Published in final edited form as:

Methods Mol Biol. 2017 ; 1587: 171–196. doi:10.1007/978-1-4939-6892-3_17.

Human Telomeric G-Quadruplex Structures and G-Quadruplex-Interactive Compounds

Clement Lin and Danzhou Yang

Abstract

G-quadruplexes are noncanonical secondary structures formed in DNA sequences containing consecutive runs of guanines. It has been shown that the 3' G-rich single-stranded overhangs of human telomeres can form G-quadruplex structures, and the human telomeric DNA G-quadruplexes are considered attractive targets for anticancer drugs. G-quadruplex-interactive compounds have been shown to inhibit telomerase access as well as telomere capping. Nuclear magnetic resonance (NMR) spectroscopy is a powerful method in determining the G-quadruplex structures under physiologically relevant conditions. We present the NMR and biophysical methodology used in our research group for the study of G-quadruplex structures in physiologically relevant solution and their interactions with small-molecule compounds.

Keywords

Human telomeres; G-Quadruplex structures; Structure polymorphism; G-Quadruplex- interactive compounds; Telomerase inhibitors; Anticancer drug targets

1 Introduction

1.1 G-Quadruplex Structures

DNA G-quadruplexes are noncanonical secondary structures formed in sequences containing consecutive runs of guanines that have recently emerged as attractive cancer therapeutic targets [1]. A G-quadruplex is a four-stranded DNA secondary structure that contains stacked G-tetrad planes of four guanines connected by a network of Hoogsteen hydrogen bonding (Fig. 1a). G-quadruplexes can be formed with one, two, or four G-rich DNA molecules (Fig. 1b). G-strands in a G-quadruplex structure can be parallel or antiparallel, and G-tetrad guanines can adopt *anti* or *syn* conformations around the glycosidic bonds. G-tetrad guanines from parallel G-strands adopt the same glycosidic conformation while those from antiparallel G-strands adopt the opposite glycosidic conformation (Fig. 1c). G-quadruplex formation requires monovalent cations, in particular K^+ and Na^+ , to coordinate with the eight electronegative O6 atoms of the adjacent stacked G-tetrads [2] (Fig. 1a). K^+ in general is more preferred than Na^+ by G-quadruplex as K^+ has better coordination with eight guanine O6s and lower dehydration energy [3]. The K^+ form is considered to be more biologically relevant due to the higher intracellular concentration of K^+ (~140 mM) than that of Na^+ (5–15 mM).

¹³Dialysis should be performed against at least 1000 volumes of exchange solvent to obtain efficient change of solution conditions.

Intramolecular G-quadruplexes are of intensive research interest currently due to their potential formation in single-stranded telomeres and oncogene promoter regions [1, 4–7]. Intramolecular G-quadruplex structures form quickly in solution and are found to be DNA sequence-specific, exhibiting great conformational diversity, such as in folding topologies, loop conformations, and capping structures [1]. High-field NMR spectroscopy has been shown to be a powerful method in determining the G-quadruplex structures under physiologically relevant solution conditions.

1.2 G-Quadruplexes Formed in Human Telomeres

The first biologically relevant G-quadruplex formation was observed in telomeric DNA [8]. Human telomeres consist of tandem repeats of the hexanucleotide d(TTAGGG)_n 5–25 kb in length, which terminate in a single-stranded 3′-overhang of 35–600 bases [9, 10]. Telomeric DNA is extensively associated with various telosome proteins, and the structure and stability of telomeres are closely related to cancer [4], aging [11], and genetic stability [12]. Vertebrate telomeric DNA repeats are highly conserved, a feature that has been suggested to be related to their ability to form the DNA G-quadruplex [4]. A study using antibodies targeted at parallel and antiparallel telomeric G-quadruplexes in ciliates has shown in vivo that the telomere end-binding proteins TEBPα and TEBPβ are required to control G-quadruplex formation and that TEBPβ phosphorylation is needed to resolve G-quadruplex structures during replication [13]. A number of other proteins have been found to specifically interact with telomeric G-quadruplex structures [14]. G-quadruplex formation was detected in vivo at human chromosomal ends by using the radiolabeled G-quadruplex ligand 360A [15] and the fluorescent G-quadruplex ligand BMVC [16]. More recently, the most direct evidence of the in vivo existence of G-quadruplexes was established by using a structure-specific fluorescent antibody in human cells, demonstrating G-quadruplex formation in human chromosomes including in telomeres and promoter regions [17, 18].

In normal somatic cells, each cell replication results in a 50- to 200-base loss of the telomere. After critical shortening of the telomeric DNA is reached, the cell undergoes apoptosis [19]. However, telomeres of cancer cells do not shorten on replication, due to the activation of telomerase, a reverse transcriptase that extends the telomeric sequence at the chromosome ends [20]. Telomerase is activated in 80–85% of human cancer cells [21], and has been suggested to play a key role in maintaining the malignant phenotype [22]. Formation of intramolecular G-quadruplex in human telomeric DNA induced by K⁺ ion has been shown to inhibit the telomerase activity [23].

1.3 G-Quadruplex Structures in Human Telomeres

Understanding of the human telomeric G-quadruplex structure under physiologically relevant K⁺ solution conditions (Fig. 2a) has been the subject of intense investigation. A basket-type intramolecular G-quadruplex was first shown to form in Na⁺ solution by NMR using a 22-nt human telomeric DNA 5′-AGGG(TTAGGG)₃ (wtTel22) (Fig. 2b) [24]; the same 22-nt telomeric sequence was shown to form a parallel-stranded three-tetrad G-quadruplex in the crystalline state in the presence of K⁺ [25]. Further work from our lab [26–28] and others [29–32] showed that the major G-quadruplexes formed in human telomeres in K⁺ solution are the hybrid-type intramolecular structures (Fig. 2a), distinct

from the Na⁺ solution structure or the K⁺ crystal structure (*see ref. 33 for a review*). The two different hybrid-type G-quadruplexes can be formed in human telomeric sequences (Fig. 2a), for example, wtTel26 (Fig. 2b) forms mainly the hybrid-2 structure (Fig. 2a) while Tel26 (Fig. 2b) forms mainly the hybrid-1 structure (Fig. 2a). The two hybrid-type structures are in dynamic equilibrium in K⁺ solution, with the hybrid-2 form being the major one for the extended four-G-tract human telomeric DNA (Fig. 2a) [28]. Both hybrid-type structures contain three parallel G-strands and one antiparallel G-strand, five *syn* guanines, and asymmetrical G-arrangements (Fig. 2a, left). They differ in loop arrangements, strand orientations, G-tetrad arrangements, and capping structures (Fig. 2a, right), all of which may provide specific ligand-binding sites [27, 28]. The hybrid-form structures also provide an efficient scaffold for a compact-stacking structure of multimers in the 3' single-stranded overhang of the human telomeric DNA (Fig. 2c) [26, 33]. The formation of higher-order telomeric G-quadruplex structures with hybrid-type forms has been supported by computational modeling and biophysical studies [34]. While short loop sizes of 1 and 2 nt are usually favored for the double-chain-reversal loop conformation [35, 36], the presence of a 3-nt double-chain-reversal TTA loop in the hybrid-type human telomeric G-quadruplexes allows the 5' and 3' ends to point in opposite directions (Fig. 2a, left). However, this 3-nt double-chain reversal loop may contribute to the structural flexibility and polymorphism of telomeric G-quadruplexes [33]. Despite the low energy difference between the two equilibrating hybrid structures, interconversion between the two appears to be kinetically slow, suggesting the presence of a high-energy intermediate [28]. We more recently determined the structure of a basket-type two G-tetrad G-quadruplex formed from a variant human telomeric sequence or specifically truncated telomeric sequences with no 5'-flanking residues [37]. This two-G-tetrad conformation is likely to be an intermediate form of the interconversion of different telomeric G-quadruplex conformations, which we propose to be a strand-reorientation process with two-tetrad and partially unfolded triplex structure intermediates [26, 37]. Nature may have selected for the human telomeric sequence because of its structural polymorphism. (See our reviews for more information [5, 33, 38]). Intramolecular G-quadruplexes formed by single-stranded DNA in human telomeres may be targeted by small molecules to inhibit telomerase function and telomere end-capping (*see refs. [4, 5] for reviews*) (Fig. 2c).

1.4 G-Quadruplex- Interactive Ligands

Recognition of the biological significance of DNA G-quadruplexes has intensified research and development of G-quadruplex interactive compounds. Targeting of DNA G-quadruplex secondary structures represents a new approach for cancer therapeutics [1, 7, 38–40], with the first report of targeting G-quadruplexes for inhibiting telomerase activity in 1997 [41]. Formation of diverse G-quadruplex structures offers an opportunity to design small molecules/ligands that can selectively bind different G-quadruplexes. Structure-based drug design has been playing an important role in the development of G-quadruplex-interactive compounds. In fact, G-quadruplex-interactive compounds themselves have contributed immensely to understanding the G-quadruplex as a therapeutic target. It is noteworthy that Quarfloxin, a G-quadruplex-targeting compound that originated from rational drug design and was developed by Cylene Pharmaceuticals (San Diego, CA), shows excellent *in vivo* activity in various solid tumors and reached Phase II clinical trials. (*See reviews [1, 42] for*

more information.) A number of small molecules that can specifically bind to and stabilize telomeric G-quadruplexes have been developed, and some have become prospective anticancer agents that display relatively low cytotoxicity. For example, TMPyP4, a G-quadruplex-interactive compound developed through structure-based rational drug design [43], has been shown to inhibit human telomerase in HeLa cell extracts. Telomestatin, one of the most potent G-quadruplex stabilizing compounds, has been shown to be a highly potent inhibitor of telomerase [44]. BRACO19, a trisubstituted acridine derivative, was rationally designed by considering the unique structural features of DNA G-quadruplexes [45, 46]. A number of other G-quadruplex-interactive compounds, including BMVC, RHPS4, 360A, and 12,459, have been shown to inhibit telomerase activity with various levels of efficacy. More recently, compounds with asymmetric aromatic ring systems have been designed to bind to telomeric G-quadruplexes [47, 48].

Early G-quadruplex compounds were designed based on the dogma that the optimal G-quadruplex-interactive compounds are those with large symmetric cyclic fused rings that maximize stacking interactions with external G-tetrads such as telomestatin or TMPyP4. Some of these stacking ligands also contained positively charged side-chain substituents that interact with phosphate groups in G-quadruplex grooves [49, 50]. However, those compounds with large symmetric cyclic rings, in general, do not appear to bind unimolecular/intramolecular G-quadruplexes specifically. We have recently determined the NMR solution structure of a 2:1 complex of quindoline and *c-MYC* G-quadruplex [51], and our results indicate that compounds containing a smaller asymmetric pharmacophore, in particular a crescent-shaped moiety, with appropriate functional groups are more likely to bind in a well-defined manner to a unimolecular G-quadruplex. Indeed, we have recently discovered that epiberberine (EPI), which also has an asymmetric crescent-shaped structure, can bind specifically to the human telomeric G-quadruplexes [52] (see Subheading 3).

In this chapter, we present practical steps involved in the NMR structural determination of human telomeric G-quadruplexes in K^+ solution. In addition, NMR, circular dichroism, and fluorescence- based methods to study G-quadruplex drug interactions are also presented.

2 Materials

2.1 Reagents

1. Phosphoramidites, resins, and other reagents (Glen Research, Sterling, VA) (see Note 1).
2. ^{15}N -Guanine phosphoramidites (Cambridge Isotopes Laboratories, Andover, MA).
3. D_2O and $DMSO-d_6$ solvents (Sigma-Aldrich, St. Louis, MO).
4. Phosphate buffer: 25 mM phosphate, 95 mM potassium, pH 7.0 (see Note 2).
5. TAMRA-labeled oligonucleotide (see Note 3).

¹Phosphoramidites are stored at room temperature until dissolution in acetonitrile.

²The buffer for NMR sample preparation is stored at room temperature.

2.2 NMR Sample Preparation

1. Lyophilized DNA samples are dissolved in aqueous solution containing 25 mM phosphate, 95 mM potassium, pH 7.0 (*see Note 4*).
2. For H₂O samples, DNA powder is dissolved in 0.6 ml of 90%/10% H₂O/D₂O. For D₂O samples, DNA is dissolved in 0.6 ml of 99% D₂O.
3. The final NMR samples contain 0.1–3 mM DNA with a final K⁺ concentration of 95 mM (*see Note 5*).
4. DNA oligonucleotides are annealed by heating to 95 °C for 15 min, then slowly cooling to room temperature.
5. Drug stock solutions (10–40 mM) can be prepared in H₂O or DMSO, depending on their solubility (*see Note 6*).

2.3 CD Sample Preparation

1. Lyophilized DNA samples are dissolved in aqueous solution containing 25 mM phosphate, 95 mM potassium, pH 7.0 (*see Note 7*). The final CD samples contain 10–30 μM DNA.
2. DNA oligonucleotides are annealed by heating to 95 °C for 15 min, then slowly cooling to room temperature.
3. Drug stocks (100–400 μM) can be prepared in H₂O or DMSO, depending on their solubility.
4. Alternatively, an NMR sample can be diluted to appropriate concentration using water or buffer solution.

2.4 Fluorescence Sample Preparation

1. Lyophilized DNA samples are dissolved in aqueous solution containing 25 mM phosphate, 95 mM potassium, pH 7.0 at a concentration of 0.5–1 μM DNA.
2. DNA oligonucleotides are annealed by heating to 95 °C for 15 min, then slowly cooling to room temperature.
3. Drug stock solutions (10–40 μM) should be prepared in the same aqueous solution as the DNA oligonucleotides, containing 25 mM phosphate, 95 mM potassium, pH 7.0.

³Labeled DNA is only necessary for fluorescence-based binding constant determination if the drug being investigated is fluorescent in complex with DNA but not in free form or vice versa.

⁴We used potassium salt as it is considered to be the more physiologically relevant ion. Na⁺ can also be used if needed.

⁵A low-concentration DNA sample is usually used for 1D experiments, while a high-concentration sample is needed for 2D NMR, particularly NOESY experiments. It needs to be confirmed that the same conformation of DNA is formed at all concentrations.

⁶For DMSO stock solution, deuterated DMSO-d₆ is used to minimize the solvent peak in ¹H-NMR spectra.

⁷For CD melting experiments, buffer containing 25 mM phosphate and 95 mM potassium ion may be too concentrated for G-quadruplexes with high (greater than 75 °C) melting temperatures, which cannot be accurately determined under such conditions. In such cases, the solution should be diluted up to 10x.

3 Methods

3.1 Synthesis and Purification of DNA with and Without Labeling

NMR spectroscopy is widely employed for structural studies of G-quadruplexes under physiologically relevant solution conditions. Significantly, global interaction information is well observed in G-quadruplex structures because of its globular folding, unlike long thread-like double-helical DNA structures, e.g., B-DNA, that lack global interaction information. Thus, a G-quadruplex structure can be derived from extended single-stranded DNA conformation by NMR-based structure calculation, analogous to a globular protein with global folding that can be well defined by the NMR-restrained structure calculation.

1. The DNA Oligonucleotides Are Synthesized at the 1 μ mol Scale on an Expedite™ 8909 Nucleic Acid System from Applied Biosystems (Foster City, CA) in DMT-On Mode using solid-phase β -cyanoethylphosphoramidite chemistry [53] (*see Note 8*).
2. For labeled DNA oligonucleotide synthesis, ^{15}N -guanine phosphoramidite is used for site-specific ^{15}N -labeling of individual bases in DNA sequences at a 6% low-level enrichment. This low enrichment is sufficient to perform special NMR experiments such as JR-HMQC (*see Note 9*).
3. The oligonucleotides are cleaved from the solid support by the treatment of concentrated ammonia for 15 h at 55 °C (*see Note 10*).
4. DMT-on oligonucleotides are purified by reverse-phase high performance liquid chromatography (RP-HPLC) on a C_{18} column (250 $\times$ 10.0 mm VARIAN Dynamax Microsorb 300–10) with a linear gradient of 5–45% acetonitrile over 45 min in 1% triethylammonium acetate (TEAA), pH 7.0, with a flow rate of 3 ml/min. DNA is monitored with UV at 254 nm. The DMT-on DNA is collected at 25 min and is then lyophilized.
5. DMT protection groups are removed from DNA by treating with 2 ml of 80% acetic acid for 2 h. The sample is then diluted with 1 ml of water and extracted with 10 ml of ether to remove DMT (*see Note 11*). The aqueous phase containing DNA is collected and concentrated on a Speedvac.
6. The concentrated DMT-off DNA samples are purified by reverse-phase HPLC on the C_{18} column with a linear gradient of 5–30% acetonitrile over 30 min in 1% triethylammonium acetate (TEAA), pH 7.0, with a flow rate of 3 ml/min. DMT-off DNA is collected at 10 min.

⁸DNA can also be synthesized at a larger scale, e.g., at 15 μ mol; however, the yield is anticipated to be lower for the larger scale synthesis.

⁹The level of enrichment can vary depending on resonances. An enrichment lower than 6% can be used for detecting resonances with high intensity, while a higher level of enrichment is needed for detecting weaker resonances.

¹⁰Cleavage of blocking groups from synthesized oligonucleotides with ammonium hydroxide can also be performed by incubation at 62 °C for 10 h.

¹¹Careful extraction of acetic acid-treated oligonucleotides with fresh ether is necessary to obtain the high-quality final material. Presence of cloudiness in the extraction process is an indication of unsuccessful extraction. Addition of 1 ml of high-purity water increases the aqueous phase volume, leading to better phase separation and better recovery of DNA products.

7. After HPLC purification, the DMT-off DNA samples are further processed by sequential dialysis in glass beakers sequentially with water, low salt (10 mM NaCl), high salt (150 mM NaCl), and water (*see* Notes 12–14). Dialysis against a weak basic solution (10 mM NaOH) can be used to remove persistent secondary structures. The purified and dialyzed samples are lyophilized to a homogeneous powder (*see* Note 15).

3.2 Determination of the G-Quadruplex Formation by 1D ¹H-NMR

NMR spectroscopy has a unique strength in studying DNA G-quadruplex secondary structures. The imino peaks associated with guanines in G-tetrad formation give rise to characteristic chemical shifts around 10.5–12 ppm (Fig. 2d). This chemical shift region is completely separated from imino chemical shifts from any other DNA conformations, such as duplex DNA, single-stranded DNA, or other secondary DNA structures. Therefore, the imino regions of G-quadruplex guanines provide a direct and clear monitoring system for not only the formation of a G-quadruplex structure but also its drug binding interactions.

1. As a first step in the NMR structure determination process, it is important to determine the G-quadruplex formation, sample purity, and line-widths by ¹H NMR spectra in solution. Formation of G-quadruplexes can be identified by chemical shifts of imino H1 protons in ¹H-NMR in H₂O solution. The appearance of imino peaks in the downfield region 10.0–12.5 ppm indicates their involvement in a Hoogsteen hydrogen-bonded network and is a characteristic for the formation of G-quadruplex. The number of imino resonances corresponds to the number of G-tetrad-associated guanines present in the system (*see* Note 16).
2. For samples in aqueous solution, the ¹H-NMR experiments are performed with water suppression techniques such as watergate or jump-and-return, and the relaxation delay is set to 3 s.

3.3 Sequence Design and Selection

It is important to note that the intramolecular G-quadruplex structure formed in a G-rich sequence can be complex. Not only do different sequences adopt different structures, but a given sequence can also fold into a variety of different conformations, as in the case of the human telomeric DNA sequence (*see* below). A DNA sequence that forms a major stable structure with good NMR spectral properties in solution, i.e., sharp line-widths, is necessary for successful structure determination by NMR. NMR and sequence analysis are unique in being able to identify the major stable G-quadruplex structures under physiologically relevant ion conditions (pH 7, 100–140 mM K⁺).

¹²Dialysis in glass beakers is necessary to prevent the leaching of chemicals which occurs while stirring in plastic beakers and produces contaminant peaks in NMR spectra.

¹⁴Oligonucleotides greater than 20-mer in lengths can be dialyzed in tubing with 3000–3500 MWCO; shorter oligonucleotides should be dialyzed in tubing with a MWCO of 1000.

¹⁵Lyophilization is a critical step for DNA product quality. All materials should be hard-frozen prior to lyophilization and dried completely before removal from the lyophilizer at all stages. Lyophilized oligonucleotides are stored under desiccation at –20 °C until dissolution.

¹⁶It is also important to determine the presence of multiple conformations. From the number of imino peaks arising from a particular sequence, it is possible to determine if there are multi-conformational species in solution and what their relative populations are.

1. The 22-nt human telomeric DNA wtTel22 (Fig. 2b), which has been shown to form a basket-type G-quadruplex in Na⁺ solution [24] or a parallel-stranded intramolecular G-quadruplex in the crystalline state in the presence of K⁺ [25], was previously examined by 1D ¹H NMR. The 1D ¹H NMR shows that wt Tel22 forms a mixture of multiple G-quadruplex structures in K⁺ solution [26].
2. We have screened a large number of variant four G-tract sequences containing the core wtTel22 with different flanking segments by ¹H NMR [26]. Sequences having no 3'-flanking strand or a 3'-flanking adenine, including the 26-nt sequence Tel26 (Fig. 2b), were found to form a major hybrid-1 type intramolecular G-quadruplex in K⁺ solution. The 1D ¹H NMR spectrum of the Tel26 sequence in K⁺ solution (Fig. 2d) shows a major intramolecular G-quadruplex structure with 12 resolved imino proton resonances between 10.5 and 12 ppm (*see* Note 17). The folding topology and molecular structure of the hybrid-1 structure was determined by NMR (PDB ID 2HY9) [26, 27] (Fig. 2a, right).
3. Sequences with extended flanking strands, such as the wtTel26 sequence (Fig. 2b), were found to form a major hybrid-2 type intramolecular G-quadruplex in K⁺ solution by NMR (Fig. 2a). The folding topology and molecular structure of the major G-quadruplex structure was determined to be hybrid-2 by NMR (PDB ID 2JPZ) [28] (Fig. 2a, right).

3.4 Assignment of Guanine Base Resonances Using Labeled DNA

Only sequences showing well-resolved NMR spectra are subjected to further structural investigation. Herein, we use the telomeric DNA sequence wtTel26 as an example for structure determination by NMR spectroscopy. NMR structure determination is comprised of NMR resonance assignment and NMR-restrained structure calculation. NMR resonance assignment of G-quadruplexes includes guanine base proton assignment using site-specific ¹⁵N-labeling and complete spectral assignment using 2D NMR methods. After the complete spectral assignment, the folding topology of a G-quadruplex is first determined, followed by three-dimensional structure determination using NMR-derived distance restraints in combination with MD calculations.

1. The 1D ¹H NMR spectrum of the wtTel26 sequence in K⁺ solution (Fig. 2d, top) shows a major intramolecular G-quadruplex structure with 12 resolved imino proton resonances between 10.5 and 12 ppm, indicating the formation of G-quadruplex with three G-tetrads (*see* Note 18).
2. The imino H1 and base aromatic H8 protons of 12 guanine residues can be unambiguously assigned by site-specific low enrichment (6%) using ¹⁵N-labeled

¹⁷Tel26 was found to form two well-defined G-quadruplex conformations when freshly dissolved in K⁺ solution, as indicated by two separate sets of relatively sharp guanine imino peaks (Fig. 2d). One conformation (~40%) slowly converts to the other (~60%) overnight, and the complete conversion takes about a day. This observation led to the careful examination of the native 26-nt human telomeric sequence wtTel26, (TTAGGG)₄TT (Fig. 2a).

¹⁸However, this major conformation only accounts for ~70% of the total population, with about 30% population of minor conformations shown as weak and broader resonances. Thus, it is a more challenging process to get a complete resonance assignment for wtTel26 and a large number of spectra at different conditions are needed [28].

guanine nucleoside at each guanine position in the wtTel26 sequence. The guanine imino H1 and aromatic H8 proton resonances have one-bond coupling to N1 and two-bond coupling to N7, respectively (Fig. 1a). Both H1 and H8 protons of the site-specific labeled guanine are readily detected by 1D ^{15}N -filtered JR-HMQC experiments (Fig. 2e) [28].

3. The assignment of guanine H8 protons can be further confirmed by long-range through-bond correlations with the already assigned imino H1 via $^{13}\text{C}5$, using 2D ^1H - ^{13}C HMBC experiment at natural abundance.

3.5 Determination of Equilibrating Structures by NMR

Site-specific ^{15}N -labeling of G-quadruplexes can be used to determine the presence of multiple G-quadruplex structures in solution. For example, 1D ^1H -NMR shows that wtTel24 (Fig. 2b) forms a mixture of multiple G-quadruplex structures in K^+ solution (Fig. 2d). ^{15}N -labeling of G18 of the wtTel24 sequence shows two imino peaks (Fig. 3), indicating a mixture of two conformations. The two imino peaks correspond to the two equilibrating hybrid forms [28].

3.6 Assignment of Thymine/Cytosine Base Resonances

1. Thymine residues can be identified by a strong cross-peak between H6-Me in COSY/TOCSY and NOESY spectra.
2. Cytosine residues can be identified by a strong cross-peak between H5 and H6 in COSY/TOCSY and NOESY spectra.
3. Base H6 protons of thymines or cytosines in a sequence can be unambiguously assigned by site-specific substitution for each thymine with deoxyuridine (dU) or for each cytosine with dT or dU one at a time (*see Note 19*).

3.7 Determination of G-Quadruplex Folding Topology

The characteristic imino H1 and base H8 connectivity pattern lead to the direct determination of the folding topology of a G-quadruplex.

1. The G-tetrad alignments and folding topology of a sequence are determined by the interresidual H1 and H8 NOEs in exchangeable NOESY spectra (Fig. 4a, b). In a G-tetrad plane, the imino proton H1 of each guanine is in close spatial vicinity to the H1s of the two adjacent guanines, and to the base H8 of one of the adjacent guanines due to the Hoogsteen H-bonded network (Fig. 1a). The formation of three G-tetrads, namely (G4-G12-G16-G22, top), (G5-G11-G17-G23, middle), and (G6-G10-G18-G24, bottom), can be determined for wtTel26.
2. The overall G-quadruplex alignment can be further defined based on the inter-tetrad NOE connections from residues that are positioned far apart in the DNA sequence. The strong NOE interactions of G5H1/G12H1, G11H1/G16H1, and G17H1/G22H1 (Fig. 4a) connect the top two G-tetrads and define their reversed

¹⁹Site-specific substitution for adenines with inosines (dI) can also be used for the assignment of adenine base protons.

G-arrangements. The sequential inter-tetrad NOE interactions between H1 (Fig. 4a) and H8 proton indicate the same G-conformations of the bottom two G-tetrads. From these data, the folding topology of the G-quadruplex formed in wtTel26 can be unambiguously determined (Fig. 2a).

3.8 Complete Spectral Assignment

1. After the site-specific assignment of base protons, the complete assignment of the base and sugar protons of a DNA sequence can be accomplished with 2D-NMR experiments (Fig. 4). Standard homonuclear 2D-NMR experiments are used to assign the non-exchangeable proton chemical shifts of the DNA, including DQF-COSY, TOCSY, and NOESY. The mixing times are set at 50, 100, 150, and 200 ms for NOESY and at 40 and 80 ms for TOCSY at various temperatures. The exchangeable proton chemical shifts are assigned for H₂O samples using 2D-NOESY experiments with WATERGATE or jump-return (NOE11) water suppression techniques. The relaxation delay is set to 2 s. The acquisition data points are set to 4096×512 . The 60°-shifted sine bell functions are applied to both dimensions of NOESY and TOCSY spectra. The fifth-order polynomial functions are employed for the base-line corrections.
2. Assignment of base protons H8/H6 (*see* Subheadings 3.4 and 3.5) gives rise to the assignment of H1-H1 (Fig. 4a) and H1-H8 crosspeaks (Fig. 4b), which can be used for the determination of the folding topology of a G-quadruplex (*see* Subheading 3.8).
3. The assignment of guanine H1 and H8 protons can give rise to the assignment of the base proton H2/H8 of adenine residues that stack with the G-tetrads (Fig. 4b).
4. A standard DNA sequential assignment procedure is used for the assignment of non-exchangeable protons (Fig. 4c). The assignment of the aromatic H8 can lead to the direct assignment of H1' and H2'/H2'', and then can be extended to H3', H4' and H5'/H5''.
5. Sequential NOEs can be assigned for the neighboring residues in a DNA sequence. For example, G_nH8-G_{n-1} H1'/H2'/H2'' NOEs allow the assignment of sequential guanines in the DNA (Fig. 4c).
6. The glycosidic torsion angles (*syn/anti*) of G-tetrad guanines can be determined by the intensity of the intraresidue H8-H1' NOEs in NOESY spectra with different mixing times. Observation of a weak intraresidue H8-H1' NOE at a short mixing time indicates an *anti* conformation, while a strong H8-H1' NOE indicates a *syn* conformation (Fig. 4c). In addition, a characteristic downfield shift is observed for H2'/H2'' of guanines that are in *syn* conformation. In wtTel26, there are five *syn* tetrad guanines out of a total of 12 tetrad guanines (Fig. 2a, left), clearly shown in the NOESY spectrum (Fig. 4c).
7. Based on the complete spectral assignment, critical inter- residue NOE interactions can be schematically drawn as in Fig. 4d. This schematic diagram of

NOE interactions clearly defines the overall shape of the hybrid-2 G-quadruplex formed in wtTel26 in K⁺.

3.9 ³¹P Heteronuclear NMR

1. ³¹P heteronuclear NMR experiments can be used to refine NMR structure calculation, particularly for loop residues with unusual conformations. ³¹P-¹H correlation experiments provide useful information for torsion angles: ³¹P-H5'/H5'' coupling constants define β , while ³¹P-¹H couplings give information about ϵ torsions [36]. All ³¹P NMR spectra are collected with an external standard of 85% H₃PO₄. The experiments include the 1D ¹H-decoupled ³¹P spectrum, 2D-heteronuclear ³¹P-¹H Correlation Spectroscopy (COSY), and Heteronuclear Single Quantum Correlation Spectroscopy (HSQC).

3.10 NMR- Restrained Structure Determination

1. For peak assignments and NOE volume integration (peak fit-ting function), the software Sparky (UCSF) is used [54]. Distances between non-exchangeable protons are estimated based on the NOE cross-peak volumes at 50–300 ms mixing times, with the upper and lower boundaries assigned to $\pm 20\%$ of the estimated distances. Distances between exchangeable protons are assigned with looser boundaries ($\pm 1\text{--}1.5$ Å). The distance between thymine base protons Me–H6 (2.99 Å) is used as a reference for wtTel26 (*see* Note 20). The distances involving the unresolved protons, e.g., methyl protons, are assigned using pseudo-atom notation to make use of the pseudo-atom correction automatically computed by X-PLOR [55]. Only relatively isolated peaks are used for NOE-restrained structure calculation.
2. Metric matrix distance geometry (MMDG) calculations are carried out using an X-PLOR protocol to produce and optimize 100 initial structures. An arbitrary global conformation of G-quadruplex is used as a starting model, as G-quadruplexes contain enough global folding and global interactions to derive a final NMR structure. Substructure embedding is performed to produce a family of 100 DG structures. The embedded DG structures are then subjected to simulated annealing.
3. All of the 100 conformations generated from the DGSA calculations are subjected to NOE-restrained simulated annealing refinement in X-PLOR with a distance-dependent dielectric constant. Hydrogen bonds in the G-tetrad planes are restrained with distances corresponding to ideal hydrogen bond geometry. Each individual hydrogen bond is restrained using two distance restraints (heavy atom–heavy atom and heavy atom–proton). The force constants are scaled at 30 and 100 kcal mol^{−1} Å^{−2} for distance restraints and hydrogen bond constraints. All possible intra- and inter-residue NOE-distance restraints, together with backbone dihedral torsions and hydrogen-bond restraints for the G-tetrads, are incorporated into the structure calculation.

²⁰When applicable, the cytosine base proton H5–H6 distance (2.45 Å) is normally used as a reference.

4. NOE-restrained simulated annealing refinement calculations are initiated at 300 K. The temperature is gradually increased to 1000 K in 4 ps, equilibrated at 1000 K for 20 ps, and then slowly cooled to 300 K in 10 ps. The 20 best conformations are selected based both on their minimal energy terms and number of NOE violations, and are further subjected to NOE-restrained MD calculations. The resulting averaged structure is subjected to minimization. The ten best molecules are selected based both on their minimal energy terms and number of NOE violations as representative solution structures of G-quadruplex [28].

3.11 G-Quadruplex- Interactive Compounds

NMR spectroscopy is an indispensable technique to obtain high-resolution structural information about biologically important molecules and their interactions with ligands in solution. The appropriate sizes and the globular folding of DNA G-quadruplexes offer an excellent molecular system for NMR structural study. In addition, the well-separated characteristic imino regions of G-quadruplex guanines provide a direct and unambiguous detection system for drug binding interactions. Furthermore, NMR allows a direct analysis of drug interactions with G-quadruplexes such as binding stoichiometry and binding kinetics, with direct monitoring of solution conditions such as temperatures and salt concentrations.

1. To study drug interactions of one particular G-quadruplex, the DNA sequence that forms a single major structure should be used (*see Note 21*). This sequence should have good NMR spectral properties, thus its drug interactions can be unambiguously characterized. For example, to study drug binding of the hybrid-1 telomeric G-quadruplex, the Tel26 sequence can be used; to study drug binding of the hybrid-2 telomeric G-quadruplex, the wtTel26 sequence can be used (Fig. 2b). The molecular structure of the G-quadruplex formed in the free DNA needs to be determined first and serves as the structural basis for studying its drug interactions.
2. 1D ^1H NMR titration experiments are used to monitor drug binding interactions. 1D NMR spectra of free drug and free DNA will be collected and used as references. Here we use wtTel26 as an example.
3. The purity of the drug is checked using ^1H -NMR or using analytical-HPLC. Impurities can cause additional peaks in ^1H -NMR spectra, which may overlap with the desired peaks or may even bind to wtTel26. The purity of the drug should be >95%.
4. For drug binding studies, concentrated drug stock solutions (20–40 mM) are prepared in H_2O or DMSO-d_6 (*see Note 22*).

²¹In some cases, the wild-type DNA sequence that forms multiple conformations is preferred for examining specific drug bindings. A drug that binds to one specific conformation may be able to shift the equilibrium of the multiple forms.

²²Highly concentrated drug stock solutions allow the use of small volume additions of drug solutions in the DNA-drug complex, so that the solvent effect is negligible.

5. The DNA (0.1–0.5 mM) solution containing 25 mM phosphate, 95 mM potassium, pH 7.0 is checked for G-quadruplex formation by ^1H -NMR.
6. The DNA-drug complex is prepared by step-by-step additions of a small volume of concentrated drug stock solution into DNA sample to achieve DNA-drug ratios of 1:0.1, 1:0.2, 1:0.3, 1:0.4, 1:0.5, 1:1, 1:2, and 1:3 (*see* Note 23). The complex sample is annealed if needed.
7. The change in chemical shifts of DNA-drug complex at various drug equivalences is monitored by 1D ^1H -NMR spectra. The imino proton regions of tetrad-guanines are well separated from non-exchangeable protons and thus can be used in monitoring drug binding interactions and line-width changes in titration profiles.
8. In general, the effects of a drug on a G-quadruplex are readily interpreted by NMR titration methods. A drug that binds specifically to a particular DNA sequence with sufficient affinity will produce a NMR spectrum with well-resolved peaks. In the case of a slow exchange binding regime, two of sets peaks for the free DNA and the bound DNA can both be observed at lower drug equivalence, e.g., 0.5; and signals from the free DNA disappear once all the binding sites are occupied. In the case of a medium-to-fast exchange binding regime, only one set of peaks are observed for DNA and drug. In the case of medium-to-fast exchange binding, the complex structure cannot be determined, but some information on binding sites can be obtained. Compounds that do not bind tightly/specifically are readily discerned, as they do not lead to any shifts or cause spectral line broadening.
9. If a drug-DNA complex gives a well-resolved ^1H -NMR spectrum, the binding site(s) or mode(s) can be determined. For example, a stacking compound (with either end of the G-quadruplex) will affect the chemical shifts of the guanine imino protons from the end G-tetrad with which the drug stacks, while an intercalating drug will shift the guanine imino resonances of the two G-tetrads between which the drug intercalates. For the groove-binding drugs, the protons located in the groove region, such as sugar protons and G's H8, rather than the imino protons, will be shifted upon drug binding. The shifting of base protons' chemical shifts will also be different for different binding modes.
10. Once the complex system is optimized, the detailed structure determination can be carried out by following the steps in Subheading 3.
11. For example, the binding of epiberberine (EPI, Fig. 5a) with wtTel23 was studied by 1D ^1H NMR titration experiments. The free wtTel23 (Fig. 5b) in 0.1 M K^+ solution forms a major hybrid-1 G-quadruplex conformation as indicated by well- resolved imino proton peaks, as shown previously [30]. Upon addition of EPI to the wtTel23 DNA solution, the imino proton peaks of the DNA first

²³The DNA-drug ratio less than 1:0.5 is important for drug binding studies: for a slow exchange binding regime, two sets of peaks for the free DNA and the bound DNA will be observed at the ratio of 1:0.5, while for a medium-to-fast exchange binding regime, only one set of peaks will be observed.

broaden at lower EPI equivalence (0.5 N) and then become sharper at higher EPI equivalence (Fig. 5b), indicating a medium exchange rate of EPI binding to wtTel23 on the NMR time-scale. The NMR titration data support a 1:1 binding stoichiometry of EPI with wtTel23, as no further qualitative change is visible in the imino region at the EPI equivalence higher than 1 (Fig. 5a). The observation of a relatively well-resolved imino proton peak at a 1:1 ratio suggests a rather specific EPI binding. The upfield-shifting of the DNA imino proton peaks indicates that EPI binds wtTel23 with a stacking binding mode [56].

3.12 Circular Dichroism with G-Quadruplex

Circular dichroism (CD) spectroscopy is highly sensitive to nucleic acid conformational states [57] and can be used as a complementary method to determine the formation of G-quadruplexes. CD experiments can provide information on G-quadruplex folding topology and thermal stability. CD can also be used to study small-molecule interactions with a G-quadruplex such as the effects of the small molecule on folding and thermostability of the G-quadruplex.

1. The general folding topology of a G-quadruplex can be determined based on a simple CD spectrum. Antiparallel-stranded G-quadruplexes have a CD spectrum characterized by a positive ellipticity maximum at 295 nm and a negative minimum at 265 nm, while parallel-stranded ones have a positive maximum at 264 nm and a negative minimum at 240 nm [33, 57]. Hybrid-type G-quadruplexes such as telomeric G-quadruplexes exhibit a combination of the two (*see* Notes 24 and 25).
2. The wavelength scan range used will depend on the desired experiment. For a G-quadruplex or G-quadruplex—drug complex, a typical scan window is 330–200 nm. Spectra are obtained by averaging three scans, after amplification and baseline correction to remove signal contributions from the buffer.
3. CD melting experiments to obtain T_m values are performed at the wavelength of maximum ellipticity, by applying a gradient of temperature from 25 to 93 °C with a heating rate of 0.5–4 °C/min and 1 s response time (*see* Note 26). For example, a CD melting experiment shows that epiberberine increases the T_m value of wtTel23 as indicated by the CD melting curve at 292 nm (Fig. 5c). The respective CD spectra also suggest a likely change in wtTel23 conformation as indicated by the change in spectral profile between the free wtTel23 and its EPI complex (Fig. 5d).

²⁴The appearance of even a small shoulder or peak at approx. 280–295 can indicate hybrid folding topology. For example, the hybrid type BCL2 Mid G-quadruplex shows a small peak at 295 [59] whereas the parallel c-Myc has no peak between 280 and 300 [60].

²⁵A CD spectrum can indicate hybrid folding topology, but it cannot definitively distinguish between hybrid-1 and hybrid-2 topology. NMR should be used to elucidate this ambiguity.

²⁶Hysteresis may be observed between melting and annealing curves. If observed, the temperature gradient should be decreased to improve accuracy [61].

3.13 Fluorescence- Based Binding Affinity Determination of G-Quadruplex-Interactive Compounds

Quantitative binding data is critical for understanding the small- molecule binding of G-quadruplexes. Fluorescence spectroscopy can provide binding affinity data for G-quadruplex-interactive drugs.

1. An excitation scan is run from 300 to 365 nM. The wavelength resulting in the highest intensity signal is used for subsequent experiments.
2. Using the determined excitation wavelength, an emission wavelength is determined starting from 5 nM above the excitation wavelength up to 700 nm. The wavelength of maximum intensity is used for binding constant determination.
3. The excitation and emission wavelengths of a complex may differ from that of the free DNA or free drug and should be con-firmed for each complex.
4. Depending on complex binding kinetics, some time may elapse before the drug and G-quadruplex associate. A time-based experiment is needed to determine the appropriate incubation time between titration points; 1 equivalent of DNA and drug are combined and their fluorescence is measured over 900 s (15 min). The time required for the fluorescence measurement to stabilize should be allowed for incubation between each addition of drug.
5. Once excitation and emission wavelengths and incubation time are determined, the binding curve is determined through subsequent measurements of step-by-step additions of a small volume of concentrated drug stock solution into the DNA sample to achieve DNA-drug ratios of 1:0.1, 1:0.2, 1:0.4, 1:0.6, 1:0.8, 1:1, 1:2, 1:3... and so on, until the complex is saturated (*see Note 27*).
6. Binding stoichiometry is determined by Job's plot analysis.
7. The data is fitted by software such as KaleidaGraph (Synergy Software) or MATLAB (MathWorks Inc.) according to an appropriate (i.e., 1:1, 2:1 drug:DNA) binding model. For example, in the case of 1:1 binding, data is fit to the general binding eq. [58]

$$A = A_{\min} + (A_{\max} - A_{\min}) \left(\frac{P_T + O_T + K_d - \sqrt{(P_T + O_T + K_d)^2 - (4P_T O_T)}}{2O_T} \right)$$

where $A_{\min}$ is the minimum fluorescence fractional intensity, $A_{\max}$ is the maximum fluorescence intensity, P_T is the drug concentration (x-axis), O_T is the concentration of the DNA, and K_d is the binding constant.

8. For example, fluorescence titration was used to determine the binding affinity of protoberberine alkaloids to wtTel23. The typical titration curves and the results

²⁷In the case where the drug is used as a fluorophore, unlabeled DNA is titrated into the drug solution instead [62].

were fit according to the 1:1 binding model (Fig. 5e). The good fitting to the data con-firms a 1:1 binding model. The binding constants with wtTel23 for coptisine (COP), palmatine (PAL), berberine (BER), and worenine (WOR) were determined to be $(1.4 \pm 0.5) \times 10^6$, $(4.1 \pm 0.4) \times 10^5$, $(4.1 \pm 0.4) \times 10^5$, and $(5.1 \pm 0.4) \times 10^5 \text{ M}^{-1}$, respectively. Remarkably, epiberberine has a much higher wtTel23 binding affinity with a binding constant of $(2.8 \pm 0.6) \times 10^7 \text{ M}^{-1}$ (dissociation constant $K_d = 37 \pm 8 \text{ nM}$).

Acknowledgments

This research was supported by the National Institutes of Health funding (1S10RR16659, CA122952, GM083117, and CA177585). We thank Dr. Megan Carver for proofreading the manuscript.

References

1. Yang D, Okamoto K. Structural insights into G-quadruplexes: towards new anticancer drugs. *Future Med Chem.* 2010; 2(4):619–646. [PubMed: 20563318]
2. Sen D, Gilbert W. A sodium-potassium switch in the formation of four-stranded G4-DNA. *Nature.* 1990; 344(6265):410–414. [PubMed: 2320109]
3. Hud, NV., Plavec, J. The role of cations in determining quadruplex structure and stability. In: Neidle, S., editor. *Quadruplex Nucleic Acids*. Royal Society of Chemistry, RSCPublishing; Cambridge: 2006. p. 100-130.
4. Neidle S, Parkinson G. Telomere maintenance as a target for anticancer drug discovery. *Nat Rev Drug Discov.* 2002; 1(5):383–393. [PubMed: 12120414]
5. Punchihewa, C., Yang, DZ. Therapeutic targets and drugs-G-quadruplex inhibitors. In: Hiyama, K., editor. *Telomeres and telomerase in cancer*. Humana Press; New York: 2009. p. 251-280.
6. Qin Y, Hurley LH. Structures, folding patterns, and functions of intramolecular DNA G-quadruplexes found in eukaryotic promoter regions. *Biochimie.* 2008; 90(8):1149–1171. [PubMed: 18355457]
7. Onel B, Lin C, Yang D. DNA G-quadruplex and its potential as anticancer drug target. *Sci China Chem.* 2014; 57(12):1605–1614. [PubMed: 27182219]
8. Henderson E, Hardin CC, Walk SK, Tinoco I Jr, Blackburn EH. Telomeric DNA oligonucleotides form novel intramolecular structures containing guanine-guanine base pairs. *Cell.* 1987; 51(6):899–908. [PubMed: 3690664]
9. Moyazis RK, Buckingham JM, Cram LS, Dani M, Deaven LL, Jones MD, Meyne J, Ratliff RL, Wu JR. A highly conserved repetitive DNA sequence (TTAGGG)_n, present at the telomere of human chromosomes. *Proc Natl Acad Sci U S A.* 1988; 85:6622–6626. [PubMed: 3413114]
10. Wright WE, Tesmer VM, Huffman KE, Levene SD, Shay JW. Normal human chromosomes have long G-rich telomeric overhangs at one end. *Genes Dev.* 1997; 11(21):2801–2809. [PubMed: 9353250]
11. Bodnar AG, Ouellette M, Frolkis M, Holt SE, Chiu CP, Morin GB, Harley CB, Shay JW, Lichtsteiner S, Wright WE. Extension of life-span by introduction of telomerase into normal human cells. *Science.* 1998; 279(5349):349–352. [PubMed: 9454332]
12. Sun D, Lopez C, Von Hoff DD, Hurley LH. Modulation of the catalytic activity and processivity of human telomerase from HeLa cell. *Proc Am Assoc Cancer Res.* 1998; 39:569.
13. Paeschke K, Simonsson T, Postberg J, Rhodes D, Lipps HJ. Telomere end-binding proteins control the formation of G-quadruplex DNA structures in vivo. *Nat Struct Mol Biol.* 2005; 12(10):847–854. [PubMed: 16142245]
14. Oganessian L, Bryan TM. Physiological relevance of telomeric G-quadruplex formation: a potential drug target. *Bioessays.* 2007; 29(2):155–165. [PubMed: 17226803]
15. Granotier C, Pennarun G, Riou L, Hoffschir F, Gauthier LR, De Cian A, Gomez D, Mandine E, Riou JF, Mergny JL, Mailliet P, Dutrillaux B, Boussin FD. Preferential binding of a G-quadruplex

- ligand to human chromosome ends. *Nucleic Acids Res.* 2005; 33(13):4182–4190. [PubMed: 16052031]
16. Chang CC, Kuo IC, Lin JJ, Lu YC, Chen CT, Back HT, Lou PJ, Chang TC. A novel carbazole derivative, BMVC: a potential antitumor agent and fluorescence marker of cancer cells. *Chem Biodivers.* 2004; 1(9):1377–1384. [PubMed: 17191915]
17. Biffi G, Tannahill D, McCafferty J, Balasubramanian S. Quantitative visualization of DNA G-quadruplex structures in human cells. *Nat Chem.* 2013; 5(3):182–186. DOI: 10.1038/nchem.1548 [PubMed: 23422559]
18. Chambers VS, Marsico G, Boutell JM, Di Antonio M, Smith GP, Balasubramanian S. High-throughput sequencing of DNA G-quadruplex structures in the human genome. *Nat Biotechnol.* 2015; 33(8):877–881. [PubMed: 26192317]
19. Harley CB, Futcher AB, Greider CW. Telomeres shorten during ageing of human fibroblasts. *Nature.* 1990; 345(6274):458–460. [PubMed: 2342578]
20. Greider CW, Blackburn EH. Identification of a specific telomere terminal transferase activity in tetrahymena extracts. *Cell.* 1985; 43(2 Pt 1):405–413. [PubMed: 3907856]
21. Kim NW, Piatyszek MA, Prowse KR, Harley CB, West MD, Ho PLC, Coviello GM, Wright WE, Weinrich SL, Shay JW. Specific association of human telomerase activity with immortal cells and cancer. *Science.* 1994; 266(5193):2011–2015. [PubMed: 7605428]
22. Hanahan D, Weinberg RA. The hallmarks of cancer. *Cell.* 2000; 100(1):57–70. [PubMed: 10647931]
23. Zahler AM, Williamson JR, Cech TR, Prescott DM. Inhibition of telomerase by G-quartet DNA structures. *Nature.* 1991; 350(6320):718–720. [PubMed: 2023635]
24. Wang Y, Patel DJ. Solution structure of the human telomeric repeat d[AG(3)(T(2) AG(3))₃] G-tetraplex. *Structure.* 1993; 1(4):263–282. [PubMed: 8081740]
25. Parkinson GN, Lee MPH, Neidle S. Crystal structure of parallel quadruplexes from human telomeric DNA. *Nature.* 2002; 417(6891):876–880. [PubMed: 12050675]
26. Ambrus A, Chen D, Dai JX, Bialis T, Jones RA, Yang DZ. Human telomeric sequence forms a hybrid-type intramolecular G-quadruplex structure with mixed parallel/antiparallel strands in potassium solution. *Nucleic Acids Res.* 2006; 34(9):2723–2735. [PubMed: 16714449]
27. Dai JX, Punchihewa C, Ambrus A, Chen D, Jones RA, Yang DZ. Structure of the intramolecular human telomeric G-quadruplex in potassium solution: a novel adenine triple formation. *Nucleic Acids Res.* 2007; 35(7):2440–2450. [PubMed: 17395643]
28. Dai JX, Carver M, Punchihewa C, Jones RA, Yang DZ. 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 Res.* 2007; 35(15):4927–4940. [PubMed: 17626043]
29. Xu Y, Noguchi Y, Sugiyama H. The new models of the human telomere d[AGGG(TTAGGG)₃] in K⁺ solution. *Bioorg Med Chem.* 2006; 14(16):5584–5591. [PubMed: 16682210]
30. Luu KN, Phan AT, Kuryavyi V, Lacroix L, Patel DJ. Structure of the human telomere in K⁺ solution: an intramolecular (3+1) G-quadruplex scaffold. *J Am Chem Soc.* 2006; 128(30):9963–9970. [PubMed: 16866556]
31. Phan AT, Kuryavyi V, Luu KN, Patel DJ. Structure of two intramolecular G-quadruplexes formed by natural human telomere sequences in K⁺ solution. *Nucleic Acids Res.* 2007; 35(19):6517–6525. [PubMed: 17895279]
32. Phan AT, Luu KN, Patel DJ. Different loop arrangements of intramolecular human telomeric (3+1) G-quadruplexes in K⁺ solution. *Nucleic Acids Res.* 2006; 34(19):5715–5719. [PubMed: 17040899]
33. Dai J, Carver M, Yang DZ. Polymorphism of human telomeric quadruplex structures. *Biochimie.* 2008; 90(8):1172–1183. [PubMed: 18373984]
34. Petraccone L, Spink C, Trent JO, Garbett NC, Mekmaysy CS, Giancola C, Chaires JB. Structure and stability of higher-order human telomeric quadruplexes. *J Am Chem Soc.* 2011; 133(51):20951–20961. DOI: 10.1021/ja209192a [PubMed: 22082001]

35. Ambrus A, Chen D, Dai JX, Jones RA, Yang DZ. Solution structure of the biologically relevant g-quadruplex element in the human c-MYC promoter. implications for g-quadruplex stabilization. *Biochem.* 2005; 44(6):2048–2058. [PubMed: 15697230]
36. Dai JX, Chen D, Jones RA, Hurley LH, Yang DZ. NMR solution structure of the major G-quadruplex structure formed in the human BCL2 promoter region. *Nucleic Acids Res.* 2006; 34(18):5133–5144. [PubMed: 16998187]
37. Zhang Z, Dai J, Veliath E, Jones RA, Yang DZ. Structure of a two-G-tetrad intramolecular G-quadruplex formed by a variant human telomeric sequence in K⁺ solution: insights into the interconversion of human telomeric G-quadruplex structures. *Nucleic Acids Res.* 2010; 38(3): 1009–1021. [PubMed: 19946019]
38. Chen Y, Yang DZ. Sequence, stability, and structure of G-quadruplexes and their interactions with drugs. *Curr Protoc Nucleic Acid Chem.* 2012; 50:17.15.11–17.15.17.
39. Balasubramanian S, Hurley LH, Neidle S. Targeting G-quadruplexes in gene promoters: a novel anticancer strategy? *Nat Rev Drug Discov.* 2011; 10(4):261–275. [PubMed: 21455236]
40. Neidle S. Quadruplex nucleic acids as novel therapeutic targets. *J Med Chem.* 2016; 59(13):5987–6011. [PubMed: 26840940]
41. Sun DY, Thompson B, Cathers BE, Salazar M, Kerwin SM, Trent JO, Jenkins TC, Neidle S, Hurley LH. Inhibition of human telomerase by a G-quadruplex-interactive compound. *J Med Chem.* 1997; 40(14):2113–2116. [PubMed: 9216827]
42. Brooks TA, Hurley LH. The role of supercoiling in transcriptional control of MYC and its importance in molecular therapeutics. *Nat Rev Cancer.* 2009; 9(12):849–861. [PubMed: 19907434]
43. Wheelhouse RT, Han FX, Sun D, Hurley LH. The interaction of telomerase inhibitory porphyrines with G-quadruplex DNA. *Proc Am Assoc Cancer Res.* 1998; 39:430.
44. Shinya K, Wierzbicka K, Matsuo K, Ohtani T, Yamada Y, Furihata K, Hayakawa Y, Seto H. Telomestatin, a novel telomerase inhibitor from *Streptomyces anulatus*. *J Am Chem Soc.* 2001; 123(6):1262–1263. [PubMed: 11456694]
45. Read M, Cuesta J, Basra I, Harrison J, Reszka A, Gowan S, Kelland LR, Neidle S. Rational design approaches to increase the potency of G-quadruplex-mediated telomerase inhibitors. *Clin Cancer Res.* 2001; 7(11):713.
46. Zhou G, Liu X, Li Y, Xu S, Ma C, Wu X, Cheng Y, Yu Z, Zhao G, Chen Y. Telomere targeting with a novel G-quadruplex-interactive ligand BRACO-19 induces T-loop disassembly and telomerase displacement in human glioblastoma cells. *Oncotarget.* 2016; 7(12):14925–14939. [PubMed: 26908447]
47. Zhang W-J, Ou T-M, Lu Y-J, Huang Y-Y, Wu W-B, Huang Z-S, Zhou J-L, Wong K-Y, Gu L-Q. 9-Substituted berberine derivatives as G-quadruplex stabilizing ligands in telomeric DNA. *Bioorg Med Chem.* 2007; 15(16):5493–5501. [PubMed: 17574421]
48. Bessi I, Bazzicalupi C, Richter C, Jonker HR, Saxena K, Sissi C, Chioccioli M, Bianco S, Bilia AR, Schwalbe H. Spectroscopic, molecular modeling, and NMR-spectroscopic investigation of the binding mode of the natural alkaloids berberine and sanguinarine to human telomeric G-quadruplex DNA. *ACS Chem Biol.* 2012; 7(6):1109–1119. [PubMed: 22486369]
49. Collie GW, Promontorio R, Hampel SM, Micco M, Neidle S, Parkinson GN. Structural basis for telomeric G-quadruplex targeting by naphthalene diimide ligands. *J Am Chem Soc.* 2012; 134(5): 2723–2731. DOI: 10.1021/ja2102423 [PubMed: 22280460]
50. Chung WJ, Heddi B, Hamon F, Teulade-Fichou MP, Phan AT. Solution structure of a G-quadruplex bound to the bisquinolinium compound Phen-DC(3). *Angew Chem Int Ed Engl.* 2014; 53(4):999–1002. DOI: 10.1002/anie.201308063 [PubMed: 24356977]
51. Dai JX, Carver M, Hurley LH, Yang DZ. Solution structure of a 2:1 quindoline-c-MYC G-quadruplex: insights into G-quadruplex-interactive small molecule drug design. *J Am Chem Soc.* 2011; 133(44):17673–17680. [PubMed: 21967482]
52. Zhang L, Liu H, Shao Y, Lin C, Jia H, Chen G, Yang D, Wang Y. Selective lighting up of epiberberine alkaloid fluorescence by fluorophore-switching aptamer and stoichiometric targeting of human telomeric DNA G-quadruplex multimer. *Anal Chem.* 2014; 87(1):730–737. [PubMed: 25429435]

53. Dai JX, Punchihewa C, Mistry P, Ooi AT, Yang DZ. Novel DNA Bis-intercalation by MLN944, a potent clinical bisphenazine anti-cancer drug. *J Biol Chem*. 2004; 279(50):46096. [PubMed: 15317822]
54. Goddard, TD., Kneller, DG. SPARKY 3. University of California; San Francisco: 2004.
55. Brünger, AT. Version 3.1: A system for X-ray crystallography and NMR. Yale University Press; New Haven, CT: 1993.
56. Lazzeretti P. Ring currents. *Prog Nucl Magn Reson Spectrosc*. 2000; 36(1):1–88.
57. Vorlickova M, Kejnovska I, Sagi J, Renciuik D, Bednarova K, Motlova J, Kypr J. Circular dichroism and guanine quadruplexes. *Methods*. 2012; 57(1):64–75. DOI: 10.1016/j.ymeth.2012.03.011 [PubMed: 22450044]
58. Lakowicz, JR. Principles of fluorescence spectroscopy. 3. Springer; New York: 2010.
59. Dai J, Dexheimer TS, Chen D, Carver M, Ambrus A, Jones RA, Yang DZ. An intramolecular G-quadruplex structure with mixed parallel/antiparallel G-strands formed in the human BCL-2 promoter region in solution. *J Am Chem Soc*. 2006; 128(4):1096–1098. [PubMed: 16433524]
60. Hatzakis E, Okamoto K, Yang D. Thermodynamic stability and folding kinetics of the major G-quadruplex and its loop isomers formed in the nuclease hypersensitive element in the human c-Myc promoter: effect of loops and flanking segments on the stability of parallel-stranded intramolecular G-quadruplexes. *Biochem*. 2010; 49(43):9152–9160. DOI: 10.1021/bi100946g [PubMed: 20849082]
61. Dexheimer TS, Carey SS, Zuohe S, Gokhale VM, Hu X, Murata LB, Maes EM, Weichsel A, Sun D, Meuillet EJ, Montfort WR, Hurley LH. NM23-H2 may play an indirect role in transcriptional activation of c-myc gene expression but does not cleave the nuclease hypersensitive element III1. *Mol Cancer Ther*. 2009; 8(5):1363–1377. [PubMed: 19435876]
62. Liu L, Shao Y, Peng J, Huang C, Liu H, Zhang L. Molecular rotor-based fluorescent probe for selective recognition of hybrid G-quadruplex and as a K⁺ sensor. *Anal Chem*. 2014; 86(3):1622–1631. [PubMed: 24405563]

Fig. 1.

(a) Schematic diagram of a G-tetrad containing a square planar alignment of four guanines connected by cyclic Hoogsteen hydrogen bonding between the N1, N2 and O6, N7 of guanine bases. *Curved lines* show the H1–H1 and H1–H8 connectivity pattern detectable in NOESY experiments. (b) Schematic diagrams of monomeric (intramolecular), dimeric, and tetrameric G-quadruplexes. (c) The guanines in a G-tetrad can adopt either *syn* or *anti* glycosidic conformation depending on the orientation of DNA strands

Fig. 2.

(a) *Left*: Schematic drawings of the folding topologies of the hybrid-1 and hybrid-2 intramolecular telomeric G-quadruplexes in K^+ solution. Darker boxes represent guanine in *syn* conformation and lighter boxes represent guanine in *anti* conformation. *Right*: Two hybrid-type intramolecular G-quadruplex structures coexist in human telomeres in K^+ solution. The molecular structures are determined by NMR in pH 7.0, 95 mM K^+ solution. (b) Human telomeric DNA sequences containing four G-tracts that were used for structure determination and other studies. The sequence numbering system is shown above wtTel26. The major conformations formed in each sequence are specified. (c) A schematic model of

human telomeres composed of compact-stacking multimers of DNA G-quadruplex secondary structure, and the proposed mechanism of drug-mediated inhibition of telomerase and interference of telomere capping by G-quadruplex-stabilizing compounds. **(d)** The imino regions of 1D ^1H -NMR spectra of wtTel26, Tel26, wtTel23, wtTel22, and wtTel24 in K^+ solution. The assignment of imino protons is shown for wtTel26, Tel26, and wtTel23. **(e)** The imino proton region with assignment of the 1D ^1H NMR spectrum of wtTel26 in K^+ solution (*top*), and imino proton assignments of wtTel26 using 1D ^{15}N -filtered experiments on site-specific labeled oligonucleotides. Conditions: 25 mM phosphate, 95 mM potassium ion, pH 7.0, 25 °C, 0.5–0.6 mM DNA. **(f)** The CD spectra of wtTel26 and Tel26 in 100 mM K^+ solution at 25 °C

Fig. 3.

The imino proton region of the 1D ^1H NMR spectrum of wtTel24 (Fig. 2b) and its site-specific labeling at G18. The imino peaks for hybrid-1 and hybrid-2 structures are labeled as 18-1 and 18-2 respectively

Fig. 4.

The expanded H1-H1 (a) and H1-H8/H2/H6 (b) regions of the exchangeable proton 2D-NOESY spectrum of wtTel26 in K^+ solution. (c) The expanded H8/H6-H1' region of the non-exchangeable 2D-NOESY spectrum of wtTel26 in K^+ solution. The sequential assignment pathway is shown. The H8-H1' NOEs of the guanines with *syn* conformation are labeled with residue names. Missing connectivities are labeled with asterisks. The characteristic *G(i)H8/G(i+1)H1'* NOEs for the *syn* *G(i)s* are labeled by arrows. Conditions: 15 °C, 25 mM phosphate, 95 mM potassium ion, pH 7.0, 2.5 mM DNA. (d) Schematic diagram of inter-residue NOE connectivities of wtTel26 G-quadruplex formed in K^+ solution. The guanines in *syn* conformation are represented using gray circles

Fig. 5.

(a) The chemical structure of epiberberine. (b) The 1D ^1H NMR titration profiles of wtTel23 with the small-molecule epiberberine at different ratios. Conditions: 25 $^{\circ}\text{C}$, 0.2 mM DNA, 25 mM phosphate, 95 mM potassium ion, pH 7.0. (c) CD melting curves of the wtTel23 G-quadruplex in the absence and presence of epiberberine at 4 equivalents. (d) CD spectra of the wtTel23 G-quadruplex in the absence and presence of epiberberine at 4 equivalents. Conditions: 10 mM K^+ , 10 M DNA (C) Fluorescence titration of epiberberine (EPI) and related alkaloids palmatine (PAL), coptisine (COP), worenine (WOR), berberine (BER), and sanguinarine (SAN)