

High-Resolution Insight into G-Overhang Architecture

Robert Hänsel,^{*,†} Frank Löhr,[†] Lukáš Trantírek,^{‡,§} and Volker Dötsch^{*,†}

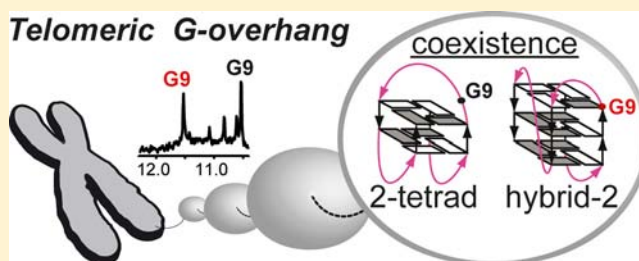
[†]Institute of Biophysical Chemistry and Center for Biomolecular Magnetic Resonance, Goethe University, Frankfurt/Main, Germany

[‡]Bijvoet Centre for Biomolecular Research, Utrecht University, Utrecht, The Netherlands

[§]Central European Institute of Technology, Masaryk University, Brno, Czech Republic

Supporting Information

ABSTRACT: NMR and fluorescence spectroscopy were used to address the effect of intracellular molecular crowding and related hydration on a model telomeric G-quadruplex (G4) DNA structure (d(AG₃(TTAGGG)₃). d(AG₃(TTAGGG)₃) prevalently adopted the hybrid-1 conformation *in vivo*, *ex vivo*, and in dilute potassium-based solution, while it formed the parallel propeller fold in water-depleted potassium-based solution, a commonly used model system for studying intracellular molecular crowding. The dilute potassium-based solution appeared to imitate the properties of the cellular environment required for d(AG₃(TTAGGG)₃) folding under *in vivo* and *ex vivo* conditions. High-resolution NMR investigations of site-specifically ¹⁵N-labeled G4 units in native-like single-stranded telomeric DNA revealed that the 3'-terminal and internal G4 unit predominantly coexist in 2-tetrad antiparallel basket and hybrid-2 structures that are arranged in "beads-on-a-string"-like fashion. Our data provide the first high-resolution insight into the telomeric G-overhang architecture under essentially physiological conditions and identify the 2-tetrad antiparallel basket and hybrid-2 topologies as the structural targets for the development of telomere-specific G4 ligands.



INTRODUCTION

Vertebrate telomeric DNA consists of repeating double-stranded d(GGGTTA)_n/d(CCCAAT)_n elements and terminates in a 50–200 nucleotide long guanine-rich 3' single strand (G-overhang).¹ The G-overhang is considered to be an evolutionary conserved structural element that is essential for chromosome end protection.² These single-stranded protrusions have been observed to adopt G-quadruplex (G4) structures *in vitro* and *in vivo*.^{3–5} Various small molecules such as telomestatin have been shown to selectively bind and stabilize telomeric G4 structures relative to double-stranded DNA. The use of these drugs with mammalian cancer cells results in DNA damage-response-mediated death, making G4 structures promising targets for anticancer chemotherapy.^{6–8} Several molecules that bind and stabilize G4 structures have shown anticancer activity in tumor xenograft models and have entered clinical trials.⁶ However, *in vivo*, all these small molecules failed to selectively target telomeric G4-DNA relative to other G4-forming regions in the genome, e.g., upstream of oncogenic transcription start sites.^{7,9} Therefore, it remains to be demonstrated if stabilization of both telomeric G4 and other putative G4-DNA sites, or stabilization of the telomeric G4 alone, is sufficient to provide anticancer activity. Despite the high interest in designing small molecules that selectively target a particular G4-DNA structure,^{10,11} the conformations adopted by G4-DNA inside living cells from vertebrates are not known.⁷

Telomeric G-repeat sequences from vertebrates rapidly fold into G4 and prefer to coordinate potassium over sodium

ions.^{12,13} Thus far, NMR and X-ray investigations of various short telomeric DNA constructs (<44 nucleotides), containing four-to-seven guanine tracts of the telomeric G-rich repeat d(GGGTTA), revealed five distinct single intramolecular G4 structures.^{14,15} One of the most studied model sequences has been the d(AG₃(TTAGGG)₃) oligonucleotide. While in sodium solution d(AG₃(TTAGGG)₃) adopts the 3-tetrad antiparallel basket G4 structure,¹⁶ the very same sequence in the presence of potassium crystallizes as the flattened propeller loop-shaped parallel-stranded conformation.¹⁷ In dilute potassium-based solution, d(AG₃(TTAGGG)₃) was, however, shown to exist in multiple (unknown) conformations distinct from those observed in sodium solution or in the crystalline state.¹⁸ A combined NMR and CD analysis carried out in potassium solution suggested that d(AG₃(TTAGGG)₃) coexists in at least two well-defined conformations.¹² Of note, independent of length and sequence composition, telomeric sequences were shown to adopt the parallel propeller G4 structure in water-depleted potassium solutions, such as in potassium solutions supplemented with 40–50% poly(ethylene glycol) (PEG) or ethanol.^{19–22} Since these *in vitro* conditions have been postulated to closely mimic the crowded environment of a cell's nucleus,²³ the parallel propeller G4 topology has been suggested to be the physiologically relevant conformation of telomeric G-DNA.^{19,21,22,24} This model recently found support

Received: December 19, 2012

Published: January 22, 2013

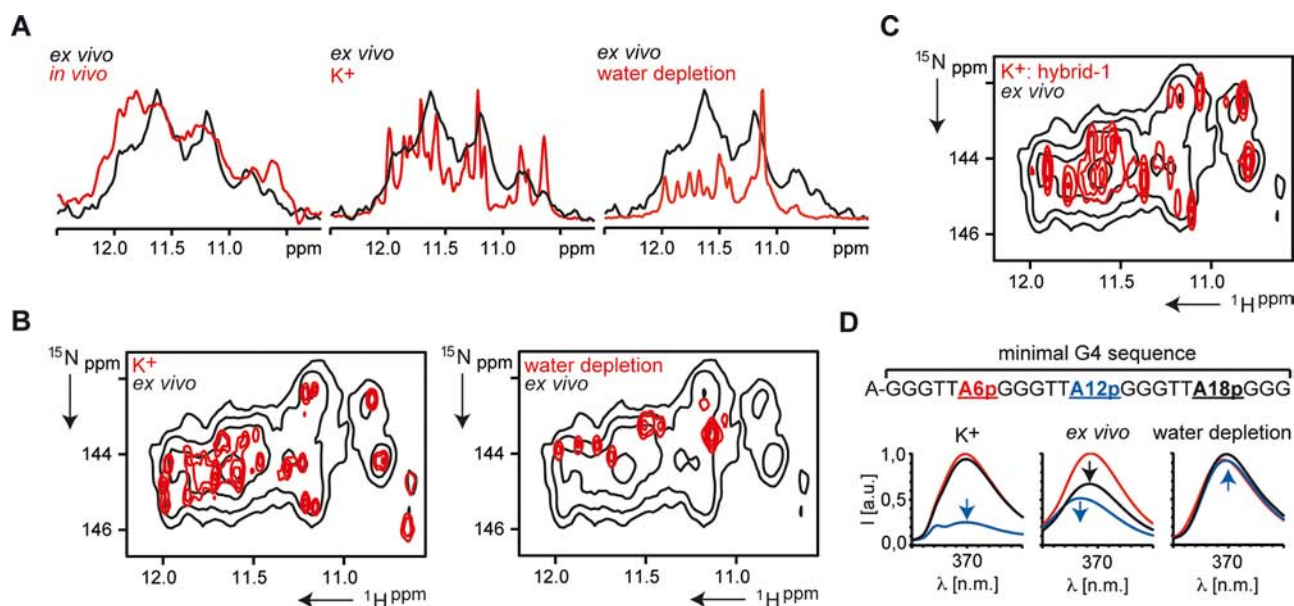


Figure 1. d(AG₃(TTAGGG)₃) adopts hybrid-1 conformation under physiologically relevant conditions. (A) Overlay of ¹⁵N-edited 1D imino proton spectra of ¹⁵N/¹³C-labeled d(AG₃(TTAGGG)₃) folded in *X. laevis* egg extract (*ex vivo*) with 1D spectra acquired inside *X. laevis* oocytes (*in vivo*), in dilute potassium-based solution (K⁺), and in potassium-based solution supplemented with 40% PEG 200 (water depletion). (B) Comparison of 2D ¹H-¹⁵N sHMQC spectra of the imino region of ¹⁵N/¹³C-labeled d(AG₃(TTAGGG)₃) folded in *X. laevis* egg extract (*ex vivo*), in dilute potassium-based solution (K⁺), and in potassium-based solution supplemented with 40% PEG 200 (water depletion). (C) Overlay of the 2D ¹H-¹⁵N sHMQC spectrum of d(AG₃(TTAGGG)₃) recorded *ex vivo* with the 2D spectrum of d(TAG₃(TTAGGG)₃) observed in dilute potassium-based solution, reported to form the hybrid-1 G4 structure. (D) Overlay of steady-state fluorescence emission spectra of d(AG₃(TTAGGG)₃), site-specifically 2-aminopurine modified at position A6, A12 or A18, acquired in dilute potassium-based solution (K⁺), *ex vivo*, and in potassium-based solution supplemented with 40% PEG 200 (water depletion).

in studies by Singh et al. (2009) and Azarkh et al. (2012), which indicated that the parallel propeller conformation is one of the major conformations of d(AG₃(TTAGGG)₃) in deeply frozen dilute potassium-based solution, *Xenopus laevis* (*X. laevis*) oocytes, and *X. laevis* oocyte extracts using double electron-electron resonance (DEER).^{25,26} In contrast, NMR investigations at physiological temperature of various non-isotopically labeled telomeric G4-DNA suggested that the parallel conformation is not the predominant G4 conformation inside living *X. laevis* oocytes²⁷ nor in the *X. laevis* egg extract.²⁰ In addition, applying different molecular crowding conditions using 400 g/L Ficoll 70 or 300 g/L BSA instead of molar concentrations of PEG failed to convert telomeric G4-DNA into the parallel topology.^{20,28} Recently, Wang and co-workers demonstrated that PEG promotes the formation of parallel G4 conformations by changing its local water structure.²⁹ To elucidate the physiologically relevant effects of the intracellular crowded environment and related intracellular local water activity on the conformation of telomeric G4-DNA, high-resolution investigations under native cell-like conditions, preferably *in vivo*, are needed.

Numerous studies using short telomeric DNA constructs have demonstrated so far that, in addition to environmental factors, the G4 conformation depends on 5'- and 3'-flanking residues immediately adjacent to the core d(G₃(TTAGGG)₃) G4 sequence.³⁰ These observations raised concerns about whether the structural behavior of G4 units in the telomeric G-overhang might be realistically assessed on the basis of studies of short model telomeric sequences forming single G4 units. Circular dichroism (CD), polyacrylamide gel electrophoresis (PAGE), and atomic force microscopy (AFM) studies indicated that long telomeric sequences preferably fold into contiguous

intramolecular G4 structures.^{19,31–35} Based on the structural information obtained for short model telomeric sequences, Dai and co-workers suggested that the telomeric G-overhang might adopt contiguously stacked hybrid G4 units.³⁶ By using CD and native PAGE, Renciuik and co-workers proposed that G4 units in the context of extended telomeric sequences adopt contiguous 3-tetrad antiparallel basket conformations in dilute potassium solution, which dynamically convert step-by-step into hybrid and parallel structures in potassium solution supplemented with 50% ethanol or upon dramatic increase of the DNA concentration.¹⁹ Using a similar approach to mimic the cell's nuclear crowded interior, Xu et al. (2011) and Yu et al. (2012) showed that G4 units in extended telomeric DNA dynamically convert into the parallel conformation in potassium solution supplemented with PEG.^{22,37} Recently CD, differential scanning calorimetry (DSC), fluorescence, and analytical ultracentrifugation (AUC) experimental data were used to formulate a model suggesting that individual G4 units in extended telomeric DNA exist in contiguous interlocked hybrid-1 and hybrid-2 structures.^{38,39} The general features of this model were supported by a recent *in vitro* electron paramagnetic resonance (EPR)⁴⁰ investigation. This study showed that experimentally derived distances between site-specifically located paramagnetic labels within long telomeric repeat sequences are not compatible with the parallel G4 fold but rather correspond to those expected in the hybrid-1 and/or hybrid-2 conformations.⁴⁰ However, none of the proposed models has been verified under physiological conditions by direct, high-resolution experimental data.

In this study, high-resolution NMR spectroscopy and steady-state fluorescence spectroscopy are used to characterize folding topologies of several model DNA constructs derived from the

telomeric G-overhang, under *in vivo* conditions in living *X. laevis* oocytes, under *ex vivo* conditions represented by *X. laevis* egg extract, and under various conditions *in vitro*. Comparison of the NMR and fluorescence spectral data allows us to assess the role of intracellular molecular crowding and local water activity on G4 unit structure in the context of sequences imitating the architecture of the telomeric G-overhang. Our data not only resolve existing controversies relating to the physiological relevance of the telomeric parallel propeller G4 folding topology but also, importantly, provide for the first time high-resolution insight into the architecture of extended telomeric DNA, which serves as a model for the telomeric G-overhang.

RESULTS

d(AG₃(TTAGGG)₃) Predominantly Adopts the Hybrid-1 Conformation under Physiologically Relevant Conditions. So far we and other groups have demonstrated that buffers that mimic intracellular molecular crowding and related hydration properties by adding PEG or ethanol stabilize the parallel conformation, independent of the telomeric G4-DNA sequence composition.^{19–22} In contrast, the parallel conformation was not observed as a dominant species in solutions containing 400 g/L Ficoll 70 or in *X. laevis* egg extract.²⁰ Due to the low resolution, however, we were not able to identify which conformation(s) G4 adopt under these more native-like molecular crowding conditions. Here, in order to directly identify the topology of G4 under native molecular crowding conditions, we characterized the conformation of the often-studied model sequence d(AG₃(TTAGGG)₃) using NMR spectroscopy. We investigated the ¹⁵N-labeled d(AG₃(TTAGGG)₃) sequence under various conditions, including a dilute potassium-based buffer resembling the intracellular salt composition of *X. laevis* oocytes, inside living *X. laevis* oocytes, and in two commonly applied conditions used to mimic intracellular molecular crowding, namely *X. laevis* egg extract and potassium-based solution supplemented with 40% PEG 200 (Figure 1A). The 1D imino spectra of 100 μM d(AG₃(TTAGGG)₃) obtained in dilute potassium-based solution, *in vivo*, and in the extract showed similar resonance profiles. However, these profiles were significantly different from that acquired in potassium-based solution supplemented with 40% PEG 200. d(AG₃(TTAGGG)₃) has been shown to adopt the parallel propeller G4 conformation in potassium solution supplemented with 40% PEG 200.²¹ Comparison of the imino spectral fingerprints thus indicates that the parallel propeller is not the preferred d(AG₃(TTAGGG)₃) conformation under *in vivo* conditions.

The measurement inside living *X. laevis* oocytes yielded a valuable reference NMR spectral fingerprint reporting on the d(AG₃(TTAGGG)₃) conformation under native conditions. However, the inherent inhomogeneity of the in-cell NMR sample and the need to use low, non-apoptotic DNA concentrations restricted us to record only 1D NMR spectra. While the information content of these spectra is sufficient to exclude parallel conformations from being significantly populated, these spectra cannot be used to identify which conformation(s) are adopted. We therefore performed 2D NMR measurements in *X. laevis* egg extract. The *X. laevis* egg extract preserves most of the activity of living *X. laevis* eggs and has been shown to be a reliable model of an intracellular environment for structural studies.^{41,42} In contrast to the NMR measurements in living *X. laevis* oocytes, the use of *X. laevis* egg

extract allows to increase the d(AG₃(TTAGGG)₃) concentration, which along with increased sample homogeneity offers the possibility to characterize the conformational state of d(AG₃(TTAGGG)₃) by 2D NMR spectroscopy. The 2D ¹H–¹⁵N sofast-heteronuclear multiple quantum coherence (sfHMQC) imino resonance spectrum of 200 μM d(AG₃(TTAGGG)₃) folded in *X. laevis* egg extract was recorded, and the resulting 2D imino spectral fingerprint was compared with corresponding fingerprints of d(AG₃(TTAGGG)₃) in dilute potassium-based solution and in potassium-based solution supplemented with 40% PEG 200 (Figure 1B). This 2D spectral analysis revealed that the imino spectral pattern recorded in *X. laevis* extract is virtually identical to the pattern observed in dilute potassium-based solution. However, the 2D spectral fingerprint acquired in *X. laevis* extract was significantly different from the spectrum observed in dilute potassium-based solution supplemented with 40% PEG 200 (Figure 1B), suggesting again that the parallel propeller is not formed within the *X. laevis* egg extract. Instead, the data support our *in vivo* 1D NMR finding that, under physiological conditions, d(AG₃(TTAGGG)₃) exists in conformation(s) observed also in dilute potassium-based solution (Figure 1A). To identify the physiologically relevant conformation(s), the 2D imino fingerprint of d(AG₃(TTAGGG)₃) obtained in extract was compared with the fingerprints of d(TAG₃(TTAGGG)₃), d(TTAG₃(TTAGGG)₃TT), and d(G₃(TTAGGG)₃T) telomeric DNA, known to adopt the hybrid-1,⁴³ hybrid-2,³⁶ and 2-tetrad antiparallel basket⁴⁴ topologies, respectively, in dilute potassium-based solution (Figure 1B,C and Supporting Information, Figure S1A,B). Moreover, we compared the spectral profile of d(AG₃(TTAGGG)₃), known to form the 3-tetrad antiparallel basket¹⁶ conformation in pure sodium solution, with the 2D imino profile of d(AG₃(TTAGGG)₃) acquired in *X. laevis* egg extract (Figure S1C). These comparisons showed that the 2D imino fingerprint acquired *ex vivo* overlaps with the pattern observed for the hybrid-1 structure, indicating that d(AG₃(TTAGGG)₃) predominantly adopted the hybrid-1 G4 structure under physiologically relevant conditions (Figures 1C and S1).

To independently confirm these NMR results, we systematically investigated the G4 loop conformations of 2-aminopurine-modified d(AG₃(TTAGGG)₃) in *X. laevis* egg extract, in dilute potassium-based solution, and in potassium-based solution supplemented with 40% PEG 200 by steady-state fluorescence spectroscopy (Figures 1D and S2). Figure 1D shows the overlay of fluorescence emission profiles for d(AG₃(TTAGGG)₃), labeled with 2-aminopurine at the A6, A12, and A18 positions, in these different conditions. While in the presence of *X. laevis* extract A12p showed a pronounced quenched and shifted emission profile, indicating that A12p was exposed to a chemical environment distinct from those of A6p and A18p, the A12p emission profile recorded in potassium solution supplemented with 40% PEG 200 was indistinguishable from the emission profiles of A6p and A18p, indicating virtually identical chemical environments for all three labels (Figure 1D). In this respect, fluorescence measurements, in accord with the NMR data, indicate that d(AG₃(TTAGGG)₃) in *X. laevis* egg extract adopts a different conformation than it does in potassium solution supplemented with 40% PEG 200. However, while NMR data indicate that d(AG₃(TTAGGG)₃) in *X. laevis* egg extract adopts the same conformation as it does in dilute potassium solution, the corresponding emission spectra, particularly that of A12p, recorded in dilute

potassium-based solution differ from corresponding emission spectra acquired in *X. laevis* egg extract (Figure 1D). To identify the source of spectral differences, the cellular proteins and their complexes were removed from the *X. laevis* egg extract via thermal denaturation. In contrast to the high emission background of the *X. laevis* egg extract, this so-called cleared lysate is essentially devoid of autofluorescence (Figure S2A,B). The emission spectrum of A12p recorded in nearly background-free cleared lysate was reminiscent of the A12p spectrum recorded in dilute potassium-based solution, suggesting that the autofluorescence of the *X. laevis* egg extract rather than distinct structural arrangement was responsible for the different shifts of the emission maxima. In contrast to the results obtained for A6p and A12p, the emission profile measured for A18p folded in cleared lysate remained more quenched compared to the corresponding emission profile acquired in dilute potassium-based solution (Figures 1D and S2). To test whether the alteration in emissions of A18p arose from pronounced structural changes of the G4-DNA construct, the 2D imino NMR region of ^{15}N -labeled $\text{d}(\text{AG}_3(\text{TTAGGG})_3)$ was acquired in cleared lysate. Analysis of the cleared lysate G4-DNA imino region with the spectral fingerprint obtained in dilute potassium-based solution suggested that $\text{d}(\text{AG}_3(\text{TTAGGG})_3)$ adopted virtually identical conformations in both environments (Figure S2C). This suggests that the differences seen in the fluorescence spectra of the A18p in cleared lysate and dilute potassium-based solution are not structure-based but possibly caused by components of the cleared lysate.

Altogether, fluorescence emission profiles further support the NMR data interpretation suggesting that the parallel propeller G4 conformation is not the preferred folding topology of $\text{d}(\text{AG}_3(\text{TTAGGG})_3)$ in *X. laevis* egg extract (Figure 1D).

5'- and 3'-Flanking-Dependent Conformational Polymorphism Observed for Single G4 Sequences Is Diminished in Long Telomeric Sequences Forming Multiple G4 Units. Numerous studies have indicated that the conformation of the minimal intramolecular telomeric G4 sequence $\text{d}(\text{G}_3(\text{TTAGGG})_3)$ depends on 5'- and 3'-flanking residues adjoining the G4 core sequence. Systematically changing the nucleotide sequence at the 3'-end from no nucleotide to $\text{d}(\text{TTA})$ nucleotides did not change the G4 conformation. In contrast, variation of the 5'-flanking from $\text{d}(\text{TA})$ to no nucleotide changed the G4 conformation from the hybrid-1 to the 2-tetrad antiparallel basket conformation, while the simultaneous addition of nucleotides at both the 5'- and the 3'-ends suppressed the 2-tetrad antiparallel basket and promoted the hybrid-2 formation (Figure S3).^{36,44} However, except for the 5'- and 3'-terminal G4 units, individual G4 units present within the telomeric G-overhang are linked with each other via $\text{d}(\text{TTA})$ linkers, thus defining the local flanking environment for all internal G4 units at both the 5'- and 3'-ends. Of note, a variable flanking sequence can, however, occur at the 3'-end of the 3'-terminal G4 unit.

To assess the effect of the 3'-nucleotide flanking sequence on the structure and stability of tandem G4 units $\text{d}(\text{G}_3(\text{TTAGGG})_3\text{TTAG}_3(\text{TTAGGG})_3)$, here referred to as the double G4 unit construct, we used a combination of 1D NMR and CD spectroscopy to monitor overall changes in structure and stability. CD and NMR spectra and melting profiles obtained for the double G4 unit constructs with different 3'-nucleotide flanking were virtually identical (Figure S4A–C). Moreover, 1D NMR imino regions of the double G4 unit

construct with 3'-TT flanking were compared to the imino region of the triple G4 unit construct containing an additional 5'-G4 unit (Figure S4D). The addition of another G4 unit to the 5'-end of the double G4 unit construct resulted in virtually identical imino spectral profiles, suggesting that, in both double and triple G4 unit constructs, G4 units adopt similar structures (Figure S4D). In conclusion, these investigations indicated that, in double G4 unit constructs, the addition of 3'-nucleotides and whole 5'-G4 units to the double G4 unit construct had no significant influence on the overall G4 unit structure.

G4 Units in G-Overhang Sequences Coexist in 2-Tetrad Antiparallel Basket and Hybrid-2 Conformations.

In this study, we separately investigated the structure of the 3'-terminal G4 unit naturally located at the very end of the G-overhang and the internal G4 unit, which represents the most abundant G4 unit in the telomeric G-overhang. To characterize the conformation of the 3'-terminal G4 unit, we used three site-specifically ^{15}N isotopically labeled G4 constructs in which the 3'-terminal G4 unit was connected at the 5'-end to a G4 unit via the $\text{d}(\text{TTA})$ linker and at the 3'-end to the flanking $\text{d}(\text{TT})$ nucleotides (Figure 2A). In the single-site-labeled constructs,

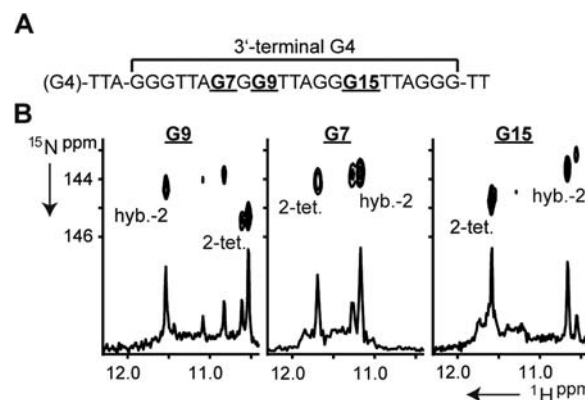


Figure 2. Conformations adopted by the 3'-terminal G4 unit. (A) Sequence of the site-specifically ^{15}N -labeled DNA constructs based on eight telomeric G-tracts. The label positions are shown in bold and are underlined. (B) ^{15}N -edited 1D ^1H and 2D ^1H - ^{15}N sfHMQC spectra of G7, G9, and G15 guanine ^{15}N -labeled constructs recorded in dilute potassium-based solution. Dominant signals were unambiguously assigned to hybrid-2 and 2-tetrad antiparallel basket topologies (Figure S5 and Table S1).

G7, G9, or G15 was ^{15}N isotopically labeled, as their unique imino proton/nitrogen NMR spectral fingerprints allow discrimination between the hybrid-1, hybrid-2, parallel, 2-tetrad antiparallel basket, and 3-tetrad antiparallel basket topologies (Figure S5 and Table S1).^{16,21,36,43,44} Figure 2B shows imino proton/nitrogen resonances of the ^{15}N -edited 1D and ^1H - ^{15}N 2D sfHMQC spectra of different site-specifically G7-, G9-, and G15-labeled 3'-terminal G4 unit constructs recorded in dilute potassium-based solution. The 1D and 2D spectra of G7, G9, and G15 site-specifically labeled 3'-terminal G4 unit constructs showed two dominant peaks having essentially the same intensities (Figure 2B). This indicates that the 3'-terminal G4 unit coexists in two distinct conformations with similar abundance. To identify these two conformations, the dominant peak positions in the ^{15}N -edited 1D ^1H sfHMQC and ^1H - ^{15}N 2D sfHMQC spectra were compared with the corresponding peak positions in reported monomeric reference G4 structures, namely hybrid-1, hybrid-2, parallel, 2-tetrad antiparallel basket,

and 3-tetrad antiparallel basket (Figure S5 and Table S1). Analysis of the chemical shifts for the specifically G9-labeled construct unambiguously assigned the proton/nitrogen signal at 10.53/145.3 ppm to the 2-tetrad antiparallel basket structure (Figures 2B and S5 and Table S1). In contrast, the other prominent G9 proton/nitrogen resonance at 11.53/144.4 ppm can be assigned to either hybrid-1 or hybrid-2 structure (Figure S5 and Table S1). Analysis of the chemical shifts observed in the G7-labeled 3'-terminal G4 unit unambiguously identified the 2-tetrad antiparallel basket and hybrid-2 conformations. In combination, the results from both G9 and G7 labeling identified the 2-tetrad antiparallel basket and hybrid-2 as coexisting conformations in the 3'-terminal G4 unit (Figures 2B and S5 and Table S1). This analysis was further confirmed by a third sample, specifically labeled at position G15. This sample again showed two dominant peaks with chemical shifts characteristic for the hybrid-2 structure (10.67/143.6 ppm) and the hybrid-1 or 2-tetrad antiparallel basket (11.58/144.7 ppm), respectively.

Four site-specifically ^{15}N -labeled oligonucleotide constructs, each comprised of twelve G-tracts, were used to characterize conformation(s) of the internal G4 unit in the context of the telomeric G-overhang. The internal G4 unit under study was site-specifically labeled with ^{15}N at positions G7, G9, G15, and G21 and connected to a 5'-G4 unit via the d(TTA) linker and to the 3'-G4-TT. The labeling scheme is depicted in Figure 3A. Similar to the situation for the 3'-terminal G4 unit, the spectra for the G9-labeled internal G4 unit showed two dominant peaks with roughly equal intensities, one of which could be

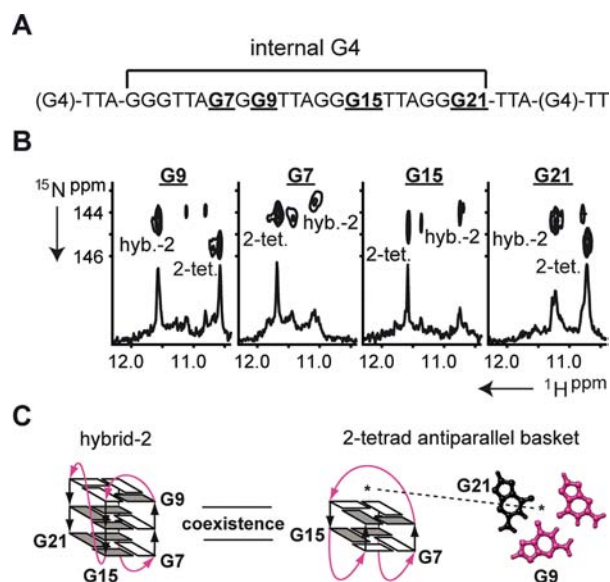


Figure 3. Conformational preferences of the internal G4 unit. (A) Sequence of the site-specifically ^{15}N -labeled DNA constructs based on twelve telomeric G-tracts. The label positions are shown in bold and are underlined. (B) ^{15}N -Edited 1D ^1H and 2D ^1H - ^{15}N sfHMQC spectra of G7, G9, G15, and G21 guanine ^{15}N -labeled constructs recorded in dilute potassium-based solution. Dominant signals of individual guanines labeled in the internal G4 unit can unambiguously be assigned to either hybrid-2 or 2-tetrad antiparallel basket topologies (Figure S5 and Table S1). (C) Schematic drawings of the telomeric G4-DNA secondary hybrid-2 and 2-tetrad antiparallel basket structures. Guanines that were studied are marked and are illustrated bold. Loops are colored magenta, and anti and syn guanines are gray and white, respectively.

unambiguously assigned to the 2-tetrad antiparallel basket (10.59/145.5 ppm) and the other to either hybrid-1 or hybrid-2 conformations (11.58/144.4 ppm) (Figure 3B). The combined results from chemical shift analysis of the G7-, G15-, and G21-labeled samples demonstrated the presence of 2-tetrad antiparallel basket and hybrid-2 conformations, but not of hybrid-1, parallel, or 3-tetrad antiparallel basket conformations, in the internal G4 unit constructs. However, while the two dominant signals of the G9-labeled constructs showed similar intensities in both 1D ^{15}N -edited 1D ^1H sfHMQC and ^1H - ^{15}N 2D sfHMQC spectra, the imino signals of G7, G15, and G21, corresponding to the hybrid-2 conformation, had notably lower intensities compared to the signals observed for the 2-tetrad antiparallel basket conformation. G9 is part of the top tetrad, while G7, G15, and G21 are located in the bottom tetrad of the hybrid-2 structure (Figure 3C). The differences in signal intensities observed for the internal G4 unit in comparison to equal signal intensities observed for the 3'-terminal G4 unit suggest that the internal G4 unit adopts, besides the well-defined 2-tetrad antiparallel basket, the hybrid-2 structure with disturbed bottom tetrad formation.

Influence of Native-like Molecular Crowding on G-Overhang Structure. To assess the relevance of our conformational analysis carried out in dilute potassium-based solution, we investigated the conformational state of the 3'-terminal G4 unit present within the d((G4)-TTA-(G4)-TT) construct in *X. laevis* egg extract by NMR spectroscopy. We recorded ^{15}N -edited ^1D ^1H imino fingerprints of 50 μM selectively G7-, G9-, and G15-labeled 3'-terminal G4 unit constructs. Due to the inherent sensitivity loss caused by the heterogeneous sample state and low concentration of the labeled construct, we failed to observe any traces of the G4 imino resonances. Nevertheless performing the same experiment in the presence of 500 μM non-labeled DNA, we were able to detect the resonances corresponding to labeled 3'-terminal G4 unit constructs folded in the presence of *X. laevis* egg extract (Figure S6). In this setup, the 500 μM non-labeled DNA, which was mixed into *X. laevis* egg extract prior the addition of 50 μM ^{15}N -labeled DNA, served only as a "bait" to trap all potential binding partners from the extract that otherwise impaired the detection of 50 μM ^{15}N -labeled DNA. Notably, for all labeled constructs acquired in *X. laevis* egg extract, the spectra were virtually identical with those recorded in dilute potassium-based solution (Figure S6).

To further investigate the influence of a native-like crowded environment, G4 loop conformations of the 2-aminopurine-labeled 3'-terminal G4 unit constructs were studied using steady-state fluorescence spectroscopy. Investigations revealed that conformations observed in dilute potassium-based solution correspond to those folded inside a cell-like crowded environment, but not with spectra obtained in potassium-based solution supplemented with 40% PEG 200 (Figure S7).

These observations, together with the results from the model d(AG₃(TTAGGG)₃) sequence, indicate that the conformation of G4 units in the telomeric G-overhang is not modulated by native molecular crowding and that dilute potassium-based solutions imitate all non-specific environmental properties essential for physiologically relevant G4 folding in the natively crowded *X. laevis* egg extract.

DISCUSSION

Previous investigations of various telomeric sequences have demonstrated a remarkable dependence of the G4 structure

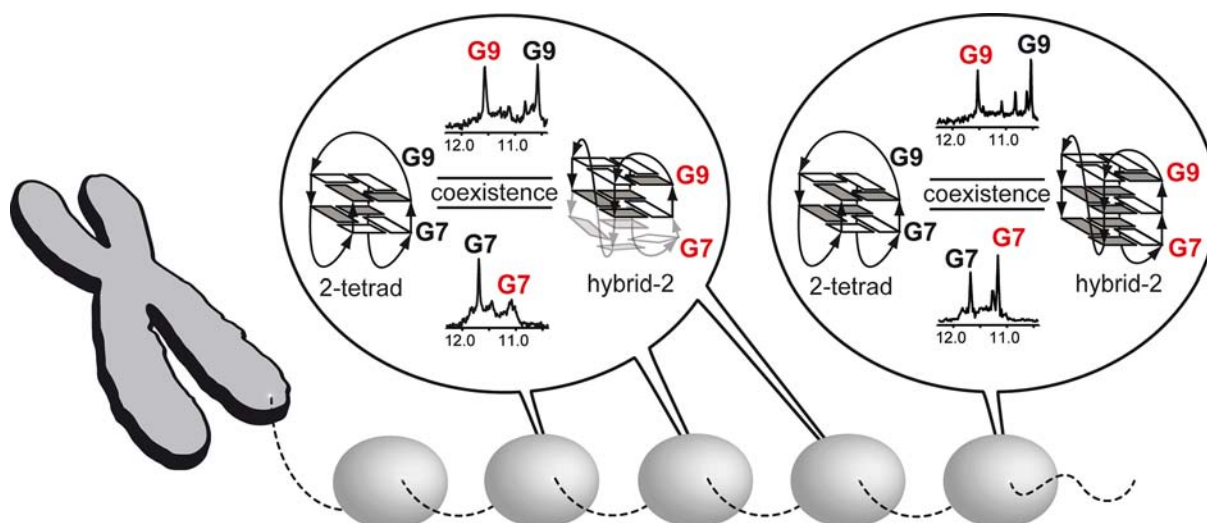


Figure 4. Beads-on-a-string model of the telomeric G-overhang from vertebrates. 3'-Terminal and internal G4 units coexist in two prevalent conformations, namely hybrid-2 and 2-tetrad antiparallel basket (2-tetrad) topology. Of note, the bottom tetrad of the hybrid-2 conformation formed by internal G4 units is structurally disturbed in comparison to the well-defined appearance in the 3'-terminal G4 unit. Schematics of G4 structures are illustrated. Backbone strand orientation is colored black, and anti and syn guanines are gray and white, respectively.

and stability^{45,46} on non-specific environmental factors such as ion composition, hydration state, or molecular crowding, for example, and on G4 loop length and residues flanking the G4 core sequence. These studies clearly suggested that both environmental conditions inside cells and the native sequence context ought to be taken into account when addressing physiologically relevant structure and stability of G-overhang DNA containing more than seven G-tracts.

In this study we investigated the conformation of $d(\text{AG}_3\text{-(TTAGGG)}_3)$, one of the most studied G4-DNA model sequences forming an intramolecular G4 structure, under various environmental conditions. The structure of $d(\text{AG}_3\text{-(TTAGGG)}_3)$ was investigated *in vivo* in living *X. laevis* oocytes, *ex vivo* in natively crowded cell-extract of *X. laevis* eggs and under different *in vitro* conditions using NMR and steady-state fluorescence spectroscopy. While previous *in vitro* studies demonstrated that $d(\text{AG}_3\text{-(TTAGGG)}_3)$ adopts 3-tetrad antiparallel basket and parallel propeller conformations in sodium solution and under water-depleted conditions, i.e., in the crystalline state and in potassium-based solution supplemented with 40% PEG 200, respectively, our data revealed that $d(\text{AG}_3\text{-(TTAGGG)}_3)$ preferentially adopts the hybrid-1 conformation under natively crowded conditions *in vivo*, *ex vivo*, and in dilute potassium-based solution. Our in-cell NMR study makes $d(\text{AG}_3\text{-(TTAGGG)}_3)$ the first reported case of a DNA molecule for which the unique set of high-resolution crystallographic, solution NMR, EPR, and *in/ex vivo* structural data is available, thus allowing us to investigate the effects of intracellular molecular crowding and hydration on its conformation. This analysis revealed that the structural behavior of $d(\text{AG}_3\text{-(TTAGGG)}_3)$ under *in/ex vivo* conditions is insensitive to native molecular crowding but rather, in agreement with the recent study by Miller and co-workers,²⁸ predominantly governed by DNA hydration. Considering the $d(\text{AG}_3\text{-(TTAGGG)}_3)$ structure as a sensitive reporter of water activity, our data strongly suggest that the hydration state of the nucleus of living cells is comparable to a dilute potassium-based solution commonly used in NMR studies.

It needs to be mentioned that the results of our in-cell NMR data that address the structural behavior of $d(\text{AG}_3\text{-}$

$\text{TTAGGG})_3$) differ from those of a recent in-cell and *in vitro* DEER study employing the very same sequence in the very same cellular model as reported here, i.e., *X. laevis* oocytes. DEER has been reliably established to characterize nucleic acid conformation *in vitro* and *in vivo* by measuring the distance (1.5–7.5 nm) between two spin-labels, serving as paramagnetic probes, that are attached to different nucleotide bases in the oligonucleotide sequence under study.⁴⁷ While our in-cell NMR data indicated hybrid-1 as the major conformation in living *X. laevis* oocytes at physiological temperature, the in-cell DEER data suggested that $d(\text{AG}_3\text{-(TTAGGG)}_3)$ coexisted as 3-tetrad antiparallel basket and parallel propeller conformations in deep-frozen *X. laevis* oocytes at cryogenic temperature. In the in-cell DEER study, the observed mean distances between the spin-label (nitroxide radical)-modified thymidines in the $d(\text{AG}_3\text{-(TTAGGG)}_3)$ construct were anticipated to reflect *in silico*-derived distances between respective thymidine C7–C7 methyl carbon atoms of the parallel propeller and 3-tetrad antiparallel basket structures.^{25,26,40} Besides the differences in experimental conditions (temperature) and the presence of nitroxide radicals attached to the DNA construct(s), the difference in interpretational outcome of our in-cell NMR and the reported in-cell DEER studies could also stem from the fact that the experimentally derived DEER distances (1.8 ± 0.2 and 3.0 ± 0.1 nm)^{25,26} cannot be unambiguously assigned to the parallel propeller (1.3–1.7 nm) and 3-tetrad antiparallel basket (2.2–3.0 nm) conformation based on *in silico*-derived distances due to the poor definition of the thymine positions in ensembles of NMR-derived hybrid-1 structures (0.8–2.8 nm) (Table S2).^{25,26}

To gain high-resolution insight into the architecture of the telomeric G-overhang, we investigated the structure of the 3'-terminal and internal G4 units. Our results on sequences comprising eight and twelve G-tracts, forming two and three G4 units, clearly showed that the conformational behavior of G4 units in these long constructs differed from that observed for four G-tract sequences, forming single G4 units. This observation is fully supported by recent CD, DSC, and AUC data.^{19,31,39} In contrast to the established models based on low-resolution data, proposing that the telomeric G-overhang

consists of consecutive blocks of either 3-tetrad antiparallel basket,¹⁹ hybrid,⁴⁰ or interlocked³⁹ hybrid-1 and hybrid-2 G4 units, our high-resolution NMR investigations of the site-specifically labeled 3'-terminal and internal G4 units suggests that both G4 units independently fold into hybrid-2 and 2-tetrad antiparallel basket conformations in dilute potassium-based solution (Figures 2 and 3). However, while the 3'-terminal G4 unit exists in well-defined hybrid-2 and 2-tetrad antiparallel basket structures, the internal G4 unit folded into well-defined 2-tetrad antiparallel basket and hybrid-2 conformations with disturbed bottom-tetrad formation. The observation that the 2-tetrad antiparallel basket conformation was identified as one of the major structures of G4 units in extended telomeric DNA finds support in a CD spectroscopy study by Renciuik and co-workers¹⁹ showing that the ellipticity profiles of telomeric sequences with increasing length are reminiscent of the profiles for the basic G4 unit d(G₃-TTAGGG)₃, which predominantly forms the 2-tetrad antiparallel basket conformation in dilute potassium-based solution (Figure S3).⁴⁴ Moreover, the recent CD, DSC, and AUC³⁹ studies support our observation of the hybrid-2 as the other coexisting conformation.

Our proposed model is reminiscent of the originally proposed beads-on-a-string model by Vorlickova et al. (2005) and Yu et al. (2006) in which G4 units independently fold into consecutive G4 structures (Figure 4).^{31,32} Our observation that 2-tetrad antiparallel basket and hybrid-2 conformations in both 3'-terminal and internal G4 units are formed not only in dilute potassium-based solution but also in a cellular crowded environment suggests that the proposed G-overhang conformational preferences will also be maintained *in vivo*. Recently, Tan and co-workers observed kinetically (μ s-to-ms time-scale folding) and thermodynamically (hour time-scale folding) favored antiparallel-type and parallel propeller G4 formation, respectively, in potassium solution supplemented with 40% PEG 200.⁴⁸ To clarify whether extended telomeric sequences could potentially adopt the thermodynamically favored parallel G4 arrangement in the presence of native-like molecular crowding, the double G4 unit construct with 3'-TT was folded into the parallel G4 arrangement in dilute potassium-based solution supplemented with 50% ethanol, lyophilized, and subsequently dissolved into *X. laevis* egg extract (Figure S8). Analysis of 1D NMR spectra revealed that the parallel G4 arrangement of the double G4 unit construct was not the thermodynamically favored conformation in egg extract. Instead, comparison of the spectral pattern of egg extract with the 1D NMR spectrum of the lyophilized parallel double G4 unit construct dissolved in dilute potassium-based solution suggests that G4 units adopted hybrid-2 and 2-tetrad antiparallel basket conformations (Figure S8). Altogether, our data suggest that physiologically relevant folding of individual G4 units in the telomeric G-overhang is both independent of nucleotide flanking and insensitive to cellular molecular crowding. Of note, our data also show that physiologically relevant conformations of G4 units in the context of the telomeric G-overhang are formed in dilute potassium-based solution, but not in dilute sodium solution and water-depleted potassium-based solutions, such as in the presence of PEG.

While the biological role of the individual conformations adopted by the G4 units remains to be addressed, the identification of hybrid-2 and 2-tetrad antiparallel basket structures as plausible G4 conformations provides a basis for rational design of selective G4-stabilizing ligands. In cancer

cells, the stabilization of G4 structures in G-overhangs with small molecular weight ligands leads to inhibition of telomerase activity, which marks G4 structures as attractive targets for anticancer therapy. Although targeting of other conformations, such as the parallel propeller G4 structure, was shown to yield desirable effects, i.e., inhibition of telomerase activity, targeting of physiologically relevant conformations of G4 units in the G-overhang could prove to be more effective, thus allowing decreased drug doses along with lowering both unwanted side effects and cost of treatment. In addition, stabilization of one of the two G4 motifs might result in phenotypes with intriguing opportunities to elucidate the biological role of these motifs *in vivo*.

CONCLUSION

To our best knowledge, this is the first study providing high-resolution insight into the conformational properties of telomeric sequences capable of forming monomeric, dimeric, and trimeric G4 units *in/ex vivo* and *in vitro*. Besides providing insight into G-overhang architecture, our data directly identified the 2-tetrad antiparallel basket structure, next to the previously implicated hybrid-2, as a physiologically relevant target for rational design of selective G4 ligands. Moreover, the native-like molecular crowding condition used here failed to convert these structures into the parallel propeller topology, suggesting that telomeric G4-DNA remains hydrated *in vivo*.

EXPERIMENTAL PROCEDURES

DNA Preparation. Non-labeled and site-specific ¹⁵N-guanine-labeled oligonucleotides were purchased from Microsynth (Switzerland). Fully 2 mg of ¹⁵N/¹³C-labeled d(AG₃(TTAGGG)₃) was purchased from Silantes Munich. Oligonucleotides were dissolved in G4-folding buffer (10 mM tetrabutylammonium phosphate, 1 mM EDTA, pH 7.0) as reported,¹³ and then dialyzed and concentrated via a 3 kDa vivaspin centrifugal ultrafiltration device to 2–5 mM stock solutions. *In vitro* experiments were carried out in three main buffers: (i) dilute potassium-based solution, mimicking the intracellular salt environment of *X. laevis* oocytes (25 mM HEPES (pH 7.5), 10.5 mM NaCl, 110 mM KCl, 130 nM CaCl₂, 1 mM MgCl₂, 10% D₂O); (ii) in the presence of this dilute potassium-based solution supplemented with 40% PEG 200 (Fluka); or (iii) in a dilute sodium-based solution (25 mM HEPES (pH 7.5), 100 mM NaCl, 130 nM CaCl₂, 1 mM MgCl₂, 10% D₂O). *In vitro* steady-state fluorescence, CD, and NMR samples were prepared, unless otherwise indicated, by mixing non-folded G4-DNA constructs with respective buffers as indicated, followed by a quick annealing step (95 °C for 10 min, cooling at room temperature). *Ex vivo* NMR and steady-state fluorescence samples were produced, unless otherwise indicated, via mixing and incubation of indicated amounts of non-folded G4-DNA with egg extract at constant 18 °C temperature for 1 h before measurement.

***X. laevis* Egg Extract Preparation.** Cytoplasmic extracts were prepared as previously described.²⁰

CD Spectroscopy. CD spectra were collected from 315 to 210 nm on a Jasco J-810 spectrometer using 1-nm bandwidth. The temperature was controlled using a digitized water bath integrated with the instrument. The molar ellipticity, $[\theta]$, was calculated as follows: $[\theta] = 10^6 \times [\theta]_{\text{obs}} \times C^{-1} \times l^{-1}$, where $[\theta]_{\text{obs}}$ is the ellipticity (mdeg), C is the oligonucleotide molar concentration, and l is the optical path length of the cell (cm). Cells with 0.1 path length and oligonucleotide concentrations of 5×10^{-5} M for single G4 and 2.5×10^{-5} M for double G4 unit constructs were used. Scan rates of 50 nm min⁻¹ were used to acquire the data. The spectra were signal-averaged over six scans, baseline-corrected by subtracting a buffer spectrum, and smoothed by mean-value averaging. The prepared non-labeled oligo samples were also used for 1D NMR measurements. Melting curves were recorded by monitoring ellipticity at 290 nm from 4 to 100 °C at

a rate of 1 °C/min. Ellipticity data points were further fitted (sigmoid via Origin6.1) and normalized.

In-Cell and Ex Vivo NMR Spectroscopy. NMR experiments were performed using Bruker 700, 800, 900, and 950 spectrometers equipped with cryogenic triple-resonance probes at 18 °C. The imino protons were observed using the 11-echo pulse sequence⁴⁹ for non-labeled oligonucleotides and sfHMQC pulse sequence⁵⁰ to detect ¹⁵N-labeled oligonucleotides, with the proton excitation maximum adjusted to the center of the Hoogsteen imino region (ca. 11.5 ppm). NMR capillary tubes (3 nm) were used for *in vitro* measurements with an active sample volume of 150 μL. *In vitro* NMR samples were prepared by mixing 50–200 μM ¹⁵N-labeled telomeric G4-DNA construct with the indicated buffer, following sample annealing at 95 °C for 10 min and subsequent cooling at room temperature. In addition, d(TAG₃-(TTAGGG)₃), d(TTAG₃(TTAGGG)₃TT), and d(G₃(TTAGGG)₃T) construct concentrations in the range of 3–5 mM were used to acquire the reference ¹⁵N-edited 1D ¹H and 2D ¹H–¹⁵N sfHMQC spectra of the hybrid-1, hybrid-2, and 2-tetrad antiparallel basket structures. Shaped sample tubes (Bruker Biospin AG) were used for *ex vivo* NMR experiments of ¹⁵N isotopically labeled oligonucleotides using *X. laevis* egg extract. The *ex vivo* NMR samples were prepared by mixing not more than 10% of labeled oligonucleotide from prepared stock solutions with 90% (or higher) of pure extract solution to a total volume of 380 μL, yielding 50–200 μM oligonucleotide final concentration. The prepared *ex vivo* NMR samples were incubated for 1 h at 18 °C, if not indicated otherwise, prior to the measurements. The imino proton background signal arising from intrinsic nucleic acids in *X. laevis* egg extract was subtracted from spectra obtained with telomeric constructs. In-cell NMR measurements of d(AG₃-(TTAGGG)₃) using *X. laevis* oocytes were performed as recently described.^{27,51,52} Briefly, 50 nL of 2 mM ¹⁵N/¹³C-labeled d(AG₃-(TTAGGG)₃) was injected into each of the roughly 200 oocytes (corresponding after dilution to roughly 100 μM d(AG₃(TTAGGG)₃) intracellular concentration) for the preparation of one in-cell NMR sample. To ensure oocyte viability after injection, 10 randomly picked oocytes were matured to eggs via the addition of 1 μM progesterone. Here, we observed that injection of d(AG₃(TTAGGG)₃) above 120 μM per oocyte resulted in apoptosis, while below this concentration oocytes matured to eggs.

2-Aminopurine (Ex Vivo) Steady-State Fluorescence Spectroscopy. G4 loop conformations have been studied via steady-state fluorescence spectroscopy using 2-aminopurine as probe.¹⁸ Fluorometric experiments were performed at 18 °C in a 1.5 × 1.5 mm quartz cuvette using a JASCO FP-6500 spectrofluorometer. 2-AP modified oligodeoxynucleotides were excited at 305 nm (3 nm bandwidth), and emission (10 nm bandwidth) spectra were measured at 1 nm intervals from 320 to 420 nm. Sample volumes were adjusted to 30 μL, and 1 μM d(AG₃(TTAGGG)₃) or 2 μM d(G₃(TTAGGG)₃TTAG₃-(TTAGGG)₃TT) was used to acquire *in vitro* spectra; 200 and 400 μM concentrations were needed to detect 2-AP-modified oligonucleotides folded within *X. laevis* egg extract to counteract extract autofluorescence (Figure S2).

■ ASSOCIATED CONTENT

■ Supporting Information

Table S1, imino proton and nitrogen chemical shift values of all telomeric G4-DNA constructs used for structural interpretations; Table S2, predicted distance information of thymidine C7–C7 methyl carbon atoms for all the currently known wild-type telomeric G4-DNA structures from vertebrates; Figures S1 and S2, 1D and 2D NMR and steady-state fluorescence spectra of the d(AG₃(TTAGGG)₃) sequence; Figures S3–S8, 1D and 2D NMR spectra, CD spectra, CD-related melting profiles, and steady-state fluorescence spectra of various model telomeric G4-DNA sequences and extended telomeric sequences. This material is available free of charge via the Internet at <http://pubs.acs.org>.

■ AUTHOR INFORMATION

Corresponding Author

vdoetsch@em.uni-frankfurt.de; rhaensel.bpc@gmail.com

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank Prof. Dr. Ernst Bamberg for support in performing microinjection of *X. laevis* oocytes and oocyte-to-egg maturation. This work was supported from the Fond der Chemischen Industrie, the Center for Biomolecular Magnetic Resonance, the Cluster of Excellence Frankfurt (Macromolecular Complexes), and the EU (Bio-NMR). L.T. was supported by the project “CEITEC” (CZ.1.05/1.1.00/02.0068) from the European Regional Development Fund and by the Maria Curie Reintegration grant.

■ REFERENCES

- (1) O’Sullivan, R. J.; Karlseder, J. *Nat. Rev. Mol. Cell. Biol.* **2010**, *11*, 171.
- (2) Verdun, R. E.; Karlseder, J. *Nature* **2007**, *447*, 924.
- (3) Williamson, J. R.; Raghuraman, M. K.; Cech, T. R. *Cell* **1989**, *59*, 871.
- (4) Schaffitzel, C.; Berger, L.; Postberg, J.; Hanes, J.; Lipps, H. J.; Pluckthun, A. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 8572.
- (5) Paeschke, K.; Simonsson, T.; Postberg, J.; Rhodes, D.; Lipps, H. J. *Nat. Struct. Mol. Biol.* **2005**, *12*, 847.
- (6) Neidle, S. *FEBS J.* **2010**, *277*, 1118.
- (7) Balasubramanian, S.; Hurley, L. H.; Neidle, S. *Nat. Rev. Drug Discov.* **2011**, *10*, 261.
- (8) Rodriguez, R.; Miller, K. M.; Forment, J. V.; Bradshaw, C. R.; Nikan, M.; Britton, S.; Oelschlaegel, T.; Xhemalce, B.; Balasubramanian, S.; Jackson, S. P. *Nat. Chem. Biol.* **2012**, *8*, 301.
- (9) Siddiqui-Jain, A.; Grand, C. L.; Bearss, D. J.; Hurley, L. H. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 11593.
- (10) Di Antonio, M.; Biffi, G.; Mariani, A.; Raiber, E. A.; Rodriguez, R.; Balasubramanian, S. *Angew. Chem., Int. Ed.* **2012**, *51*, 11073.
- (11) Muller, S.; Sanders, D. A.; Di Antonio, M.; Matsis, S.; Riou, J. F.; Rodriguez, R.; Balasubramanian, S. *Org. Biomol. Chem.* **2012**, *10*, 6537.
- (12) Ambrus, A.; Chen, D.; Dai, J.; Bialis, T.; Jones, R. A.; Yang, D. *Nucleic Acids Res.* **2006**, *34*, 2723.
- (13) Gray, R. D.; Chaires, J. B. *Nucleic Acids Res.* **2008**, *36*, 4191.
- (14) Phan, A. T. *FEBS J.* **2010**, *277*, 1107.
- (15) Yue, D. J.; Lim, K. W.; Phan, A. T. *J. Am. Chem. Soc.* **2011**, *133*, 11462.
- (16) Wang, Y.; Patel, D. J. *Structure* **1993**, *1*, 263.
- (17) Parkinson, G. N.; Lee, M. P.; Neidle, S. *Nature* **2002**, *417*, 876.
- (18) Li, J.; Correia, J. J.; Wang, L.; Trent, J. O.; Chaires, J. B. *Nucleic Acids Res.* **2005**, *33*, 4649.
- (19) Renciuik, D.; Kejnovska, I.; Skolakova, P.; Bednarova, K.; Motlova, J.; Vorlickova, M. *Nucleic Acids Res.* **2009**, *37*, 6625.
- (20) Hänsel, R.; Löhr, F.; Foldynova-Trantirkova, S.; Bamberg, E.; Trantirek, L.; Dötsch, V. *Nucleic Acids Res.* **2011**, *39*, 5768.
- (21) Heddi, B.; Phan, A. T. *J. Am. Chem. Soc.* **2011**, *133*, 9824.
- (22) Yu, H.; Gu, X.; Nakano, S.; Miyoshi, D.; Sugimoto, N. *J. Am. Chem. Soc.* **2012**, *134*, 20060.
- (23) Miyoshi, D.; Nakao, A.; Sugimoto, N. *Biochemistry* **2002**, *41*, 15017.
- (24) Xue, Y.; Kan, Z. Y.; Wang, Q.; Yao, Y.; Liu, J.; Hao, Y. H.; Tan, Z. *J. Am. Chem. Soc.* **2007**, *129*, 11185.
- (25) Singh, V.; Azarkh, M.; Exner, T. E.; Hartig, J. S.; Drescher, M. *Angew. Chem., Int. Ed.* **2009**, *48*, 9728.
- (26) Azarkh, M.; Singh, V.; Okle, O.; Dietrich, D. R.; Hartig, J. S.; Drescher, M. *ChemPhysChem* **2012**, *13*, 1444.
- (27) Hänsel, R.; Foldynova-Trantirkova, S.; Löhr, F.; Buck, J.; Bongartz, E.; Bamberg, E.; Schwalbe, H.; Dötsch, V.; Trantirek, L. *J. Am. Chem. Soc.* **2009**, *131*, 15761.

- (28) Miller, M. C.; Buscaglia, R.; Chaires, J. B.; Lane, A. N.; Trent, J. O. *J. Am. Chem. Soc.* **2010**, *132*, 17105.
- (29) Wang, Z. F.; Chang, T. C. *Nucleic Acids Res.* **2012**, *40*, 8711.
- (30) Dai, J.; Carver, M.; Yang, D. *Biochimie* **2008**, *90*, 1172.
- (31) Vorlickova, M.; Chladkova, J.; Kejnovska, I.; Fialova, M.; Kypr, J. *Nucleic Acids Res.* **2005**, *33*, 5851.
- (32) Yu, H. Q.; Miyoshi, D.; Sugimoto, N. *J. Am. Chem. Soc.* **2006**, *128*, 15461.
- (33) Pedroso, I. M.; Duarte, L. F.; Yanez, G.; Burkewitz, K.; Fletcher, T. M. *Biopolymers* **2007**, *87*, 74.
- (34) Xu, Y.; Ishizuka, T.; Kurabayashi, K.; Komiyama, M. *Angew. Chem., Int. Ed.* **2009**, *48*, 7833.
- (35) Bauer, L.; Tluckova, K.; Tohova, P.; Viglasky, V. *Biochemistry* **2011**, *50*, 7484.
- (36) Dai, J.; Carver, M.; PUNCHIHewa, C.; Jones, R. A.; Yang, D. *Nucleic Acids Res.* **2007**, *35*, 4927.
- (37) Xu, L.; Feng, S.; Zhou, X. *Chem. Commun.* **2011**, *47*, 3517.
- (38) Petraccone, L.; Trent, J. O.; Chaires, J. B. *J. Am. Chem. Soc.* **2008**, *130*, 16530.
- (39) Petraccone, L.; Spink, C.; Trent, J. O.; Garbett, N. C.; Mekmaysy, C. S.; Giancola, C.; Chaires, J. B. *J. Am. Chem. Soc.* **2011**, *133*, 20951.
- (40) Singh, V.; Azarkh, M.; Drescher, M.; Hartig, J. S. *Chem. Commun.* **2012**, *48*, 8258.
- (41) Selenko, P.; Serber, Z.; Gadea, B.; Ruderman, J.; Wagner, G. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 11904.
- (42) Selenko, P.; Frueh, D. P.; Elsaesser, S. J.; Haas, W.; Gygi, S. P.; Wagner, G. *Nat. Struct. Mol. Biol.* **2008**, *15*, 321.
- (43) Luu, K. N.; Phan, A. T.; Kuryavyi, V.; Lacroix, L.; Patel, D. J. *J. Am. Chem. Soc.* **2006**, *128*, 9963.
- (44) Lim, K. W.; Amrane, S.; Bouaziz, S.; Xu, W.; Mu, Y.; Patel, D. J.; Luu, K. N.; Phan, A. T. *J. Am. Chem. Soc.* **2009**, *131*, 4301.
- (45) Tran, P. L.; Mergny, J. L.; Alberti, P. *Nucleic Acids Res.* **2011**, *39*, 3282.
- (46) Lane, A. N. *Biochimie* **2012**, *94*, 277.
- (47) Krstic, I.; Hänsel, R.; Romainczyk, O.; Engels, J. W.; Dötsch, V.; Prisner, T. F. *Angew. Chem., Int. Ed.* **2011**, *50*, 5070.
- (48) Xue, Y.; Liu, J. Q.; Zheng, K. W.; Kan, Z. Y.; Hao, Y. H.; Tan, Z. *Angew. Chem., Int. Ed.* **2011**, *50*, 8046.
- (49) Sklenar, V.; Bax, A. J. *Magn. Reson.* **1987**, *74*, 469.
- (50) Schanda, P.; Brutscher, B. *J. Am. Chem. Soc.* **2005**, *127*, 8014.
- (51) Serber, Z.; Selenko, P.; Hänsel, R.; Reckel, S.; Lohr, F.; Ferrell, J. E., Jr.; Wagner, G.; Dötsch, V. *Nat. Protoc.* **2006**, *1*, 2701.
- (52) Hänsel, R.; Foldynova-Trantirkova, S.; Dötsch, V.; Trantirek, L. *Top. Curr. Chem.* **2012**, DOI: 10.1007/128_2012_332.