Published in final edited form as:

Methods Mol Biol. 2019 ; 2035: 157–176. doi:10.1007/978-1-4939-9666-7_9.

NMR Studies of G-Quadruplex Structures and G-Quadruplex-Interactive Compounds

Clement Lin¹, Jonathan Dickerhoff¹, Danzhou Yang^{2,3,4}

¹Medicinal Chemistry and Molecular Pharmacology, College of Pharmacy, Purdue University, West Lafayette, IN, USA

²Department of Medicinal Chemistry and Molecular Pharmacology, College of Pharmacy, Purdue University, West Lafayette, IN, USA

³Purdue Center for Cancer Research, West Lafayette, IN, USA

⁴Purdue Institute for Drug Discovery, West Lafayette, IN, USA

Abstract

G-quadruplexes are noncanonical, four-stranded nucleic acid secondary structures formed in sequences containing consecutive runs of guanines. These G-quadruplex structures have been found to form in nucleic acid regions of biological significance, including human telomeres, gene promoters, and untranslated regions of mRNA. Thus, they are considered attractive therapeutic targets. Nuclear magnetic resonance (NMR) spectroscopy is a powerful method for understanding the structures of G-quadruplexes and their interactions with small molecules under physiologically relevant conditions. Here, we present the NMR methodology used in our research group for the study of DNA G-quadruplex structures in physiologically relevant solution and their ligand interactions.

Keywords

NMR spectroscopy; G-quadruplex structures; G-quadruplex-interactive compounds

1 Introduction

G-quadruplexes are noncanonical, four-stranded nucleic acid secondary structures formed in sequences containing consecutive runs of guanines, which have recently emerged as attractive cancer therapeutic targets [1]. They are built from stacked G-tetrad planes, consisting of four guanine bases connected by a square network of Hoogsteen hydrogen bonding (Fig. 1a). G-quadruplexes can be formed with one, two, or four G-rich nucleic acid molecules (Fig. 1b). Within a G-quadruplex structure, G-strands 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). Formation of G-quadruplex structures requires the presence of monovalent cations, such as K^+ and Na^+ , to coordinate with the eight electronegative O6 atoms of the adjacent stacked G-tetrads [2] (Fig. 1a). K^+ is preferred over Na^+ by G-quadruplexes in general for its better coordination with the guanine O6s and lower dehydration energy [3]. Due to their potential formation in biologically relevant regions [1, 4–7], intramolecular G-quadruplexes are a subject of significant research interest. These 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]. Recognition of the biological significance of G-quadruplexes has intensified research and development of G-quadruplex interactive compounds. Targeting of G-quadruplex secondary structures in DNA represents a new approach for cancer therapeutics [1, 7–10], with the first report of targeting G-quadruplexes for inhibiting telomerase activity in 1997 [11].

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 which originated from rational drug design and was developed by Cylene Pharmaceuticals (San Diego, CA), showed excellent in vivo activity in various solid tumors and reached Phase II clinical trials (*see reviews* [1, 12].) 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 [13], and Telomestatin, one of the most potent G-quadruplex stabilizing compounds [14], have been shown to inhibit human telomerase. Several other G-quadruplex-interactive compounds have been shown to inhibit expression of oncogenes, including MYC [15] and BCL-2 [16].

Nuclear magnetic resonance (NMR) spectroscopy is widely used for the structural study of macromolecules, such as proteins and nucleic acids, as well as their binding interactions with small molecules. NMR spectroscopy employs a strong magnetic field to analyze suitable nuclei, such as 1H , ^{13}C , or ^{15}N , with atomic resolution. The information of the chemical environment is gained and additionally through-space interactions between protons are revealed, which are fundamental for NMR-based structure calculations. The extracted distance information can then be used to guide a molecular mechanics calculation to the final quadruplex conformation. Unlike X-ray crystallography, NMR spectroscopic study benefits from its ability to utilize the G-quadruplex under physiologically relevant solution conditions. This enables the determination of high-resolution structures in the solution state and can provide information for kinetics and dynamics studies, all of which are critical for understanding G-quadruplexes as a drug target. In this chapter we present practical steps involved in the NMR structural determination of DNA G-quadruplexes in K^+ solution. In addition, we describe NMR methods used in our group to examine G-quadruplex–ligand interactions.

2 Materials

2.1 Reagents

1. Phosphoramidites, resins, and other reagents (Glen Research, Sterling, VA) (*see Note 1*).
2. ^{15}N -Guanine labeled phosphoramidites (Cambridge Isotopes Laboratories, Andover, MA).
3. Milli-Q Water, D_2O , and $\text{DMSO}-d_6$ solvents.
4. Phosphate buffer: 25 mM phosphate, 95 mM potassium, pH 7.0 (*see Note 2*).

2.2 NMR Sample Preparation

1. Lyophilized DNA samples are dissolved in aqueous solution containing 25 mM phosphate and 95 mM potassium, pH 7.0 (*see Subheading 3.1, Notes 3 and 4*).
2. For H_2O samples, DNA powder is dissolved in 0.6 mL of 90%/10% $\text{H}_2\text{O}/\text{D}_2\text{O}$. For D_2O samples, DNA is dissolved in 0.6 mL of 99.9% D_2O .
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 D_2O or $\text{DMSO}-d_6$, depending on their solubility (*see Note 6*).

3 Methods

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.

3.1 Synthesis and Purification of DNA with and without Labeling

1. The DNA oligonucleotides are synthesized at the 1 μmol scale on an Expedite™ 8909 Nucleic Acid System from Applied Biosystems (Foster City, CA) in DMT-on mode using solid-phase β -cyanoethylphosphoramidite chemistry [17] (*see Note 7*).
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 1D JR-HMQC (*see Note 8*).

3. The oligonucleotides are cleaved from the solid support by the treatment of concentrated ammonia for 15 h at 55 °C (*see Note 9*).
4. DMT-on oligonucleotides are purified by reverse-phase high performance liquid chromatography (RP-HPLC) on a C₁₈ column (250 × 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 10*). 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 C₁₈ 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.
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 11–13*). 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 14*).

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. 2c). 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/D₂O 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 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 15*).
2. For samples in aqueous solution, the ¹H-NMR experiments are performed with water suppression techniques such as Watergate or Jump-and-Return.

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⁺). In this section, we present the determination of the G-quadruplex formed in human telomeric DNA.

1. The 22-nt human telomeric DNA wtTel22 (Fig. 2b), which has been shown to form a basket-type G-quadruplex in Na⁺ solution [18] or a parallel-stranded intramolecular G-quadruplex in the crystalline state in the presence of K⁺ [19], was previously examined by 1D ¹H NMR. The 1D ¹H NMR shows that wtTel22 forms a mixture of multiple G-quadruplex structures in K⁺ solution [20].
2. We screened a large number of variant four G-tract sequences containing the core wtTel22 with different flanking segments by ¹H NMR (Fig. 2b) [20].
3. 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. 2c) shows a major intramolecular G-quadruplex structure with twelve resolved imino proton resonances between 10.5–12 ppm (*see Note 16*). The folding topology and molecular structure of the hybrid-1 structure was determined by NMR (PDB ID 2HY9) [20, 21] (Fig. 2a left).
4. 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) [21] (Fig. 2a center).

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 comprises 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. 2c top) shows a major intramolecular G-quadruplex structure with twelve resolved imino proton resonances between 10.5 and 12 ppm, indicating the formation of a G-quadruplex with three G-tetrads (*see Note 17*).
2. The imino H1 and base aromatic H8 protons of 12 guanine residues can be unambiguously assigned by site-specific low enrichment (6%) using ^{15}N -labeled 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. 2d) [21].
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 [22].

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. 2c). ^{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 [21].

3.6 Assignment of Thymine/Cytosine/Adenine Base Resonances

Several approaches are possible for assigning base resonances of thymine, cytosine, and adenine. It is generally possible to assign these bases using COSY, TOCSY, and NOESY spectra. ^1H - ^{13}C -HSQC spectra of the aromatic region can facilitate the assignment of all base protons, as they are clustered at specific regions of ^1H - ^{13}C -HSQC spectra depending on their type and glycosidic torsion angle (Fig. 4) [23].

1. Thymine residues can be identified by a strong cross-peak between H6-Me in TOCSY and NOESY spectra (*see Note 18*).
2. Cytosine residues can be identified by a strong cross-peak between H5 and H6 in COSY/TOCSY and NOESY spectra. Also, the H6/H5 and H6/Me cross-peaks are the only signals of aromatic protons in TOCSY or COSY 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*).
4. The adenine H2, which can be difficult to distinguish from H8, can be easily assigned based on their cross-peak with the strongly downfield shifted C2 (~155 ppm) compared to the other base protons in ^1H - ^{13}C -HSQC spectra.

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 inter-residual H1 and H8 NOEs in exchangeable NOESY spectra (Fig. 5a, 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. 5a) connect the top two G-tetrads and define their reversed G-arrangements. The sequential intertetrad NOE interactions between H1 (Fig. 5a) 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. 5). 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 (see **Note 20**). 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. 5a) and H1–H8 cross-peaks (Fig. 5b), 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 may give rise to the assignment of the base proton H2/H8 of adenine residues that stack with the G-tetrads (Fig. 5b).
4. A standard DNA sequential assignment procedure is used for the assignment of non-exchangeable protons (Fig. 5c). Within a base, the assignment of the

aromatic H6/H8 protons 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. 5c) (see **Note 21**).
6. The glycosidic torsion angles (*syn/anti*) of the G-tetrad guanines can be determined by the C6/C8 ^{13}C chemical shift in the HSQC. Since they are strongly dependent on the glycosidic angle, a downfield shift up to 4 ppm is observed for *syn* oriented bases (Fig. 4) [23]. Additionally, strong intraresidue H8-H1' NOE cross-peaks are found even at short mixing times together with rather weak intraresidue H8-H2'/H2'' contacts. 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. 5c). 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 twelve tetrad guanines (Fig. 2a center), clearly shown in the NOESY spectrum (Fig. 5c).
7. Based on the complete spectral assignment, critical interresidue NOE interactions can be schematically drawn as in Fig. 5d. This schematic diagram of NOE interactions clearly defines the overall shape of the hybrid-2 G-quadruplex formed in wtTel26 in K^+ .

3.9 NMR-Restrained Structure Determination

1. For peak assignments and NOE volume integration (peak fitting function), the software Sparky (UCSF) is used [24].
2. Distances restraints between non-exchangeable protons are calculated based on the NOE cross-peak volumes at 50–300 ms mixing times, based on the equation:

$$R_{ij} = R_{ref} \left(\frac{V_{ref}}{V_{ij}} \right)^{1/6}$$

where r_{ij} is the interproton distance of the two protons i and j , r_{ref} is a known reference distance between two protons, v_{ij} is the NOE cross-peak volume corresponding to protons i and j , and v_{ref} is the volume of the reference NOE peak. The upper and lower boundaries assigned to $\pm 20\%$ of the estimated distances (see **Note 22**). Distances between exchangeable protons are assigned with looser boundaries ($\pm 30\%$). The cytosine base proton H5–H6 distance (2.45 Å) is used as a reference (see **Note 23**). 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 [25].

3. 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.
4. 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.
5. 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 10 best molecules are selected based both on their minimal energy terms and number of NOE violations as representative solution structures of G-quadruplex [21].

3.10 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 24*). This sequence should have good NMR spectral properties, thus its drug interactions can be unambiguously characterized. For example, to study drug binding of the c-Myc promoter G-quadruplex, the Myc14/23 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 Myc14/23 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 25*).
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, 1:3, and 1:4 (*see Note 26*). 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 sets of 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 (*see Note 27*).
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, typically causing an upfield shift. 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 quindoline (EPI, Fig. 6a) with Myc 14/23 was studied by 1D ^1H NMR titration experiments. The free Myc 14/23 (Fig. 6b) in 0.1 M K^+ solution forms a major parallel G-quadruplex conformation as indicated by well-resolved imino proton peaks (Fig. 6c), as shown previously [26]. Upon addition of quindoline to the Myc 14/23 DNA solution, the imino proton peaks of the DNA first broaden at lower drug equivalence (0.5 N) and then become sharper at higher equivalence (Fig. 6c), indicating a medium exchange rate of quindoline binding to Myc 14/23 on the NMR time-scale. The NMR titration data support a 2:1 binding stoichiometry of quindoline with Myc 14/23, as no further qualitative change is visible in the imino region at the EPI equivalence higher than 2 (Fig. 6a). The observation of a relatively well-resolved imino proton peaks at a 2:1 ratio suggests a rather specific binding of quindoline. The upfield-shifting of the DNA imino proton peaks indicates that quindoline binds Myc 14/23 with a stacking binding mode [27].

4 Notes

1. Phosphoramidites are stored at room temperature until dissolution in acetonitrile.
2. The buffer for NMR sample preparation is stored at room temperature.
3. Some G-quadruplex forming sequences are prone to aggregation under high salt conditions. In such cases, this buffer may be diluted.
4. We used potassium salt as it is considered to be the more physiologically relevant ion. Na^+ can also be used if needed.
5. A 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.
6. For DMSO stock solution, deuterated DMSO- d_6 is used to minimize the solvent peak in ^1H -NMR spectra.
7. DNA can also be synthesized at a larger scale, e.g., at 15 μmol ; however, the yield is anticipated to be lower for the larger scale synthesis.
8. The level of enrichment can vary depending on resonances. An enrichment level lower than 6% can be used for detecting resonances with high intensity, while a higher level of enrichment is needed for detecting weaker resonances.
9. Cleavage of blocking groups from synthesized oligonucleotides with ammonium hydroxide can also be performed by incubation at 62 $^\circ\text{C}$ for 10 h.

10. Careful 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.
11. Dialysis 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.
12. Dialysis should be performed against at least 1000 volumes of exchange solvent in order to obtain efficient change of solution conditions.
13. Oligonucleotides 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.
14. Lyophilization 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.
15. It 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.
16. Tel26 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. 2c). One conformation ($\sim 40\%$) slowly converts to the other ($\sim 60\%$) 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, $(\text{TTAGGG})_4\text{TT}$ (Fig. 2a).
17. However, this major conformation only accounts for $\sim 70\%$ 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 [21].
18. It is possible to mistake the thymine H6-HMe cross-peak for the $\text{H6-H2'}/\text{H2''}$ cross-peaks in the NOESY spectra in some systems.
19. Site-specific substitution for adenines or guanine with inosines (dI) can also be used for the assignment of their respective base protons.
20. Setting acquisition data points are to 4096×512 is sufficient with fold back spectra. (We use a spectral width of $18 \text{ ppm} \times 9 \text{ ppm}$). For non-folded spectra, the resulting resolution may be poor without additional points (i.e., 4096×1024 points).

21. Sequential connectivity is often interrupted by various structural features of a G-quadruplex. Guanine H8 protons in *syn* bases exhibit reversed connectivity compared to standard b-DNA. Bases in strand-reversal loops may not show sequential NOE connectivity with guanines involved in G-tetrad formation, especially in the case of short (1–2 nt) strand-reversal loops.
22. It may be difficult to obtain accurate distance values from significantly overlapped peaks (i.e., 3 or more) for NOE-restrained structure calculation. In such cases, restraints from these peaks may be discarded or introduced during late structural refinement stages.
23. If the G-quadruplex sequence does not have cytosine bases, the thymine base protons Me-H6 (2.99 Å) can also be used as a reference.
24. In 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.
25. Highly 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.
26. The 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.
27. Line broadening may be observed for ligands in intermediate exchange. In such cases, peak resolution would recover upon saturation of the binding site with excess ligand.

Acknowledgments

This research was supported by the National Institutes of Health (R01CA122952 (DY), R01CA177585 (DY), and P30CA023168 (Purdue Center for Cancer Research)).

References

1. Yang DZ, Okamoto K (2010) Structural insights into G-quadruplexes: towards new anticancer drugs. *Future Med Chem* 2 (4):619–646 [PubMed: 20563318]
2. Sen D, Gilbert W (1990) A sodium-potassium switch in the formation of four-stranded G4-DNA. *Nature* 344(6265):410–414 [PubMed: 2320109]
3. Hud NV, Plavec J (2006) The role of cations in determining quadruplex structure and stability In: Neidle S (ed) *Quadruplex nucleic acids*. Royal Society of Chemistry, RSC Publishing, Cambridge, pp 100–130
4. Neidle S, Parkinson G (2002) Telomere maintenance as a target for anticancer drug discovery. *Nat Rev Drug Discov* 1(5):383–393 [PubMed: 12120414]
5. Puchihiwewa C, Yang DZ (2009) Therapeutic targets and drugs-G-quadruplex inhibitors In: Hiyama K (ed) *Telomeres and telomerase in cancer*. Springer, NJ, USA, pp 251–280
6. Qin Y, Hurley LH (2008) Structures, folding patterns, and functions of intramolecular DNA G-quadruplexes found in eukaryotic promoter regions. *Biochimie* 90(8):1149–1171. 10.1016/j.biochi.2008.02.020 [PubMed: 18355457]

7. Onel B, Lin C, Yang D (2014) DNA G-quadruplex and its potential as anticancer drug target. *SciChinaChem* 57(12):1605–1614
8. Balasubramanian S, Hurley LH, Neidle S (2011) Targeting G-quadruplexes in gene promoters: a novel anticancer strategy? *Nat Rev Drug Discov* 10(4):261–275 [PubMed: 21455236]
9. Neidle S (2016) Quadruplex nucleic acids as novel therapeutic targets. *J Med Chem* 59 (13):5987–6011 [PubMed: 26840940]
10. Chen Y, Yang DZ (2012) Sequence, stability, and structure of G-quadruplexes and their interactions with drugs. *Curr Protoc Nucl Acid Chem* 50:17.15.11–17.15.17
11. Sun DY, Thompson B, Cathers BE, Salazar M, Kerwin SM, Trent JO, Jenkins TC, Neidle S, Hurley LH (1997) Inhibition of human telomerase by a G-quadruplex-interactive compound. *J Med Chem* 40(14):2113–2116 [PubMed: 9216827]
12. Brooks TA, Hurley LH (2009) The role of supercoiling in transcriptional control of MYC and its importance in molecular therapeutics. *Nat Rev Cancer* 9(12):849–861 [PubMed: 19907434]
13. Wheelhouse RT, Han FX, Sun D, Hurley LH (1998) The interaction of telomerase inhibitory porphyrines with G-quadruplex DNA. *Proc Am Assoc Cancer Res* 39:430
14. Shin-ya K, Wierzbica K, Matsuo K, Ohtani T, Yamada Y, Furihata K, Hayakawa Y, Seto H (2001) Telomestatin, a novel telomerase inhibitor from *Streptomyces anulatus*. *J Am Chem Soc* 123(6):1262–1263 [PubMed: 11456694]
15. Local A, Zhang H, Benbatoul KD, Folger P, Sheng X, Tsai C- Y, Howell SB, Rice WG (2018) APTO-253 stabilizes G-quadruplex DNA, inhibits MYC expression and induces DNA damage in acute myeloid leukemia cells. *Mol Cancer Ther* 17(6):1177–1186. 10.1158/1535-7163.MCT-17-1209 [PubMed: 29626127]
16. Gunaratnam M, Collie GW, Reszka AP, Todd AK, Parkinson GN, Neidle S (2018) A naphthalene diimide G-quadruplex ligand inhibits cell growth and down-regulates BCL-2 expression in an imatinib-resistant gastrointestinal cancer cell line. *Bioorg Med Chem* 26 (11):2958–2964 [PubMed: 29724653]
17. Dai JX, Punchihewa C, Mistry P, Ooi AT, Yang DZ (2004) Novel dNa Bis-intercalation by MLN944, a potent clinical bisphenazine anticancer drug. *J Biol Chem* 279(50):46096 [PubMed: 15317822]
18. Wang Y, Patel DJ (1993) Solution structure of the human Telomeric repeat D[AG(3)(T(2)AG(3))3] G-Tetraplex. *Structure* 1(4):263–282 [PubMed: 8081740]
19. Parkinson GN, Lee MPH, Neidle S (2002) Crystal structure of parallel quadruplexes from human telomeric DNA. *Nature* 417(6891):876–880 [PubMed: 12050675]
20. Ambrus A, Chen D, Dai JX, Bialis T, Jones RA, Yang DZ (2006) Human telomeric sequence forms a hybrid-type intramolecular G-quadruplex structure with mixed parallel/antiparallel strands in potassium solution. *Nucleic Acids Res* 34(9):2723–2735 [PubMed: 16714449]
21. Dai JX, Punchihewa C, Ambrus A, Chen D, Jones RA, Yang DZ (2007) Structure of the intramolecular human telomeric G-quadruplex in potassium solution: a novel adenine triple formation. *Nucleic Acids Res* 35(7):2440–2450 [PubMed: 17395643]
22. Tu A (2000) Long-range imino proton-¹³C J-couplings and the through-bond correlation of imino and non-exchangeable protons in unlabeled DNA. *J Biomol NMR* 16 (2):175–178 [PubMed: 10723997]
23. Greene KL, Wang Y, Live D (1995) Influence of the glycosidic torsion angle on ¹³C and ¹⁵N shifts in guanosine nucleotides: investigations of G-tetrad models with alternating syn and anti bases. *J Biomol NMR* 5(4):333–338 [PubMed: 7647551]
24. Goddard TD, Kneller DG (2004). University of California, San Francisco
25. Brünger AT (1993) Version 3.1: a system for X-ray crystallography and NMR. Yale University Press, New Haven, CT, USA Version 31: A system for X-ray crystallography and NMR Yale University Press, New Haven, CT, USA
26. Luu KN, Phan AT, Kuryavii V, Lacroix L, Patel DJ (2006) Structure of the human telomere in K⁺ solution: an intramolecular (3+1) G-quadruplex scaffold. *J Am Chem Soc* 128(30):9963–9970 [PubMed: 16866556]
27. Lazzarotti P (2000) Ring currents. *Prog Nucl Magn Reson Spectrosc* 36(1):1–88

28. Ambrus A, Chen D, Dai JX, Jones RA, Yang DZ (2005) Solution structure of the biologically relevant g-quadruplex element in the human c-MYC promoter. Implications for g-quadruplex stabilization. *Biochemist* 44(6):2048–2058
29. Maruši M, Šket P, Bauer L, Viglasky V, Plavec J (2012) Solution-state structure of an intramolecular G-quadruplex with propeller, diagonal and edgewise loops. *Nucleic Acids Res* 40(14):6946–6956 [PubMed: 22532609]

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) Schematic drawings of the folding topologies of the hybrid-1 (left) and hybrid-2 (center) intramolecular telomeric G-quadruplexes and parallel (right) c-Myc promoter G-quadruplex in K^+ solution. Darker boxes represent guanine in *syn* conformation and lighter boxes represent guanine in *anti* conformation. The molecular structures are determined by NMR in pH 7.0, 95 mM K^+ solution. (b) 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) 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. **(d)** 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

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.

Representative aromatic sections from ^1H - ^{13}C HSQC spectra showing the aromatic region for a (a) parallel G4 from the Myc promoter sequence as well as for a (b) human telomeric and an (c) artificial hybrid topology [26, 28, 29]. The base H6-C6/H8-C8 cross-peaks are clustered in specific regions and anti/syn guanines within the core can be easily distinguished

Fig. 5.

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

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Fig. 6. (a) The chemical structure of quindoline. (b) Schematic drawing of the parallel-stranded Myc 14/23 G-quadruplex. (c) The 1D ^1H NMR titration profiles of Myc 14/23 with the small molecule quindoline at different ratios. Conditions: 25 °C, 0.2 mM DNA, 25 mM phosphate, 95 mM potassium ion, pH 7.0