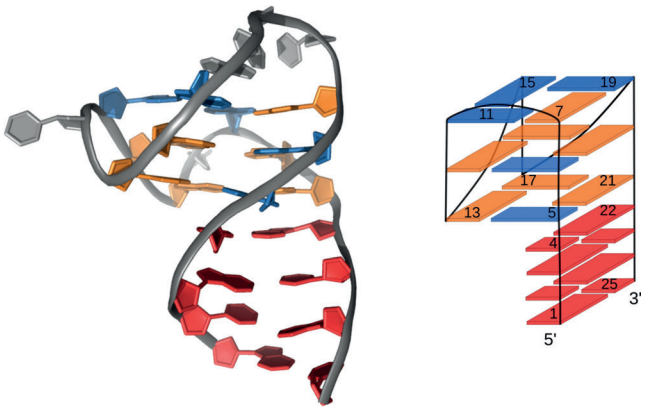



***Figure 5.*** Side view of a representative low-energy solution structure of dsQ(15,19) (left, PDB ID 6R9L) and schematic representation of the hybrid-1R′ conformation with some of the residues numbered (right). The *anti* and *syn* residues of the G-core are shown in orange and blue, respectively.

population of the coexisting species indicates a higher thermodynamic stability of hybrid-1R compared to the hybrid-1R′ form. Thus given their closely related conformational features, the lower energetic penalty imposed by a smaller number of *syn*-G residues in the hybrid-1R structure seems to outweigh the larger number of energetically unfavorable *syn*–*syn* steps.

The even higher stability of the native parallel G4 is also reflected by DSC measurements of the parallel ssQ quadruplex and the dsQ mixture, the latter showing a slightly broadened unresolved quadruplex melting transition lowered by 4 °C when compared to the ssQ G4 (Table S1 and Figure S8). On the other hand, melting of the duplex extension for the two (3+1) hybrid structures is indicated by a second low-temperature melting process about 20 °C below the melting point of the quadruplex core. Yet, the additional formation of base pairs between the short overhang sequences allows for the disfavored fold to effectively compete with the parallel topology. It might therefore be speculated that fast transient base pairing of the overhang domains preorganizes the sequence to fold into a hybrid topology with the duplex serving as a kinetic trap.[24] Similarly, prefolding into a homodimeric structure through base pairing of the short 3′-tails of a kit* sequence has recently been suggested to drive folding into a single G-quadruplex.[25]

In addition to base pairing between overhang sequences, internal hairpin duplexes formed along quadruplex loops have been shown to stabilize quadruplex architectures and seem to be a more recurrent motif in genomic DNA.[26–28] Likewise, the here reported two novel three-layered (3+1) hybrid quadruplexes with one lateral and two propeller loops may be conceivable in G-rich genomic sequences if short single-stranded regions flanking the two outermost functional G-tracts allow for Watson–Crick base pairing. As demonstrated by previous studies, the duplex–quadruplex junction formed may also be a unique interaction site for various ligands.[29,30] On the other hand, the principle of guiding quadruplex folding through additional interactions in duplex extensions may support a rational design of an ever increasing toolbox of quadruplex topologies for their use in technological applications.

## Acknowledgements

We thank the Deutsche Forschungsgemeinschaft for financial support (grant no. WE 1933/15-1).

## Conflict of interest

The authors declare no conflict of interest.

**Keywords:** glycosidic conformation · G-quadruplexes · NMR spectroscopy · quadruplex–duplex junctions · topology

[1] Y. Chen, D. Yang, *Current Protocols in Nucleic Acid Chemistry*, Wiley, Hoboken, **2012**, pp. 17.5.1 – 17.5.17.
[2] A. Guédin, P. Alberti, J. L. Mergny, *Nucleic Acids Res.* **2009**, *37*, 5559 – 5567.
[3] B. Gatto, M. Palumbo, C. Sissi, *Curr. Med. Chem.* **2009**, *16*, 1248 – 1265.
[4] U. Feldkamp, C. M. Niemeyer, *Angew. Chem. Int. Ed.* **2006**, *45*, 1856 – 1876; *Angew. Chem.* **2006**, *118*, 1888 – 1910.
[5] L. Haase, B. Karg, K. Weisz, *ChemBioChem* **2019**, *20*, 985 – 993.
[6] A. Ambrus, D. Chen, J. Dai, R. Jones, D. Yang, *Biochemistry* **2005**, *44*, 2048 – 2058.
[7] A. T. Phan, K. N. Luu, D. J. Patel, *Nucleic Acids Res.* **2006**, *34*, 5715 – 5719.
[8] M. L. Greco, A. Kotar, R. Rigo, C. Cristofari, J. Plavec, C. Sissi, *Nucleic Acids Res.* **2017**, *45*, 10132 – 10142.
[9] A. Beniaminov, A. Shchyolkina, D. Kaluzhny, *Biochimie* **2019**, *160*, 122 – 128.
[10] K. W. Lim, A. T. Phan, *Angew. Chem. Int. Ed.* **2013**, *52*, 8566 – 8569; *Angew. Chem.* **2013**, *125*, 8728 – 8731.
[11] K. W. Lim, Z. J. Khong, A. T. Phan, *Biochemistry* **2014**, *53*, 247 – 257.
[12] M. Adrian, B. Heddi, A. T. Phan, *Methods* **2012**, *57*, 11 – 24.
[13] C. Liu, B. Zhou, Y. Geng, D. Yan Tam, R. Feng, H. Miao, N. Xu, X. Shi, Y. You, Y. Hong, et al., *Chem. Sci.* **2019**, *10*, 218 – 226.
[14] J. Brčić, J. Plavec, *Nucleic Acids Res.* **2018**, *46*, 11605 – 11617.
[15] M. Webba da Silva, *Chem. Eur. J.* **2007**, *13*, 9738 – 9745.
[16] A. I. Karsisiotis, N. M. Hessari, E. Novellino, G. P. Spada, A. Randazzo, M. Webba da Silva, *Angew. Chem. Int. Ed.* **2011**, *50*, 10645 – 10648; *Angew. Chem.* **2011**, *123*, 10833 – 10836.
[17] A. I. Karsisiotis, C. O'Kane, M. Webba da Silva, *Methods* **2013**, *64*, 28 – 35.
[18] J. Dickerhoff, K. Weisz, *Angew. Chem. Int. Ed.* **2015**, *54*, 5588 – 5591; *Angew. Chem.* **2015**, *127*, 5680 – 5683.
[19] S. Bielskutė, J. Plavec, P. Podbevšek, *J. Am. Chem. Soc.* **2019**, *141*, 2594 – 2603.
[20] K. Greene, Y. Wang, D. Live, *J. Biomol. NMR* **1995**, *5*, 333 – 338.
[21] J. M. Fonville, M. Swart, Z. Vokáčová, V. Sychrovský, J. E. Šponer, J. Šponer, C. W. Hilbers, F. M. Bickelhaupt, S. S. Wijmenga, *Chem. Eur. J.* **2012**, *18*, 12372 – 12387.
[22] X. Cang, J. Šponer, T. E. Cheatham III, *Nucleic Acids Res.* **2011**, *39*, 4499 – 4512.
[23] J. Šponer, A. Mládek, N. Špačková, X. Cang, T. E. Cheatham III, S. Grimme, *J. Am. Chem. Soc.* **2013**, *135*, 9785 – 9796.
[24] 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 – 1263.
[25] A. Kotar, R. Rigo, C. Sissi, J. Plavec, *Nucleic Acids Res.* **2019**, *47*, 2641 – 2653.
[26] E. Butovskaya, B. Heddi, B. Bakalar, S. N. Richter, A. T. Phan, *J. Am. Chem. Soc.* **2018**, *140*, 13654 – 13662.
[27] K. W. Lim, P. Jenjaroenpun, Z. J. Low, Z. J. Khong, Y. S. Ng, V. A. Kuznetsov, A. T. Phan, *Nucleic Acids Res.* **2015**, *43*, 5630 – 5646.
[28] K. W. Lim, T. Q. N. Nguyen, A. T. Phan, *J. Am. Chem. Soc.* **2014**, *136*, 17969 – 17973.
[29] S. Asamitsu, S. Obata, A. T. Phan, K. Hashiya, T. Bando, H. Sugiyama, *Chem. Eur. J.* **2018**, *24*, 4428 – 4435.
[30] A. T. Phan, V. Kuryavyi, J. C. Darnell, A. Serganov, A. Majumdar, S. Ilin, T. Raslin, A. Polonskaia, C. Chen, D. Clain et al., *Nat. Struct. Mol. Biol.* **2011**, *18*, 796 – 804.

Manuscript received: April 30, 2019
Accepted manuscript online: May 28, 2019
Version of record online: ■■ ■■, ■■■■

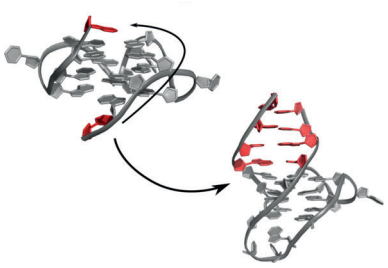
  

# Communications

## G-Quadruplexes

B. Karg, S. Mohr,
K. Weisz* ——————— ■■■■–■■■■

Duplex-Guided Refolding into Novel G-Quadruplex (3+1) Hybrid Conformations



**A parallel G-quadruplex** was guided into a novel (3+1) hybrid topology with a lateral–propeller–propeller loop architecture and two different *syn*/*anti* patterns within the G-core by introducing additional Watson–Crick hydrogen bonding interactions between complementary 5′- and 3′-flanking sequences.

Supporting Information

69451 Weinheim, Germany

# Duplex-Guided Refolding into Novel G-Quadruplex (3+1) Hybrid Conformations

Beatrice Karg, Swantje Mohr, and Klaus Weisz*

## Table of Contents

## Experimental Procedures

### Materials

Unmodified and HPLC-purified modified DNA oligonucleotides substituted with 8-$^{Br}$G and $^{15}$N-dA (10% $^{15}$N enrichment) were purchased from TIB MOLBIOL (Berlin, Germany) and IBA (Göttingen, Germany). Following ethanol precipitation, concentrations were determined spectrophotometrically by measuring their absorbance $A_{260}$ at 90 °C in salt-free $H_2O$ using a molar extinction coefficient $\varepsilon_{260}$ given by the manufacturer. Samples were obtained by dissolving oligonucleotides in a low salt buffer with 10 mM potassium phosphate, pH 7.0. Prior to measurements, the samples were annealed by heating to 90 °C followed by cooling to room temperature. Final concentrations of oligonucleotides were 5 $\mu$M for optical, 50 $\mu$M for calorimetric, and between 0.12 and 1 mM for NMR measurements.

### Circular dichroism

CD spectra were acquired with a Jasco J-810 spectropolarimeter (Jasco, Tokyo, Japan). Spectra of oligonucleotides were recorded at 20 °C with a bandwidth of 1 nm, a scanning speed of 50 nm/min and 5 accumulations.

### UV melting

Melting curves were acquired in triplicate between 15 and 89 °C on a Cary 100 spectrophotometer equipped with a Peltier temperature control unit (Varian, Darmstadt, Germany). Using quartz cuvettes of 10 mm path length, data points were recorded in 0.5 °C intervals with heating rates of 0.2 °C/min. Melting temperatures $T_m$ were determined by the maximum in a first derivative plot of the absorbance at 295 nm as a function of temperature.

### Differential Scanning Calorimetry

DSC measurements were performed on a VP-DSC instrument (Malvern Instruments, Malvern, Great Britain) using 50 $\mu$M oligonucleotide in a buffer with 10 mM potassium phosphate (pH 7). The sample was heated from 20 °C to a maximum of 80 °C with a scan rate of 10°C/h. A buffer versus buffer scan was subtracted from the sample scan prior to baseline correction. The melting temperature $T_m$ was determined from the maximum of the DSC peak in three independent measurements with four heating cycles each.

### NMR Experiments

All NMR spectra were acquired on a Bruker Avance 600 MHz spectrometer equipped with an inverse quadruple resonance cryoprobehead and z-field gradients. Data were processed with the Topspin 4.0.3 software and analyzed with CcpNmr Analysis.[1] Typically, NOESY experiments in 90% $H_2O$/10% $D_2O$ were performed at 30 °C with mixing times from 80 ms to 300 ms and a spectral width of 10 kHz. 2K×1K data points with 16-64 transients per $t_1$ increment and a recycle delay of 2 s were collected in $t_2$ and $t_1$. DQF-COSY spectra were acquired in $D_2O$ with a 3-9-19 solvent suppression scheme that was also used for the phase-sensitive $^1$H-$^{13}$C HSQC experiments in 90% $H_2O$/10% $D_2O$. The latter were optimized for a $^1J$(C,H) coupling of 170-220 Hz. $^{13}$C chemical shifts were referenced relative to DSS by using the indirect referencing method.

### NMR Structure Calculations

The RED software was used to calculate the partial atomic charges for the modified 8-bromoguanosine residues.[2] Ten starting models from 300 extended structures were generated with a distance geometry simulated annealing protocol in Xplor-NIH 2.48.[3] Subsequently, restrained simulated annealing was performed with the Amber15 software in implicit water using the parmbsc0 force field including the $\chi_{OL4}$, $\varepsilon\zeta_{OL1}$, and $\beta_{OL1}$ corrections.[4-7] The ten lowest energy quadruplex structures were selected for refinement in water and neutralized with two $K^+$ cations manually placed between the tetrads. Subsequently, the system was solved with water molecules (TIP3P) in a truncated octahedral box of 10 Å. For the final simulation of 1 ns duration only restraints for NMR-derived distances, torsion angles, Watson-Crick hydrogen bonds, and Hoogsteen hydrogen bonds were employed. Finally, the trajectories were averaged over the last 40 ps and shortly minimized in vacuo for 50 steps.

## Results

**Table S1.** G4 forming sequences with G-tracts underlined and their melting temperatures (10 mM $KP_i$ buffer, pH 7).

| name | sequence 5' → 3'[a] | $T_m$ (°C) |
|---|---|---|
| ssQ | GCGTGGGTCAGGGTGGGTGGGCTCA | 60.7±0.3[b] |
| dsQ | GCGTGGGTCAGGGTGGGTGGGACGC | 36.7±1.1<br>56.6±1.5[b] |
| dsQ(6,7) | GCGTG$^{Br}$G$^{Br}$GTCAGGGTGGGTGGGACGC | 58.3±0.6[c] |
| dsQ(15,19) | GCGTGGGTCAGGGT$^{Br}$GGGT$^{Br}$GGGACGC | 44.9±0.8[c] |

[a] $^{Br}$**G** = 8-Bromo-2'-deoxyguanosine. [b] Determined by DSC in triplicate. [c] Determined by UV in triplicate.

**Figure S1.** CD spectra of ssQ, dsQ, dsQ(6,7), and dsQ(15,19) quadruplexes at 20 °C in $K^+$ solution (pH 7). Despite the coexistence of a minor (3+1) hybrid structure with heteropolar tetrad stacking, the CD signature for dsQ is largely governed by the parallel G4 and the dsQ(6,7)-type hybrid with their exclusive homopolar stacking interactions.

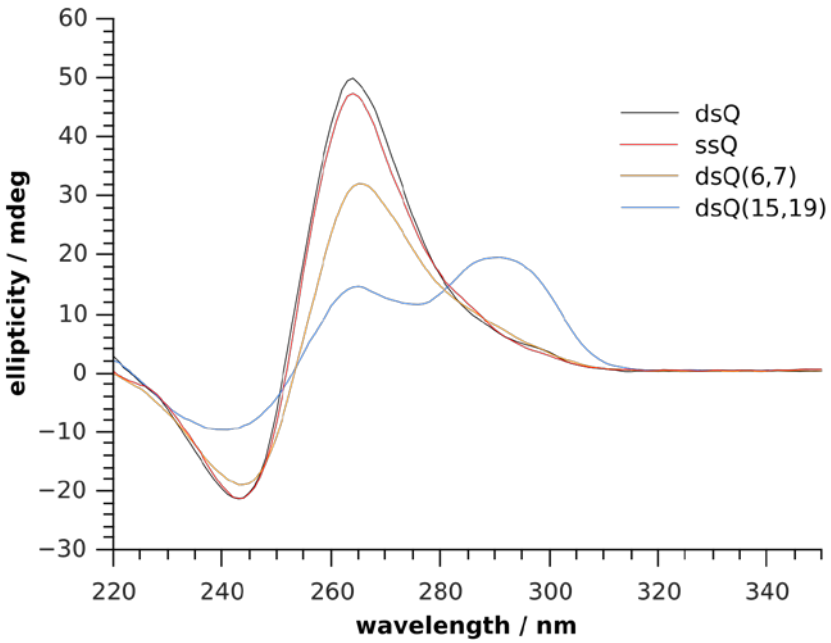
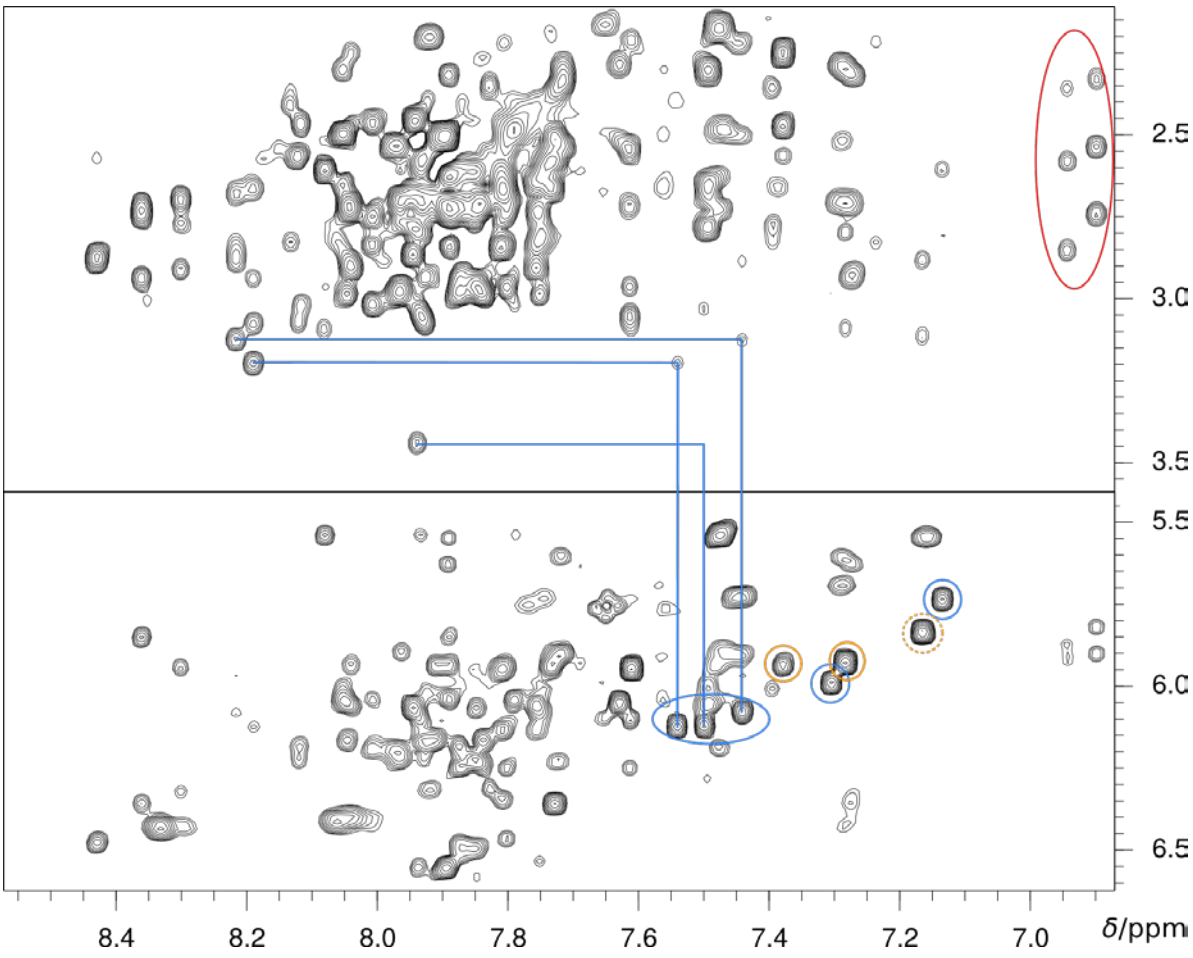


**Figure S2.** H6/H8-H1' and H6/H8-H2'/H2" regions of a 2D NOE spectrum of dsQ acquired at 30 °C with a mixing time of 300 ms. Conspicuous high-intensity H8-H1' crosspeaks of *syn*-Gs are encircled orange and blue for hybrid-1R and hybrid-1R' species, respectively. Blue lines trace sequential contacts to 3'-neighboring *anti*-G residues for the hybrid-1R' structure. A typical rectangular H8-H1' pattern for *syn-anti* steps is mostly hidden due to signal overlap. Two sets of T4 H6 crosspeaks within the duplex overhang are encircled in red.



### Resonance Assignments for dsQ(6,7)

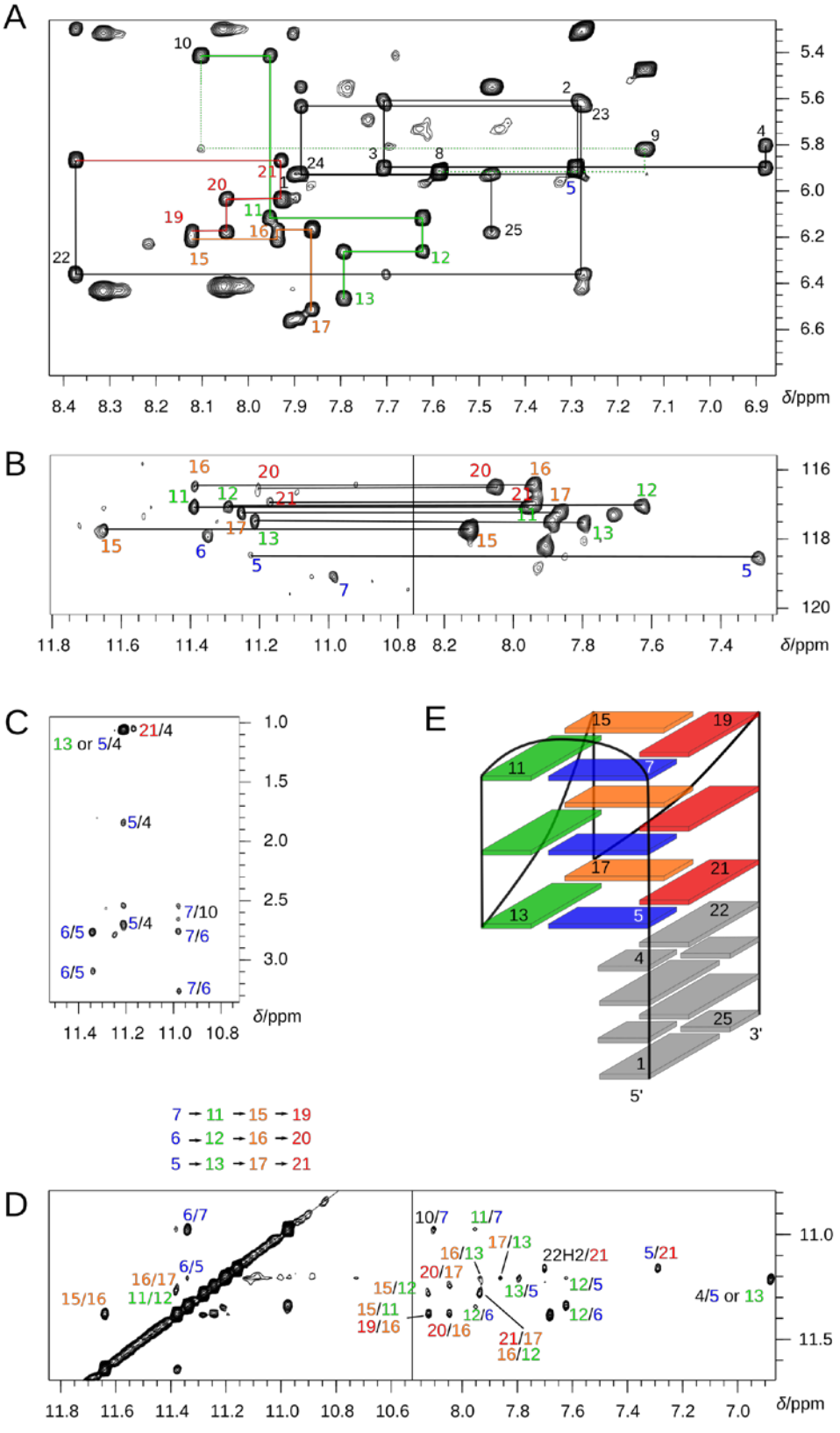
Looking at the imino proton spectral region, the presence of Watson-Crick GC and AT iminos with NOE contacts between G H1 and C amino as well as T H3 and A H2 protons strongly suggest formation of a duplex from the 4-nt complementary 5'- and 3'- termini of dsQ(6,7) (Figure 1). With the exception of G1, all duplex imino protons can be observed. A22 H8 shows sequential contacts to G21 sugar protons and thus into the fourth G4 tract with its uninterrupted H8-H1' connectivities, indicative of a G19(*anti*)-G20(*anti*)-G21(*anti*) glycosidic bond conformational pattern (Fig. S3A, red). Two more *anti-anti-anti* G tracts can be found, one of those identified to be the second G tract through its sequential contacts to A10 of the TCA loop (Figure S3A, orange and green). In contrast, no sequential contacts can be observed between the 5'-terminal T8 of this loop and $^{Br}$G7. The yet unidentified G residue G5 shows no sequential H8-H1' contacts. However, a strong additional intra-nucleotide H1'-H8 contact (Figure S3A) together with a single characteristic $^{13}$C8 chemical shift as observed in $^{1}$H-$^{13}$C HSQC spectra (Figure 3) point to a *syn* glycosidic conformation for this residue.

Corresponding imino protons can be unambiguously related to their H8 through JR-HMBC experiments[8] except for analogs $^{Br}$G6 and $^{Br}$G7 lacking H8 but also for G19 with a missing correlation of H8 to its imino proton (Figure S3B) . Interestingly, contacts to T4 methyl as well as H2'/H2" sugar protons can be observed for G5 H1 overlapping with the G13 imino (Figure S3C). Such sequential contacts are only enabled for the guanine base of G5 in a *syn* conformation. Two additional iminos with corresponding sequential H2'/H2" NOE connectivities but yet unidentified due to their missing H8 correlations must be attributed to $^{Br}$G6 and $^{Br}$G7 analogs. The latter is unambiguously identified through its additional contact to an A10 sugar proton. Taken together, this unusual pattern of $(H2'/H2")_n$-$(H1)_{n+1}$ NOE contacts along G5-$^{Br}$G6-$^{Br}$G7 is only compatible with an all-*syn* conformation for the first G tract.

All of the observed imino-imino contacts are sequential, indicating exclusive homopolar tetrad stacking (Figure S3D left). Fom H1-H8 contacts, H-bonding can be observed within a G11-G15-G19 tetrad interrupted by the $^{Br}$G7 residue (Figure S3D right). The same applies to G12-G16-G20 with $^{Br}$G6 being the fourth residue within this central tetrad. An uninterrupted cyclic hydrogen bond pattern is observed for the unmodified G5-G13-G17-G21 tetrad bordering the duplex overhang. Consequently, the imino-aromatic NOE contacts demonstrate a quadruplex topological change with a strand inversion and an all-*syn* G-tract.

Initially, ambiguities may arise due to the overlap of G13 and G5 iminos. However, two NOE contacts are only compatible with the proposed structure. These are: (i) a contact of G13 H8 to H1 of a *syn-G*; as G13 H8 cannot have any NOE contact to G13 H1, this identifies G5 H1; (ii) a contact of G21 H1 to a *syn*-G H8 which is only compatible with G5 H8, thus verifying a *syn* conformation of G5 within a flipped strand.

**Figure S3.** A) H6/H8 –H1' region of a 2D NOE spectrum of dsQ(6,7) acquired at 30 °C with a mixing time of 300 ms. NOE walks along the duplex and the four G-tracts are colored black, blue, green, orange, and red; residue numbers identify intra-nucleotide NOE contacts. B) Through-bond correlations of guanine H1 and H8 protons via long-range couplings to their $^{13}$C5 with residue numbers indicated; different threshold levels were used for the left and right panels. C) 2D NOE spectral region showing $(H1)_n$-$(H2'/H2''/Me)_{n-1}$ sequential contacts. D) H1-H1 (left) and H8-H1 (right) NOE spectral regions. E) Color-coded schematic structure of dsQ(6,7).



### Resonance Assignments for dsQ(15,19)

In analogy to dsQ(6,7), the presence of Watson-Crick GC and AT imino protons with NOE contacts between G H1 and C aminos as well as T H3 and A H2 protons strongly suggest formation of a duplex from the 4-nt complementary 5'- and 3'- termini of dsQ(15,19) (Figure 1). Again, the G1 imino at the terminus is not observed due to fraying effects. G22 H8 shows sequential contacts into the fourth G4-tract with continuous H8-H1' NOEs ending with the 5'-terminal bromoguanosine $^{Br}$G19 (Figure S4A, red). The same connectivity pattern can also be observed for the third G-tract G17-G16-$^{Br}$G15 (Figure S4A, orange). The 5'-overhang shows no sequential H1'-H8 contacts to G5.

Three *syn*-G residues can be observed through both their strong intra-nucleotide H1'-H8 contacts and their characteristic downfield-shifted $^{13}$C8 chemical shifts (Figure 3). Two of these are part of sequential H1'-H8 connectivities with a typical rectangular NOE pattern as anticipated for *syn-anti* steps (Figure S4A, blue and green; the *syn-anti* pattern of the latter is partially hidden due to overlap by similar H1' chemical shifts). Three G residues are linked by a continuous NOE connectivity walk along one of the corresponding G-tracts whereas no sequential contacts to a third base are observed for the other G-tracts, indicating a *syn-syn-anti* glycosidic bond pattern. Notably, there are strong T4 methyl contacts to both G5 and G13 imino protons only compatible with a strand inversion of the first G-tract (Figure S4C).

Aromatic protons can be related to their imino protons through JR-HMBC experiments except for $^{Br}$G15, $^{Br}$G19, as well as for overlapping residues G13/G17, G7/G21, and G16/G20 (Figure S4B, orange). On the other hand, sequential NOE contacts to $(H2'/H2'')_{n-1}$ can be observed for the imino proton of a *syn*-G5 and *syn*-G6 residue, also allowing for a sequential assignment through the T4 sugar and methyl protons (Figure S4C). As G7 is clearly in *anti* conformation, this tract must feature a *syn-syn-anti* run. At least one other strand without a bromination site follows a *syn-anti-anti* pattern (Figure S4A, green). This can only be achieved by inverting the 5'-terminal G-tract, yielding a 5'-*syn-syn-anti*-3' pattern with all other tracts showing a 5'-*syn-anti-anti*-3' sequence of glycosidic bond angles. Thus, both $^{Br}$G residues adopt a preferred *syn*-conformation.

The resulting tetrad polarities are corroborated by H1-H8 intra-tetrad NOE contacts (Figure S4D right). The tetrad at the duplex-quadruplex junction shows cyclic contacts along G5-G13-G17-G21. Corresponding contacts can also be found for the central tetrad. H1-H8 contacts of the third tetrad are interrupted by the two $^{Br}$G analogs but a polarity inversion is evidenced by a G11 H1-G7 H8 connectivity. Additionally, non-sequential imino contacts between G15 and G12, G7 and G20, as well as G11 and G6 can be observed and give further evidence for a tetrad- and strand-flip to give the rearranged hybrid structure.

**Figure S4.** A) H6/H8 –H1' region of a 2D NOE spectrum of dsQ(15,19) acquired at 30 °C with a mixing time of 300 ms. NOE walks along the duplex and the four G-tracts are colored black, blue, green, orange, and red; residue numbers identify intra-nucleotide NOE contacts. B) Through-bond correlations of guanine H1 and H8 protons via long-range couplings to their $^{13}$C5 with residue numbers indicated; different threshold levels were used for the left and right panels. C) 2D NOE spectral region showing $(H1)_n$-(H2'/H2"/Me)$_{n-1}$ sequential contacts. D) H1-H1 (left) and H8-H1 (right) NOE spectral regions. E) Color-coded schematic structure of dsQ(15,19).

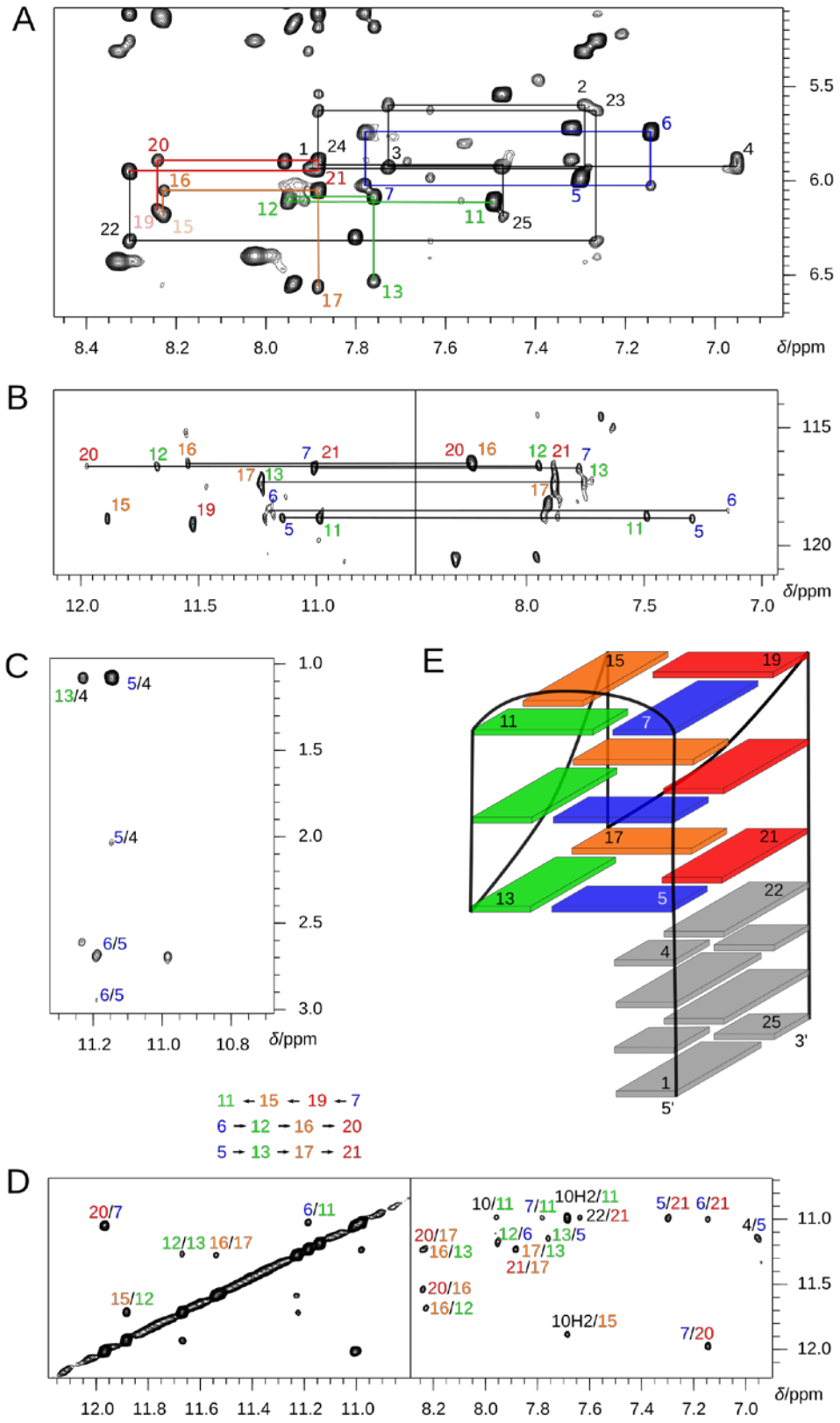


**Table S2**. NMR restraints and structural statistics for dsQ(6,7) and dsQ(15,19)

| NMR distance and torsion angle restraints | | |
|---|---|---|
| NOE-derived distance restraints | dsQ(6,7) | dsQ(15,19) |
| total | 341 | 231 |
| intra-residue | 214 | 125 |
| inter-residue | | |
| sequential | 94 | 76 |
| long-range | 33 | 30 |
| hydrogen bond restraints | 64 | 64 |
| torsion angle restraints | 25 | 25 |
| G-quartet planarity restraints | 36 | 36 |
| **structural statistics** | | |
| violations | | |
| mean NOE restraint violation (Å) | 0.0081±0.0033 | 0.0054±0.0013 |
| max. NOE restraint violation (Å) | 0.27 | 0.35 |
| deviations from idealized geometry | | |
| bond length (Å) | 0.01±0.0001 | 0.01±0.0001 |
| bond angle (°) | 2.29±0.03 | 2.38±0.0018 |
| pairwise heavy atom RMSD (Å) | | |
| overall | 1.75±0.42 | 1.71±0.26 |
| G-quartets | 0.99±0.18 | 0.92±0.16 |

**Figure S5.** Three-dimensional structures of dsQ(6,7) (left) and dsQ(15,19) (right). (A) Superposition of the ten lowest-energy structures. (B) Representative low energy structure in a cartoon representation. (C) Stacking between the central (red) and outer G-tetrad opposite the duplex overhang (grey).

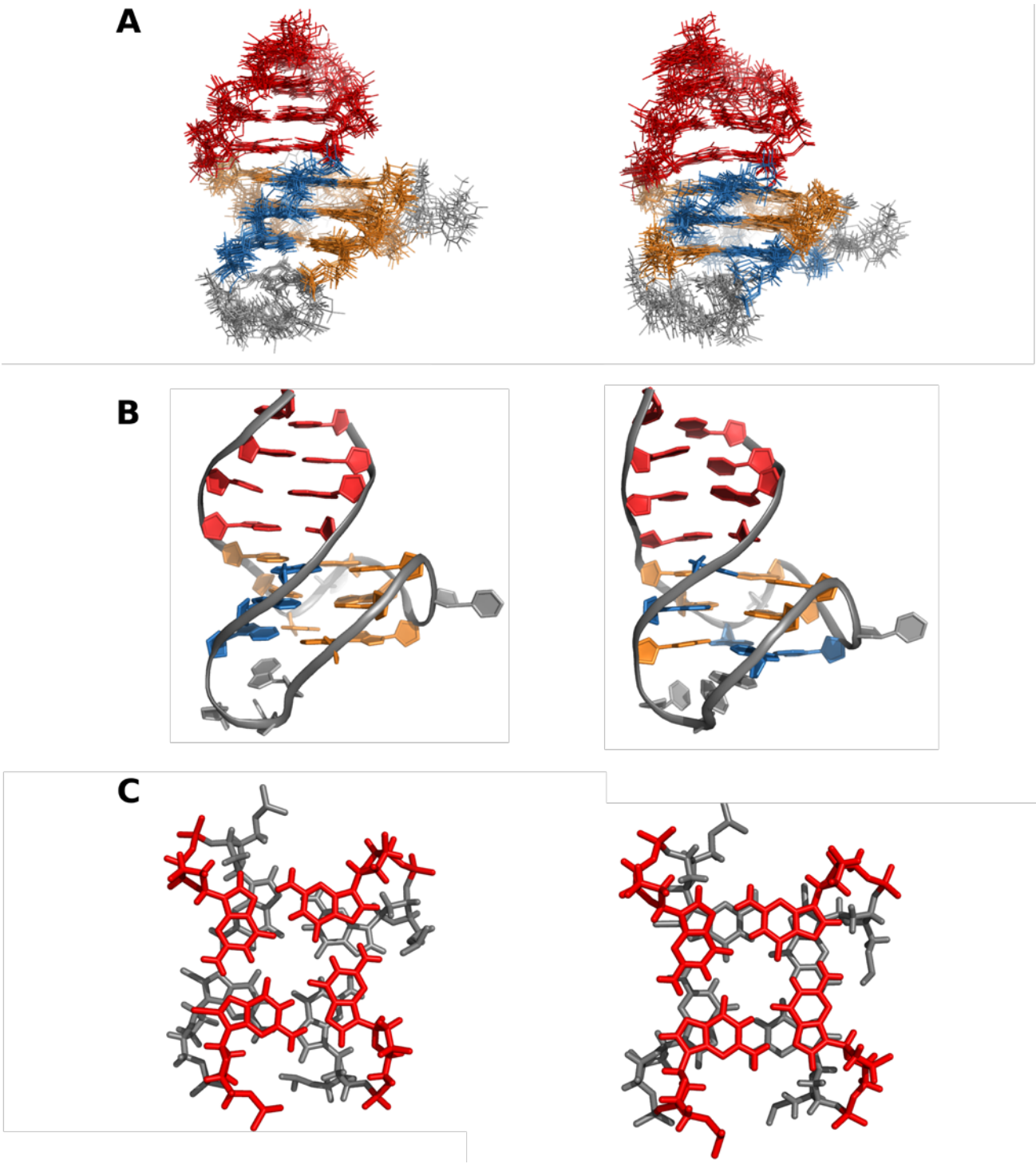


**Figure S6.** Exchange crosspeaks of imino protons with water in a 2D NOE spectrum of dsQ(6,7) acquired at 30 °C with a mixing time of 300 ms. The exchange-based broadening of the G19 imino, which is hydrogen-bonded to *syn*-G7 preceding the lateral loop is conspicuous. It should be noted that corresponding broadening effects due to fast solvent exchange are absent in the dsQ(15,19) hybrid-type quadruplex featuring an *anti*-G7 linked to the lateral loop (not shown).

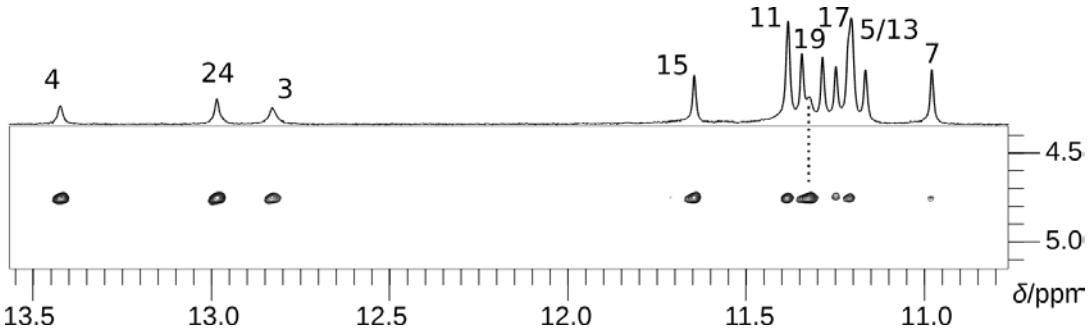



**Figure S7.** $^{15}$N-edited 1D HMQC spectrum of dsQ $^{15}$N labeled (10% enrichment) at adenosine position A10 (top) and comparative H8–C8 spectral region of $^{1}$H-$^{13}$C HSQC spectra for ssQ (black), dsQ(6,7) (orange), and dsQ(15,19) (blue) (bottom). The A10 H8 resonance for dsQ(15,19) in its hybrid-1R' form partly overlaps with an A H2 signal (marked with asterisks) that is only partially suppressed due to similar $^{1}$H-$^{15}$N couplings for adenine H8 and H2 protons. Note, that the deshielded and more shielded A H8 resonances are characteristic for A residues located in propeller and lateral loops, respectively.

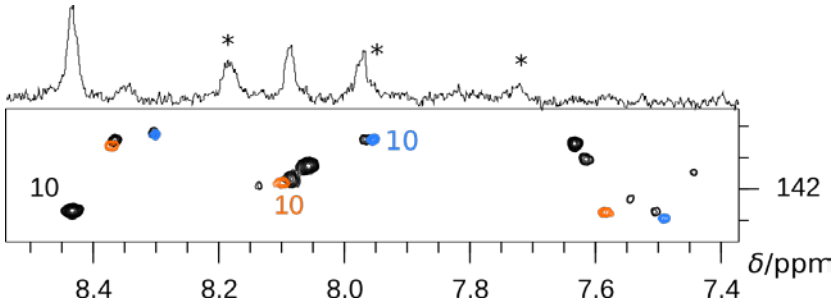
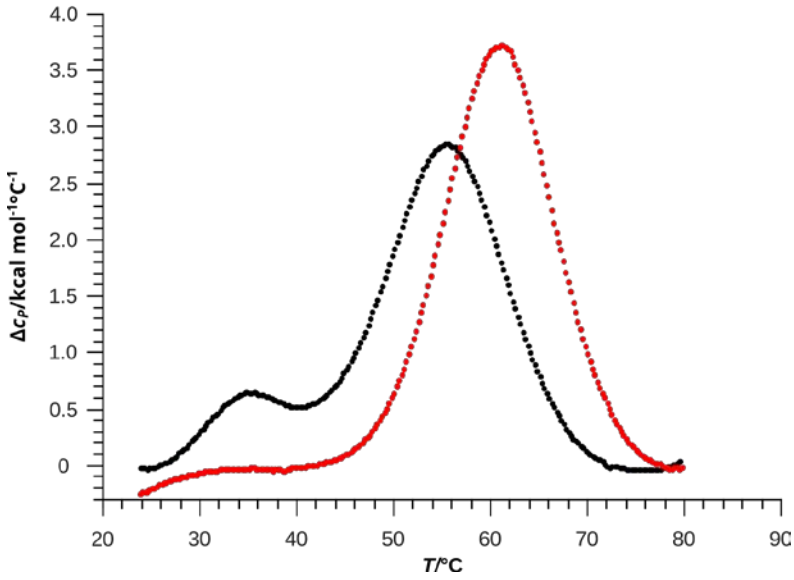



**Figure S8** DSC melting curves for dsQ (black) and ssQ (red) in $K^+$ buffer (pH 7).



## References


1. W. F. Vranken, W. Boucher, T. J. Stevens, R. H. Fogh, A. Pajon, M. Llinas, E. L. Ulrich, J. L. Markley, J. Ionides, E. D. Laue, *Proteins: Struct., Funct., Bioinf.* **2005**, *59*, 687–696.
2. E. Vanquelef, S. Simon, G. Marquant, E. Garcia, G. Klimerak, J. C. Delepine, P. Cieplak, F.-Y. Dupradeau, *Nucleic Acids Res.* **2011**, *39*, W511–W517.
3. C. D. Schwieters, J. J. Kuszewski, N. Tjandra, G. M. Clore, *J. Magn. Reson.* **2003**, *160*, 65–73.
4. M. Zgarbová, , M. Otyepka, J. Šponer, A. Mládek, P. Banáš, T. E. Cheatham, III, P. Jurečka, *J. Chem. Theory Comput.* **2011**, *7*, 2886–2902.
5. W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell, P. A. Kollman, *J. Am. Chem. Soc.* **1995**, *117*, 5179–5197.
6. M. Krepl, M. Zgarbová, P. Stadlbauer, M. Otyepka, P. Banáš, J. Koča, T. E. Cheatham, III, P. Jurečka, J. Šponer, *J. Chem. Theory Comput*. **2012**, *8*, 2506–2520.
7. M. Zgarbová, J. Šponer, M. Otyepka, T. E. Cheatham, III, R. Galindo-Murillo, P. Jurečka, *J. Chem. Theory Comput.* **2015**, *11*, 5723–5736.
8. A. T. Phan, *J. Biomol. NMR* **2000**, *16*, 175–178.

# Publication IV

ChemPubSoc Europe

DOI: 10.1002/cbic.201800642

CHEMBIOCHEM Minireviews

# Manipulating DNA G-Quadruplex Structures by Using Guanosine Analogues

Linn Haase, Beatrice Karg, and Klaus Weisz*[a]

The ability to control the folding topology of DNA G-quadruplexes allows for rational design of quadruplex-based scaffolds for potential use in various therapeutic and technological applications. By exploiting the distinct conformational properties of some base- and sugar-modified guanosine surrogates, conformational transitions can be induced through their judicious incorporation at specific sites in the quadruplex core. Changes may involve tetrad polarity inversions with conservation of the global fold or complete refolding to new topologies. Reliable predictions relating to low-energy conformers formed upon specific chemical perturbations of the system and the rational design of modified sequences suffer from our still limited understanding of the subtle interplay of various favorable and unfavorable interactions within a particular quadruplex scaffold. However, aided by an increasing number of systematic substitution experiments and high-resolution structures of modified quadruplex variants, critical interactions, in addition to glycosidic bond angle propensities, are starting to emerge as important contributors to modification-driven quadruplex refolding.

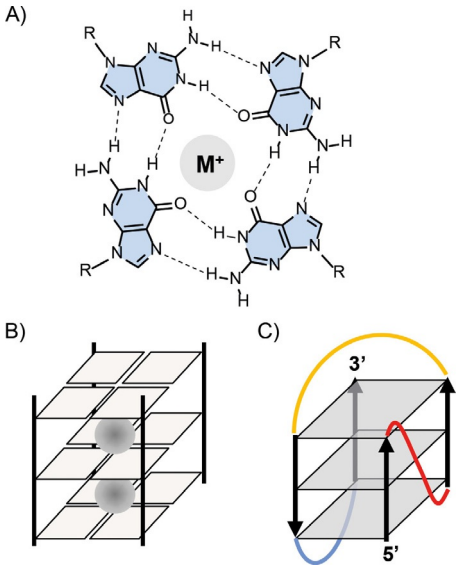
## Introduction

G-quadruplexes (G4s) are noncanonical nucleic acids formed by G-rich sequences. These four-stranded structures have generated a lot of interest during the last two decades, due to their potential regulatory roles in vivo, as well as their wide range of applications in bio- and nanotechnology. Various studies have convincingly demonstrated their formation within genomic DNA, in, for example, telomeric sequences or in promoter regions of various tumor-related genes such as *MYC*, *PDGF-A*, or *BCL-2*.[1–4] Aside from its biological significance, the G-quadruplex motif constitutes a key element for many aptamers selected for proteins or metabolites. In addition, G4-based sensor systems for metal ions, DNAzymes, or conductive nanowires all attest to the versatility of these fascinating structures for many current technological developments.[5–8]

Unlike Watson–Crick base-paired duplexes, DNA G-quadruplexes display considerable polymorphism, with a variety of topological features. The core of each quadruplex consists of a stack of, typically, two to four G-tetrads, each formed by a circular and coplanar arrangement of four guanine units connected through a total of eight Hoogsteen-type hydrogen bonds (Scheme 1). In addition to stacking interactions between tetrads, monovalent cations such as potassium or sodium are coordinated within the central channel by the carbonyl oxygen atoms of the guanine bases and contribute to the stability of the G-quadruplex core.

Typical G-quadruplexes can be divided into three main topological families: parallel quadruplexes with all four G-tracts being parallel, antiparallel quadruplexes with two of the four columnar strands running in opposite directions, and (3+1)-hybrid quadruplexes having one antiparallel and three parallel G-tracts. Quadruplexes may consist of one (monomolecular), two (bimolecular), or four (tetramolecular) separate G-rich strands. In the case of monomolecular quadruplexes, intervening sequences between adjacent parallel G-tracts form propeller loops, whereas lateral and diagonal loops link neighboring and diagonally positioned antiparallel strands (Scheme 1 C).

In the past, a variety of base- and sugar-modified G analogues have been substituted into tetra-stranded oligonucleotides for biophysical and biochemical studies. Thus, 8-oxo-, 8-




**Scheme 1.** A) Guanine tetrad with hydrogen bonds and a centrally located metal ion. B) G-quadruplex core containing three G-tetrad layers and centrally located metal ions. C) G-quadruplex with propeller (red), diagonal (orange), and lateral loops (blue) connecting the four columnar G-tracts.

[a] *L. Haase, B. Karg, K. Weisz*
*Institute of Biochemistry, Universität Greifswald*
*Felix-Hausdorff-Strasse 4 17487 Greifswald (Germany)*
*E-mail: weisz@uni-greifswald.de*

*The ORCID identification numbers for the authors of this article can be found under https://doi.org/10.1002/cbic.201800642.*

methyl, or 8-aryl-substituted G analogues represent naturally occurring DNA lesions, and knowing their impact on the folding of G-rich sequences into G-quadruplexes gives important clues on the regulatory role and physiological significance of these noncanonical nucleic acid structures.[9–11] Chemical modifications may also alleviate limitations to detailed structural and/or thermodynamic characterization as imposed by the highly polymorphous nature of quadruplexes. Because even small structural perturbations can affect the relative stabilities of different G4 conformations, appropriate substitutions have frequently been employed to stabilize a particular topology and to shift equilibria in a mixture of different quadruplex folds.[12–15] On the other hand, growing interest in quadruplexes for their use in technological and therapeutic applications has stimulated the incorporation of G analogues to alter G4 molecular properties deliberately. Thus, target affinity, nuclease resistance, or cellular uptake of G4-based aptamers might benefit from specific substitutions with noncanonical residues.[16,17] In addition, modular functionalization of DNA quadruplexes through N2-substituted guanine modifications was shown to yield a highly stable nanoscaffold.[18]

Whereas guanine N2-modifications conserve the native quadruplex topology, other G4 substitutions can completely redirect folding, thus expanding the toolbox of quadruplex architectures with implications for our basic understanding of the accessible G4 conformational landscape and G4 folding pathways. Additionally, G analogues can be powerful agents for tuning quadruplexes for particular applications in cases of binding partners with discriminatory potential towards quadruplex structural features such as groove geometries, but also in cases of G4 molecular properties affected by the specific tetrad stacking within the G-quadruplex core, such as electron-hole transfer in conductive G-wire systems.[19] To exploit fully the great potential of modification-driven G4 structural transitions for the future engineering of tailor-made architectures, a comprehensive understanding of relevant interactions within quadruplexes based on detailed structural characterization seems indispensable. In that respect, this article does not attempt to cover the various chemical modifications introduced into quadruplex structures and summarized in more recent reviews.[20,21] Rather, emphasis is placed on those base- and sugar-modified G analogues for which selective incorporation into G4-forming oligonucleotides has promoted defined structural rearrangements that have been subjected to more detailed characterization.

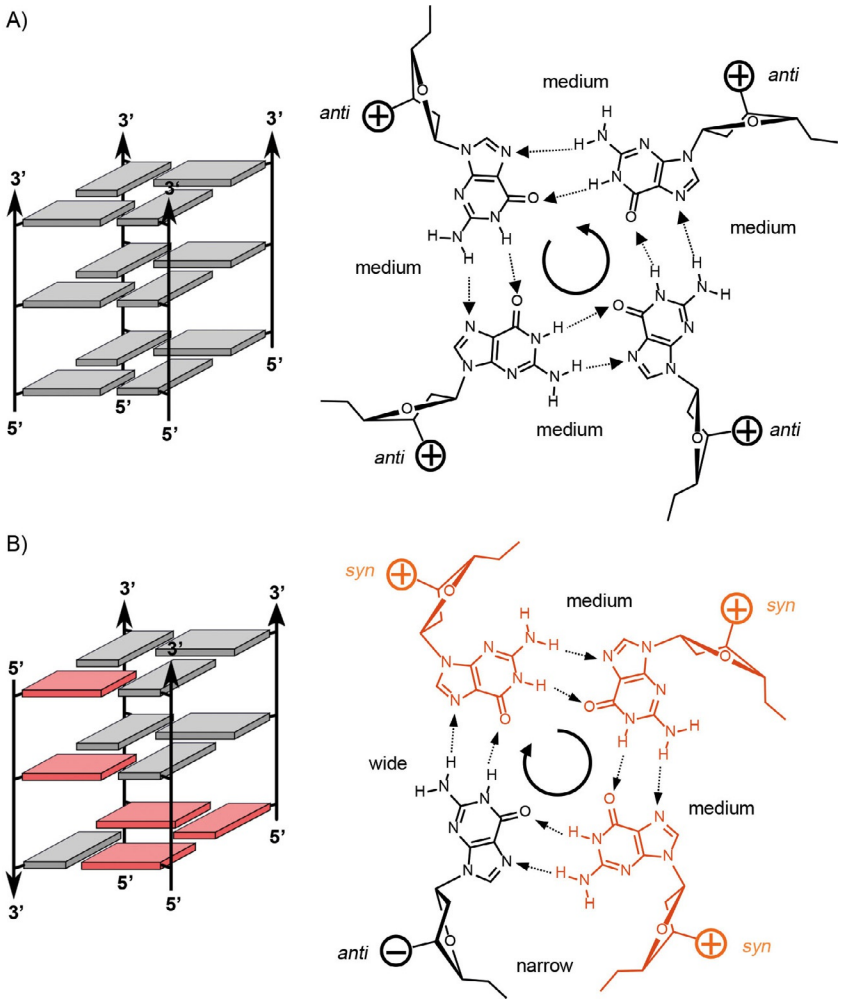
## Conformational features of G-quadruplexes

In a G-quadruplex, the orientation of individual G-tracts is correlated with the glycosidic torsion angles of G residues within a G-quartet through their specific arrangement. When all four G-tracts are parallel, the guanosine residues are required to have the same glycosidic conformation for G-tetrad formation, usually being all-*anti*. Consequently, inverting the orientation of one G-tract must be associated with a compensating base flip around the glycosidic bond (Scheme 2). The *syn-anti* pattern within a G-quartet also affects the dimensions of the four grooves formed by the four-stranded structure. Whereas a parallel quadruplex exhibits four medium-sized grooves, antiparallel and (3+1)-hybrid-type quadruplexes also form narrow and wide grooves between *syn-anti* and *anti-syn* hydrogen-bonded G pairs when viewed from H-bond donor to H-bond acceptor (Scheme 2).

Generally, 2′-deoxyriboguanosine residues within the quadruplex core adopt a C2′-*endo* sugar pucker and, although favored, *anti* glycosidic conformations can easily convert into *syn* glycosidic torsion angles. Whereas the circular *syn-anti* pattern is fixed within a G-tetrad through the conformation of one arbitrary residue and the relative orientation of strands, sequential *syn-anti* glycosidic conformations along individual G-tracts are variable. Thus, whereas three-layered antiparallel quadruplexes typically feature 5′-*anti-syn-anti* and 5′-*syn-anti-syn* arrays along the strands, (3+1)-hybrid quadruplexes mostly adopt 5′-*syn-anti-anti* and 5′-*syn-syn-anti* arrangements (Scheme 2B). Favored glycosidic bond angle patterns along G-tracts are expected to be influenced by stacking interactions, with *syn-anti* and *anti-anti* interactions preferred over *anti-syn* and *syn-syn* stacking.[22,23] In addition, the potential formation of an intramolecular hydrogen bond between a free sugar 5′-OH group and N3 promotes a *syn* orientation of a guanosine unit if located at the 5′ terminus.[24] It has to be noted that the polarity of a G-tetrad as defined by the clockwise or anticlockwise direction of Hoogsteen hydrogen bonds changes with alteration of the glycosidic conformation along the G-tracts, resulting in heteropolar stacking interactions between adjacent G-quartets. Accordingly, quadruplexes can also be divided into three groups with characteristic CD spectral signatures.[25] In general, quadruplexes of group I and group III are represented by parallel and antiparallel topologies, exclusively featuring homopolar and heteropolar stacking interactions, respectively. On the other hand, group II incorporates quadruplexes with both homopolar and heteropolar interactions as typified by (3+1)-hybrid quadruplexes.[26]

## Sugar conformations and *syn-anti* equilibria of guanosine analogues

Because of the dependency between relative orientation and glycosidic torsion angles within the individual G-quartets, the deliberate incorporation of *syn*- and *anti*-favoring G surrogates can be used in return to control G4 folding topologies. The most widely used G analogues preferring *syn* and *anti* conformations are 8-substituted or sugar C2′-modified guanosine derivatives, respectively. 8-Bromo- ($^{Br}$G, Scheme 3) or 8-methyl-2′-deoxyguanosine ($^{Me}$G) units are well known for their propensity to adopt the *syn* conformation, due to steric hindrances of their large 8-substituents when positioned above the furanose sugar. On the other hand, governed by anomeric and *gauche* effects, the electron-attracting 2′-OH and 2′-F substituents in riboguanosine (rG) and 2′-deoxy-2′-fluororiboguanosine ($^{F}$rG) affect glycosidic torsion angles through their preferred axial orientation within a pseudorotational conformer in the north (*N*) domain. The corresponding C3′-*endo* pseudorotamers are strongly correlated with *anti* conformations because they place




**Scheme 2.** Guanine tetrad and schematic representation of A) a parallel, and B) a (3 + 1)-hybrid G4 without loop-forming intervening sequences. The directions of hydrogen bonds between guanosine units and the groove widths are indicated, (+) and (−) denote strand directionalities, and gray and red rectangles in the schematic structures represent *anti*- and *syn*-guanosine units, respectively.

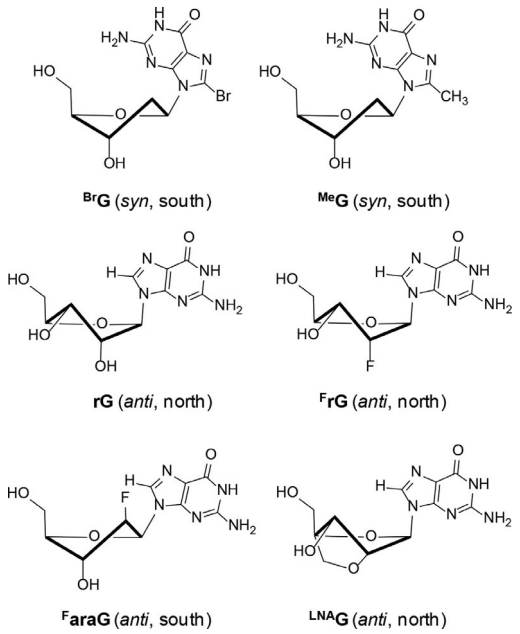
the nucleobase in a pseudoaxial position that causes unfavorable steric interactions if the guanine base is positioned above the sugar ring in a *syn* orientation. However, quantum-mechanical studies have questioned the general validity of such an interpretation, reporting a *syn* conformation for free nucleosides close in energy for south and north sugars.[27]

Whereas *syn* conformers of rG and $^{F}$rG analogues are destabilizing but might still be tolerated, depending on the individual structural context, a bicyclic 2′-*O*-4′-*C*-methylene rG analogue (locked nucleic acid, $^{LNA}$G) is locked in a C3′-*endo* (*N*-type) conformation. Combined with its very high pucker amplitude of about 60°, a *syn* glycosidic conformation is mostly prevented, due to severe clashes between the base and H3′ above the furanose plane.[28] Another popular *anti*-G substitute is based on the C2′ epimer of $^{F}$rG, 2′-deoxy-2′-fluoroarabinoguanosine ($^{F}$araG). Steered by strong *gauche* and possibly also steric effects due to the β-2′-fluorine, this rather rigid nucleoside tends to adopt a more DNA-like south-east/east or C1′-*exo*/O4′-*endo* conformation.[29–31] Nevertheless, expected unfavorable interactions between the β-positioned fluorine above the furanose sugar and a *syn*-guanine moiety, together with a potential unconventional intranucleotide hydrogen bond between fluorine and H8 of an *anti*-guanine base, may be advocated as major determinants for its preferred *anti* conformation.[32,33]

## Quadruplex tetrad polarity inversion

Induced *anti*↔*syn* conformational transitions of all four residues within a G-tetrad invert its polarity without affecting the G4 global fold. As a result, stacking interactions will change, as often revealed by a CD signature characteristic of a different quadruplex group. Inverting the G-tetrad polarity can affect fluorescence and charge-transfer properties of the quadruplex but might also influence selectivities in asymmetric organic reactions catalyzed by G4-based DNAzymes.[7,34,35]

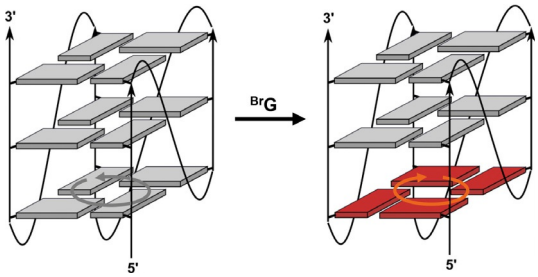
Tetrad inversions were first reported upon incorporation of *syn*-favoring $^{Br}$G or $^{Me}$G residues into tetramolecular model quadruplexes featuring a fourfold symmetry with a parallel strand arrangement and all-*anti* conformations.[36,37] Driven by the *syn* propensity of the 8-substituted G analogues, a fully modified tetrad at the 5′ terminus (5′-tetrad) switches polarity




**Scheme 3.** 2′-Deoxyguanosine analogues with favored glycosidic conformations and sugar puckers.

to give an all-*syn* arrangement, whereas correspondingly modified central tetrads seem to resist a concerted base flip. In addition, thermal stabilities tend to decrease when modified tetrads are shifted from the 5′- to the 3′-end, with partial unfolding often observed for quadruplexes with an 8-bromo- or 8-methyl-substituted 3′-terminal quartet. The latter would be expected to suffer from unfavorable *anti–syn* steps upon a 3′-tetrad inversion, whereas an increased stability for a G4 with a modified 5′-inverted tetrad can be traced to favorable stacking energies of its *syn–anti* steps. Notably, various observations on parallel quadruplexes may be attributed to such a "5′-tetrad effect". These include 1) reports on a minor species with a 5′-terminal all-*syn* quartet in an unmodified $[d(TGGGT)]_4$ quadruplex,[38] 2) observation of a glycosidic angle in the *high-anti* (far *syn*) range for the 5′-terminal residues preceding $^{LNA}$G modifications in $[d(TG^{LNA}GG^{LNA}GT)]_4$,[39] and 3) a G4 destabilization observed upon replacement of the 5′-tetrad by an *anti*-affine rG quartet as well as a polarity inversion to an all-*syn* 5′-tetrad in the case of an adjacent quartet fully substituted with four rG residues.[40] Remarkably, the latter 5′-tetrad inversion is unusual in likely being induced by a relay mechanism starting from conformational perturbations of a neighboring rG quartet. On the basis of these experimental findings, the 5′-tetrad in a parallel quadruplex may be considered a hinge for geometric adjustments, favoring concerted *anti*→*syn* interconversions and possibly disfavoring modifications with a pronounced *anti* conformational preference.

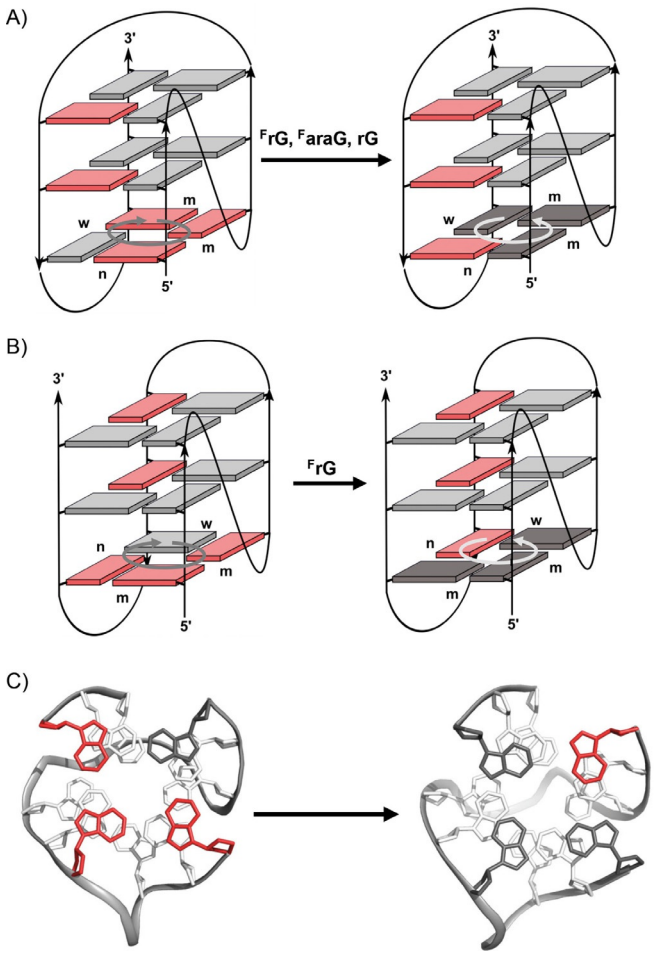
Extending such conformational rearrangements to monomolecular quadruplexes allows for substitutions at single sites but also adds restrictions imposed by the loop sequences. Incorporation of $^{Br}$G analogues at all four *anti*-G positions of the 5′-tetrad enforces a complete tetrad polarity inversion in a robust parallel-stranded *MYC* quadruplex with the formation of an all-*syn* G-tetrad (Scheme 4).[41] In line with additive effects of substitutions, only a partial inversion was observed with a corresponding triply substituted *MYC* variant. In that case, native and flipped structures present in about a 1:1 molar ratio were




**Scheme 4.** 5′-Terminal G-tetrad polarity inversion in the parallel *MYC* quadruplex through *syn*↔*anti* conformational transitions after substitution of $^{Br}$G analogues for the *anti* G residues in the 5′-quartet; *syn* $^{Br}$G and *anti* G residues are colored red and gray, respectively.

found to be in dynamic exchange. In addition to the much stronger propensity of the 5′-tetrad to undergo a polarity inversion (vide supra), significant position-dependent effects for inducing a tetrad flip have been found and attributed to the preceding propeller loops.[42] Thus, a short and less flexible 1 nt loop will strongly enforce an *anti*→*syn* conformational transition of the $^{Br}$G analogue to evade unfavorable bromine interactions with the 5′-phosphate group. In contrast, longer loops or 5′-overhang sequences allow for conformational adjustments that can alleviate bromine steric clashes in an *anti* conformer of the native fold.

In analogy to substitutions of *syn*-favoring G analogues into parallel quadruplexes, the introduction of *anti*-favoring 2′-modified guanosine analogues into unmatching *syn* positions can likewise induce a tetrad flip in nonparallel quadruplexes. Thus, $^{F}$rG incorporation into all three *syn* positions within the 5′-tetrad of a (3+1)-hybrid quadruplex (ODN) with a propeller, diagonal, and lateral loop triggered a concerted flip of the glycosidic torsion angle for all residues in the 5′-tetrad (Scheme 5).[43] The formed tetrad polarity isomer represents an unusual type of G4 conformer in having an all-*syn* G-tract and, despite conserving its native topology, features CD spectra reminiscent of a parallel G-quadruplex, due to its exclusively homopolar stacking interactions. Demonstrating that such conformational transitions are not restricted to one particular folding topology, a 5′-tetrad polarity inversion was also observed upon corresponding modifications of a human telomeric (3+1)-hybrid structure (HT) featuring overhang sequences, one propeller, and two lateral loops (Scheme 5).[44] A different strategy for inducing a tetrad flip in the HT quadruplex employed the simultaneous incorporation of 8-oxoguanosine (O) and xanthine (X) residues at neighboring positions of its 5′-tetrad.[45] Here, one out of two stable X·O base pair arrangements features a pattern of free hydrogen bond donor and acceptor sites complementary to a coplanar G·G Hoogsteen base pair, enforcing concerted changes in glycosidic torsion angles to form a stable G·G·X·O quartet with defined polarity.




**Scheme 5.** Schematic representation of a polarity inversion upon incorporation of *anti*-favoring G analogues into *syn* positions of the 5′-tetrad for A) the ODN, and B) the HT quadruplex; *syn* and *anti* residues are colored red and gray, respectively, with modified positions highlighted in darker color, whereas m, w, and n denote medium, wide, and narrow grooves. C) Top view of the 5′-terminal tetrad of a representative NMR solution structure for the native (left, PDB ID: 2LOD) and the tetrad-inverted $^{F}$rG-modified ODN quadruplex (right, PDB ID: 5MCR).

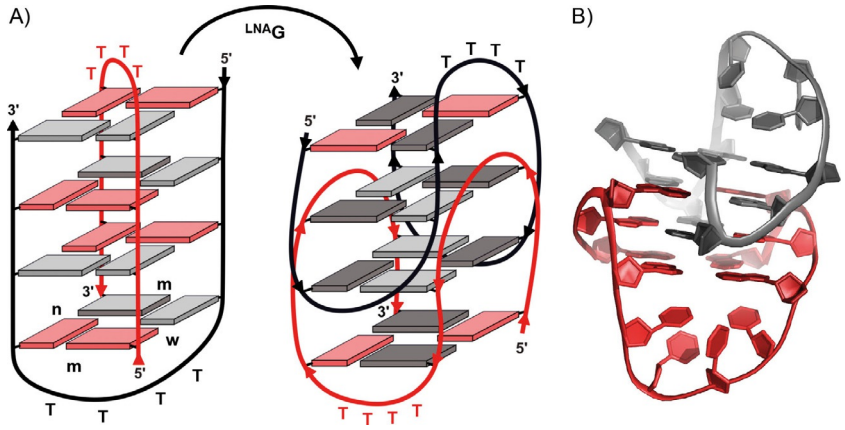
Recently, the ODN quadruplex has been subjected to further modifications also employing $^{F}$araG and rG as *anti*-favoring surrogates for incorporation into unmatching *syn* positions of its 5′-tetrad. In analogy to $^{F}$rG substitutions, a tetrad polarity flip was promoted in both cases.[33,46] Noticeably, stabilities of $^{F}$araG-modified quadruplexes were hardly compromised, in contrast to $^{F}$rG modifications with their destabilizing effect depending on number and position of substitution sites.[33,43] Thermal destabilization was also observed with rG modifications, but effects here were not additive, and a complete transition to the isomer with inverted 5′-tetrad polarity was induced by the disubstituted quadruplex variant but only partially for the trisubstituted one.[46]

## Changes in quadruplex folding topologies

Shifting of equilibria between coexisting G4 species is mostly based on their different *syn*/*anti* patterns, and judicious substitutions at selected positions have been successfully applied in many cases. In addition, swift transitions to parallel quadruplexes as observed upon the incorporation of *anti*-favoring rG, $^{LNA}$G, $^{F}$rG, and $^{F}$araG analogues at multiple *syn* positions of a quadruplex with an antiparallel folding topology are easily explained in terms of *anti* glycosidic preferences of the sugar-modified nucleosides.[28,47–49] Notably, as a result of a better conformational fit and more favorable fluorine interactions, $^{F}$araG is typically found to be superior to $^{F}$rG in the stabilization of a parallel fold.[12,33,49,50] However, complete G4 refolding into a new topology after local structural perturbations is often difficult to predict. Thus, substitution of quadruplexes with G analogues adopting unmatching glycosidic torsion angles at single positions might yield a new major fold but might also result in quadruplex unfolding or in unresolvable polymorphous mixtures. Although such rearrangements are thought to be governed by the modulation of *syn*/*anti* equilibria through the introduced G derivatives, the observation of noticeable destabilizing effects on the native quadruplex even when G surrogates are incorporated into matching positions demonstrates that additional, more subtle interactions are operative and contribute to the structural reorganization.

Strand reorientations in tetramolecular quadruplex structures lacking loop domains are expected to be hardly restricted. However, a parallel strand orientation seems to be strongly preferred, with a newly formed antiparallel structure having only been reported for $[d(TGGGGT)]_4$ upon the incorporation of *syn*-favoring $^{Me}$G analogues at the first and third position of the G-run.[51] As proposed on the basis of NMR experimental data, canonical guanosine residues and $^{Me}$G analogues exclusively adopt *anti* and *syn* conformations, respectively. Featuring a *syn-anti-syn-anti* pattern along the four strands, parallel and antiparallel strand orientations are expected to exhibit similar stacking interactions. In fact, NMR spectra suggest the presence of other G4 structures, pointing to coexisting species of similar free energy. It should also be noted that four of the eight methyl substituents are relocated from a medium into a wide groove upon the conversion from a parallel to the antiparallel structure, with possible implications for relative stabilities.

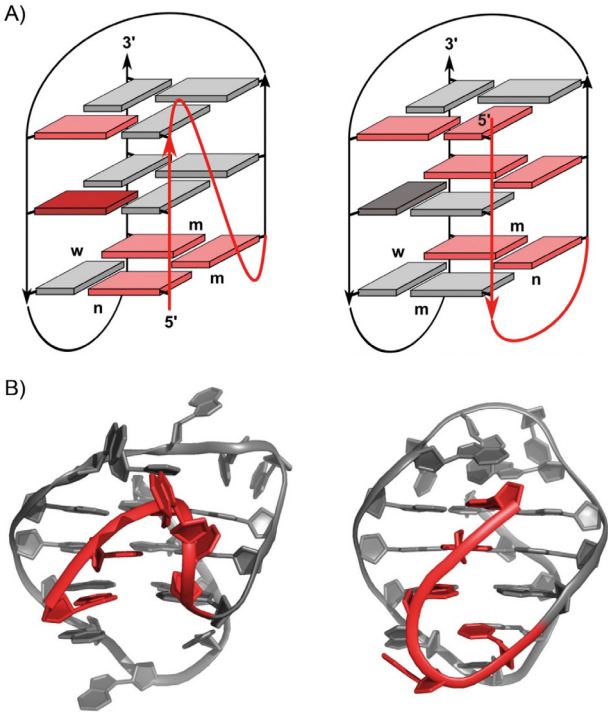
Locked nucleic acid modifications are exceptional, due to their rigid bicyclic ring system and associated restricted conformational flexibility. As expected, LNA-modified quadruplexes typically prefer parallel folds with all-*anti* core residues over antiparallel topologies.[13,28] If suitably positioned, a strong stabilization due to conformational preorganization of the LNA residue associated with favorable entropic effects is generally observed.[39,52] On the other hand, in the case of an imperfect conformational match, limited adaptive capabilities, possibly amplified by the propensity of the LNA residue also to enforce an *N*-type sugar pucker for the subsequent nucleotide,[39] might result in considerable perturbations of the native structure, often with unpredictable outcomes. Thus, LNA substitutions in *anti* positions upstream of short propeller loops have been found to exert a strongly destabilizing effect.[12] Apparently, conformational restraints imposed by the propeller loop cannot be compensated by the rigid LNA analogue. Likewise,




**Scheme 6.** A) Conversion of a dimeric antiparallel quadruplex into a V4 fold after substitution of $^{LNA}G$ for all *anti* G residues; *syn* and *anti* residues are colored red and gray, respectively, with modified positions highlighted in dark gray, whereas m, w, and n denote medium, wide, and narrow grooves. B) Side view of a high-resolution structure of the V4-folded $^{LNA}G$ modified quadruplex (PDB ID: 2WCN).

significant destabilization, even driving a complete refolding of the native G4 structure, has been observed in LNA-modified quadruplexes despite substitution at matching *anti* positions.[53,54] Notably, formation of an unexpected topology, termed a V4 fold, has been reported for a telomeric $d(G_4T_4G_4)$ sequence from *Oxytricha nova* when the four matching *anti* G positions were replaced by $^{LNA}G$ to give $d(G^{LNA}GG^{LNA}GT_4G^{LNA}GG^{LNA}G)$ (Scheme 6).[54] Whereas the unmodified sequence forms a four-tetrad dimeric quadruplex with antiparallel G-tracts and diagonal $T_4$ loops, the modified sequence forms an unusual four-layered G4 structure featuring interrupted G columns with the two symmetry-related strands folding back in four V-shaped loops. Because some of the LNA residues would border the narrow groove in a native topology, unfavorable narrow-groove interactions might trigger the topological transition to the V4 fold.

Partial refolding of the (3 + 1)-hybrid quadruplex ODN when substituted at a single *syn* position within its central tetrad by an $^{F}rG$ or $^{F}araG$ analogue was recently observed and characterized in detail.[55] For both modified sequences two species coexist in solution: the native fold and a new topology with an inverted first G-tract (Scheme 7). Strand inversion associated with the conversion of a 2 nt propeller into a lateral loop allows the incorporated 2′-F-modified residue in the third strand to adopt a preferred *anti* conformation. At the same time, two G residues within the same central tetrad change to a *syn* conformation. The newly formed G4 constitutes a rare type of antiparallel quadruplex in combining both homo- and heteropolar G-tetrad stacking interactions. It is characterized by pairs of *syn-anti-anti* and *syn-syn-anti* G-tracts running in opposite directions and bridged by the conserved diagonal loop. In contrast with tetrad polarity isomers, no dynamic equilibrium between the native and rearranged structure was observed, and their mutual interconversion requires complete strand separation. In support of a kinetic partitioning mechanism, the rather unusual substitution-induced refolding can be traced to a stabilization of alternate G-triplex intermediates, by taking account of the glycosidic torsion angle propensities of the G analogue and more favorable stacking interactions for *anti–anti* and *syn–anti* nucleotide steps along the G-tracts.[55]




**Scheme 7.** Native (left) and rearranged ODN quadruplex with an inverted first G-tract (right). A) Schematic representation with *syn* and *anti* residues colored red and gray, respectively; the single $^{F}rG$ or $^{F}araG$ analogue at the *syn* position of the central tetrad is highlighted by the darker color. B) Side view of the first two G-tracts of a representative NMR solution structure for the native (left, PDB ID: 2LOD) and the tetrad-inverted $^{F}araG$-modified quadruplex (right, PDB ID: 6F4Z); the 5′-tract and the following loop are colored red.

## Formation of unconventional hydrogen bonds

Remarkably, the coexistence of two stable quadruplex structures with either an *anti* or a *syn* conformation for 2′-F-modified G incorporated into the central tetrad of the ODN G4 illus-

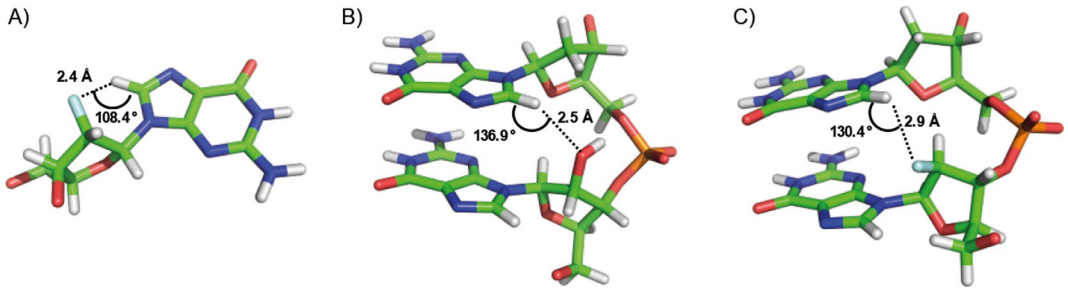



**Figure 1.** Unconventional hydrogen bonds formed in sugar-modified quadruplexes. A) Intraresidual C8–H8···F2′ hydrogen bond in a $^{F}$araG-modified ODN (PDB ID: 5OV2). B) Sequential C8–H8···O2′ hydrogen bond in a rG-modified ODN (PDB ID: 6FFR). C) Sequential C8–H8···F2′ hydrogen bond in a parallel $^{F}$araG-modified two-repeat telomeric sequence (PDB ID: 2M1G).

trates that both $^{F}$rG and $^{F}$araG might also assume the disfavored *syn* glycosidic torsion angle in a conserved topology. Several examples suggest that the stability of different conformations both for 2′-F- and for 2′-OH-substituted G analogues, which might possibly determine distinct global conformational rearrangements, seems to be partially linked to the formation of stabilizing unconventional hydrogen bonds between the 2′-substituent in an *S*-puckered ribose unit as hydrogen bond acceptor and guanine C8–H8 (Figure 1).[46,50,55,56] Strikingly, in the two coexisting ODN topologies with a single modification in the central tetrad (Scheme 7), both *syn* $^{F}$rG and *syn* $^{F}$araG in the native G4, but also *anti* $^{F}$rG in the rearranged G4, participate in a sequential F2′···H8–C8 hydrogen bond to the 3′-adjacent guanine base.[55] In contrast, *anti* $^{F}$araG was found to be stabilized by an intraresidue hydrogen bridge between fluorine and its own H8 (Figure 1 A).[33,55] Such interactions were unambiguously demonstrated experimentally through the observation of *trans*-hydrogen-bond spin–spin scalar couplings $^{h1}J$(H,F).

Although unconventional hydrogen bonds appear to play a crucial role in determining conformational equilibria, it is difficult to predict their formation and impact on the stability of specific G4 topologies reliably. Thus, the expected participation of fluorine in a sequential hydrogen bond to H8 was not observed when $^{F}$rG was incorporated into the 5′-tetrad of the ODN quadruplex, yet an O2′···H8–C8 interaction was found for corresponding rG modifications (Figure 1 B).[43,44,46] Moreover, whereas the *anti* $^{F}$araG residue showed an intraresidue hydrogen bond when incorporated into central or outer positions of the ODN G4,[33,55] a sequential F2′···H8–C8 pseudo-hydrogen bond was proposed on the basis of a high-resolution NMR structure of an $^{F}$araG-modified two-repeat telomeric quadruplex with a parallel fold (Figure 1 C).[50] The dimeric structure also revealed a close contact between H2″ and O4′ of the 3′-neighboring guanosine residue, indicating an additional stabilizing F2′–C2′–H2″···O4′ interaction. In fact, these fluorine-associated interactions have been proposed to contribute to the superior G4 stabilization of $^{F}$araG relative to araG and rG substitutions in the telomeric quadruplex. On the other hand, strongly destabilizing effects as observed for $^{F}$rG placed at corresponding positions were attributed to unfavorable interactions of its α-2′-fluorine atom. Thus, whereas the 2′-OH group in araG and rG might be engaged in hydrogen bonding with neighboring phosphate groups, electrostatic repulsion between F2′ in $^{F}$rG and the backbone phosphate was found by quantum-mechanical calculations to compromise G4 stabilities.[50]

## Interactions within narrow grooves

Not only backbone but also groove interactions seem to affect the conformations of introduced G analogues with their influence on quadruplex transitions. Structural data suggest that a quadruplex narrow groove does not readily host sugar or base substituents.[44,46] For instance, $^{F}$rG and rG, when incorporated into the ODN or HT 5′-position of a G-tract bordering a narrow and medium groove, were each found to adopt a C3′-*endo* sugar pucker. It is this conformation that orients the 2′-fluoro or 2′-hydroxy substituent away from the narrow groove and into the medium groove, likely avoiding unfavorable interactions within the narrow groove (Figure 2 A). In contrast, G analogues of other G-tracts are S-puckered with the 2′-substituent turned upwards, allowing for sequential pseudo-hydrogen bonds within a medium groove (Figure 2 B).[44,46]

Notably, the role of unfavorable narrow groove interactions in driving rG or $^{F}$rG puckering seems to be most pronounced in the 5′-tetrad. Thus, rG substituted for an *anti* G in the central position of the first G-tract of the ODN quadruplex that borders a narrow groove adopts a C2′-*endo* conformation to allow participation of the 2′-oxygen atom in a sequential pseudo-hydrogen bond with C8–H8.[56] Interestingly, rG incorporation at the single *anti* position of the ODN 5′-tetrad causes disruption of the quadruplex fold. This could be attributed to the subsequent strained 3 nt loop spanning a wide groove and restricting the sugar pucker to south, a conformation that would shift the 2′-OH substituent into the narrow groove (Scheme 5).[46,56] On the other hand, the single *anti* G position within the 5′-tetrad of the telomeric (3 + 1) quadruplex HT precedes a more relaxed lateral loop that spans a narrow groove, resulting in only a very small destabilization upon incorporation of an $^{F}$rG analogue.[12] Apparently, the impact of modifications is determined by a combination of several factors that also include conformational restraints as imposed by adjacent loops and grooves.

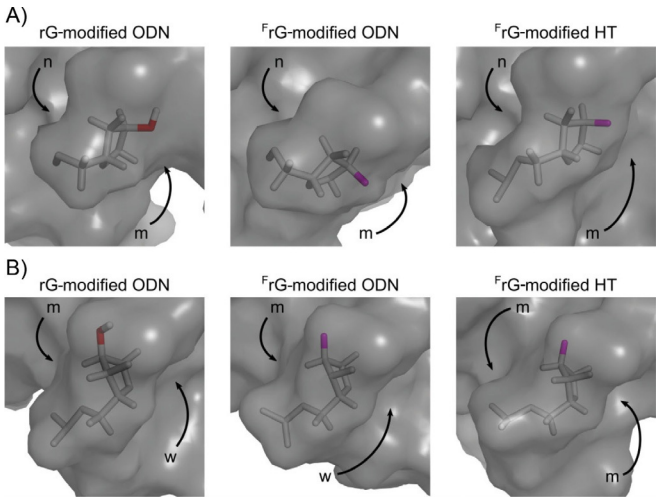



**Figure 2.** Sugar conformation and orientation of 2′-substituents for rG and FrG residues incorporated in the 5′-tetrad of ODN and HT quadruplexes. The sugar moiety of the analogue is shown in stick representation with O2′ and F2′ substituents highlighted in red and magenta, respectively. A) *N*-puckered rG and FrG residues bordering narrow and medium grooves (located in the 5′-terminal and 3′-terminal G-tract of ODN and HT, respectively). The C3′-*endo* conformation orients the 2′-substituent away from the narrow groove and sideways into the medium groove. B) *S*-puckered rG and FrG residues bordering medium and wide grooves (third G-tract of ODN) or two medium grooves (5′-terminal G-tract of HT). The sugar pucker in the south domain causes an upward orientation of the 2′-substituent, placing it in a medium groove.

## Conclusions

Site-specific modifications of DNA G-quadruplexes with appropriate G analogues offer the possibility to stabilize a particular fold but to also induce refolding to a new conformer with different polarity or topological features. Whereas *syn*/*anti* preferences of residues are known to be of prime importance for the stability of a particular fold, other, more subtle interactions also contribute to folding and must often be considered when trying to interpret the preferred formation of particular quadruplex structures. The formation of unconventional hydrogen bonds, steric interactions of substituents within narrow grooves, and conformational restrictions imposed by short loops have all been revealed as potential contributors to observed G4 structural interconversions. Identifying such interactions calls for a rigorous analysis of sugar puckers but also of the conformation along the sugar-phosphate backbone. These parameters are often overlooked when studying quadruplexes with low-resolution methods, with emphasis often being put exclusively on the more easily accessible glycosidic torsion angles.

To exploit fully the potential of chemical modifications and their impact on G4 structural diversity in the future, in-depth understanding of interactions and their interplay within a quadruplex scaffold is required. With an ever increasing number of high-resolution structures for unmodified and modified quadruplexes, more information on relevant electronic and steric effects can be collected to allow better prediction of low-energy conformers and, eventually, rational engineering of G4 architectures tailored to particular needs. In addition to the guanosine surrogates presented in this review article, other analogues including backbone modifications might also be recruited to drive well-defined quadruplex conformational transitions depending on sequence context and environmental conditions. It can be anticipated that advances in the conformational control of the quadruplex scaffold through appropriate substitutions into G-rich oligonucleotides will considerably expand bio- and nanotechnological applications based on the G4 platform in the future.

## *Conflict of Interest*

*The authors declare no conflict of interest.*

**Keywords:** aptamers · conformation analysis · G-quadruplexes · guanosine · oligonucleotides · topology

[1] G. Biffi, D. Tannahill, J. McCafferty, S. Balasubramanian, *Nat. Chem.* **2013**, *5*, 182 – 186.
[2] V. S. Chambers, G. Marsico, J. M. Boutell, M. Di Antonio, G. P. Smith, S. Balasubramanian, *Nat. Biotechnol.* **2015**, *33*, 877 – 881.
[3] A. Laguerre, K. Hukezalie, P. Winckler, F. Katranji, G. Chanteloup, M. Pirrotta, J.-M. Perrier-Cornet, J. M. Y. Wong, D. Monchaud, *J. Am. Chem. Soc.* **2015**, *137*, 8521 – 8525.
[4] 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 – 1272.
[5] J. L. Neo, K. Kamaladasan, M. Uttamchandani, *Curr. Pharm. Des.* **2012**, *18*, 2048 – 2057.
[6] F. Wang, X. Liu, I. Willner, *Angew. Chem. Int. Ed.* **2015**, *54*, 1098 – 1129; *Angew. Chem.* **2015**, *127*, 1112 – 1144.
[7] L.-X. Wang, J.-F. Xiang, Y.-L. Tang, *Adv. Synth. Catal.* **2015**, *357*, 13 – 20.
[8] G. I. Livshits, J. Ghabboun, N. Borovok, A. B. Kotlyar, D. Porath, *Adv. Mater.* **2014**, *26*, 4981 – 4985.
[9] A. M. Fleming, Y. Ding, C. J. Burrows, *Proc. Natl. Acad. Sci. USA* **2017**, *114*, 2604 – 2609.
[10] N. Shrivastav, D. Li, J. M. Essigmann, *Carcinogenesis* **2010**, *31*, 59 – 70.
[11] M. Sproviero, A. M. R. Verwey, K. M. Rankin, A. A. Witham, D. V. Soldatov, R. A. Manderville, M. I. Fekry, S. J. Sturla, P. Sharma, S. D. Wetmore, *Nucleic Acids Res.* **2014**, *42*, 13405 – 13421.
[12] Z. Li, C. J. Lech, A. T. Phan, *Nucleic Acids Res.* **2014**, *42*, 4068 – 4079.
[13] M. Marušič, R. N. Veedu, J. Wengel, J. Plavec, *Nucleic Acids Res.* **2013**, *41*, 9524 – 9536.
[14] A. Matsugami, Y. Xu, Y. Noguchi, H. Sugiyama, M. Katahira, *FEBS J.* **2007**, *274*, 3545 – 3556.
[15] K. W. Lim, S. Amrane, S. Bouaziz, W. Xu, Y. Mu, D. J. Patel, K. N. Luu, A. T. Phan, *J. Am. Chem. Soc.* **2009**, *131*, 4301 – 4309.
[16] J. Lietard, H. A. Assi, I. Gómez-Pinto, C. González, M. M. Somoza, M. J. Damha, *Nucleic Acids Res.* **2017**, *45*, 1619 – 1632.
[17] E. B. Pedersen, J. T. Nielsen, C. Nielsen, V. V. Filichev, *Nucleic Acids Res.* **2011**, *39*, 2470 – 2481.
[18] C. J. Lech, A. T. Phan, *Nucleic Acids Res.* **2017**, *45*, 6265 – 6274.
[19] C. J. Lech, A. T. Phan, M. E. Michel-Beyerle, A. A. Voityuk, *J. Phys. Chem. B* **2013**, *117*, 9851 – 9856.
[20] O. Doluca, J. M. Withers, V. V. Filichev, *Chem. Rev.* **2013**, *113*, 3044 – 3083.
[21] J. Sagi, *J. Nucleic Acids* **2017**, 1641845.
[22] X. Cang, J. Šponer, T. E. Cheatham III, *Nucleic Acids Res.* **2011**, *39*, 4499 – 4512.
[23] J. Šponer, A. Mládek, N. Špačková, X. Cang, T. E. Cheatham III, S. Grimme, *J. Am. Chem. Soc.* **2013**, *135*, 9785 – 9796.
[24] G. A. Jeffrey, W. Saenger, *Hydrogen Bonding in Biological Structures*, Springer, New York, **1994**, Chapter 17.
[25] S. Masiero, R. Trotta, S. Pieraccini, S. De Tito, R. Perone, A. Randazzo, G. P. Spada, *Org. Biomol. Chem.* **2010**, *8*, 2683 – 2692.

[26] A. I. Karsisiotis, N. Ma'ani Hessari, E. Novellino, G. P. Spada, A. Randazzo, M. Webba da Silva, *Angew. Chem. Int. Ed.* **2011**, *50*, 10645 – 10648; *Angew. Chem.* **2011**, *123*, 10833 – 10836.
[27] N. Foloppe, B. Hartmann, L. Nilsson, A. D. MacKerell, Jr., *Biophys. J.* **2002**, *82*, 1554 – 1569.
[28] D. Pradhan, L. H. Hansen, B. Vester, M. Petersen, *Chem. Eur. J.* **2011**, *17*, 2405 – 2413.
[29] I. Berger, V. Tereshko, H. Ikeda, V. E. Marquez, M. Egli, *Nucleic Acids Res.* **1998**, *26*, 2473 – 2480.
[30] J.-F. Trempe, C. J. Wilds, A. Y. Denisov, R. T. Pon, M. J. Damha, K. Gehring, *J. Am. Chem. Soc.* **2001**, *123*, 4896 – 4903.
[31] C. Moriou, A. D. Da Silva, M. J. V. Prado, C. Denhez, O. Plashkevych, J. Chattopadhyaya, D. Guillaume, P. Clivio, *J. Org. Chem.* **2018**, *83*, 2473 – 2478.
[32] C. G. Peng, M. J. Damha, *Nucleic Acids Res.* **2007**, *35*, 4977 – 4988.
[33] J. Dickerhoff, K. Weisz, *J. Phys. Chem. Lett.* **2017**, *8*, 5148 – 5152.
[34] N. T. Dao, R. Haselsberger, M. E. Michel-Beyerle, A. T. Phan, *ChemPhysChem* **2013**, *14*, 2667 – 2671.
[35] C. J. Lech, A. T. Phan, M. E. Michel-Beyerle, A. A. Voityuk, *J. Phys. Chem. B* **2015**, *119*, 3697 – 3705.
[36] A. Virgilio, V. Esposito, A. Randazzo, L. Mayol, A. Galeone, *Nucleic Acids Res.* **2005**, *33*, 6188 – 6195.
[37] P. L. Thao Tran, A. Virgilio, V. Esposito, G. Citarella, J.-L. Mergny, A. Galeone, *Biochimie* **2011**, *93*, 399 – 408.
[38] P. Šket, A. Virgilio, V. Esposito, A. Galeone, J. Plavec, *Nucleic Acids Res.* **2012**, *40*, 11047 – 11057.
[39] J. T. Nielsen, K. Arar, M. Petersen, *Nucleic Acids Res.* **2006**, *34*, 2006 – 2014.
[40] J. Zhou, S. Amrane, F. Rosu, G. F. Salgado, Y. Bian, H. Tateishi-Karimata, E. Largy, D. N. Korkut, A. Bourdoncle, D. Miyoshi, J. Zhang, H. Ju, W. Wang, N. Sugimoto, V. Gabelica, J.-L. Mergny, *J. Am. Chem. Soc.* **2017**, *139*, 7768 – 7779.
[41] B. Karg, L. Haase, A. Funke, J. Dickerhoff, K. Weisz, *Biochemistry* **2016**, *55*, 6949 – 6955.
[42] B. Karg, K. Weisz, *Chem. Eur. J.* **2018**, *24*, 10246 – 10252.
[43] J. Dickerhoff, K. Weisz, *Angew. Chem. Int. Ed.* **2015**, *54*, 5588 – 5591; *Angew. Chem.* **2015**, *127*, 5680 – 5683.
[44] J. Dickerhoff, L. Haase, W. Langel, K. Weisz, *ACS Chem. Biol.* **2017**, *12*, 1308 – 1315.
[45] V. V. Cheong, C. J. Lech, B. Heddi, A. T. Phan, *Angew. Chem. Int. Ed.* **2016**, *55*, 160 – 163; *Angew. Chem.* **2016**, *128*, 168 – 171.
[46] L. Haase, J. Dickerhoff, K. Weisz, *Chem. Eur. J.* **2018**, *24*, 15365 – 15371.
[47] J. Qi, R. H. Shafer, *Biochemistry* **2007**, *46*, 7599 – 7606.
[48] C.-F. Tang, R. H. Shafer, *J. Am. Chem. Soc.* **2006**, *128*, 5966 – 5973.
[49] C. J. Lech, Z. Li, B. Heddi, A. T. Phan, *Chem. Commun.* **2012**, *48*, 11425 – 11427.
[50] N. Martín-Pintado, M. Yahyaee-Anzahaee, G. F. Deleavey, G. Portella, M. Orozco, M. J. Damha, C. González, *J. Am. Chem. Soc.* **2013**, *135*, 5344 – 5347.
[51] A. Virgilio, V. Esposito, G. Citarella, A. Pepe, L. Mayol, A. Galeone, *Nucleic Acids Res.* **2012**, *40*, 461 – 475.
[52] A. Randazzo, V. Esposito, O. Ohlenschläger, R. Ramachandran, L. Mayol, *Nucleic Acids Res.* **2004**, *32*, 3083 – 3092.
[53] P. K. Dominick, M. B. Jarstfer, *J. Am. Chem. Soc.* **2004**, *126*, 5050 – 5051.
[54] J. T. Nielsen, K. Arar, M. Petersen, *Angew. Chem. Int. Ed.* **2009**, *48*, 3099 – 3103; *Angew. Chem.* **2009**, *121*, 3145 – 3149.
[55] J. Dickerhoff, K. Weisz, *ChemBioChem* **2018**, *19*, 927 – 930.
[56] J. Dickerhoff, B. Appel, S. Müller, K. Weisz, *Angew. Chem. Int. Ed.* **2016**, *55*, 15162 – 15165; *Angew. Chem.* **2016**, *128*, 15386 – 15390.

Manuscript received: October 22, 2018
Accepted manuscript online: December 4, 2018
Version of record online: February 15, 2019

# Affirmation

The affirmation has been removed in the published version.

# Curriculum Vitae

The curriculum vitae has been removed in the published version.

## Published Articles

1. Duplex-Guided Refolding into Novel G-Quadruplex (3+1) Hybrid Conformations; **Karg**, Weisz; *Angewandte Chemie*, 2019, DOI: 10.1002/anie.201905372
2. Manipulating DNA G-Quadruplex Structures Through Guanosine Analogs; Haase, **Karg**, Weisz; *ChemBioChem*, 2018, 20, 985-993.
3. Loop Length Affects Syn-Anti Conformational Rearrangements in Parallel G-Quadruplexes; **Karg**, Weisz; *Chemistry*, 2018, 24, 10246–10252.
4. Ligand-Induced Dimerization of a Truncated Parallel MYC G-Quadruplex; Funke*, **Karg***, Dickerhoff, Balke, Müller, Weisz; *ChemBioChem*,2018, 19,505–512 [*co-first authors].
5. Selective Targeting of G-Quadruplex Structures by a Benzothiazole-Based Binding Motif; Buchholz, **Karg**, Dickerhoff, Sievers-Engler, Lämmerhofer, Weisz; *Chemistry*, 2017, 23, 5814–5823.
6. Observation of a Dynamic G-Tetrad Flip in Intramolecular G-Quadruplexes; **Karg**, Haase, Funke, Dickerhoff, Weisz; *Biochemistry*, 2016, 55, 6949–6955.
7. Molecular Recognition and Visual Detection of G-Quadruplexes by a Dicarbocyanine Dye; **Karg**, Funke, Ficht, Sievers-Engler, Lämmerhofer, Weisz; *Chemistry*, 2015, 21, 13802–13811.

## Posters

1. Enforcing G-quadruplex Topological Changes, G4thering, Padua, Biology of Non-Canonical Nucleic Acids: from humans to pathogens, 26th-28th September 2018
2. NMR as a Versatile Tool in G4-Quadruplex Folding and Dynamics, International Council on Magnetic Resonance in Biological Systems, Dublin, 19th-24th August 2018
3. Tetrad Polarity Reversal and Exchange Dynamics of Single Residues in Monomolecular G4s, G4thering, Prag, 31st May - 3rd June 2017
4. Binding of an Indoloquinoline to a Parallel Quadruplex: Impact of Overhang Sequences on Ligand Affinity, XXII. International Round Table of Nucleic Acids, Paris, 18th-22th July 2016

5. Changing the Tetrad Polarity in a Parallel Monomolecular Quadruplex, Euromar, Aarhus, 3rd-7th July 2016

6. Fluorescence-Based Studies on Ligand Binding: The G-Quadruplex Perspective, 5th International Meeting on Quadruplex Nucleic Acids, Bordeaux, 26th-28th May 2015

7. Quadruplex Binding of a Symmetrical Cyanine Dye, 6. Nukleinsäurechemietreffen, Greifswald, 19th-20th September 2013

# Acknowledgements

Allen voran möchte ich Klaus Weisz danken, zuerst einmal ganz förmlich für die Bereitstellung des Themas und dann noch einmal ganz informell, für alles, für die Geduld, die ich zugegeben recht stark strapaziert habe, die Unterstützung bei allen Projekten, die in vier Jahren so aufkommen und die Freiheiten, vor allem was diese ganzen Projekte angeht, die ich so angefangen habe, aber auch die Möglichkeit meine Zeit komplett selbst einzuteilen. Danke, dass ich die vier Jahre bleiben durfte.
Besonderer Dank gilt natürlich dem ganzen AKW, Jonathan, für die vielen Ideen, die vielen Erklärungen, die vielen Gespräche; Andrea, für die wunderschöne Zeit, die Motivation und die Prokrastionation; Petra, für diese tolle familiäre Atmosphäre in der Gruppe und der ganze Rest der Bande, Linn, Lena und die alten Hasen. Ihr habt die Zeit bereichert.
Außerdem möchte ich bei einigen anderen Gruppen danken, vor allem der Organsichen Chemie für die Flurfreundschaft, insbesondere Robert, Jenny, Claudia und Bettina, die immer ein offenes Ohr und nützliche Problemlösungen parat hatten. Der physikalischen Chemie, vor allem Norman, möchte ich für die vielen Nachhilfestunden in Informatik und vor allem für die Rechenkapazität auf Griffin danken. Mein Intermezzo mit der Lehre in der Biochemie hat sich als erstaunlich angenehm entpuppt durch euch, Gottfried, Leona und Britta.
Zuletzt möchte ich meiner Familie danken, die mir das Studium überhaupt erst ermöglicht hat. Euer bedingungsloses Verständnis für alles, was ich hier mache, ist der Grund, warum ich es überhaupt machen kann. Und dann noch Hendrik, ohne dich hätte ich wohl fünfmal so lange gebraucht, davon mal ganz abgesehen, danke.