HHS Public Access

Author manuscript

J Am Chem Soc. Author manuscript; available in PMC 2015 December 29.

Published in final edited form as:

J Am Chem Soc. 2005 December 14; 127(49): 17277–17285. doi:10.1021/ja0543090.

(3 + 1) Assembly of Three Human Telomeric Repeats into an Asymmetric Dimeric G-Quadruplex

Na Zhang, **Anh Tuaân Phan***, and **Dinshaw J. Patel**

Contribution from the Structural Biology Program, Memorial Sloan-Kettering Cancer Center, New York, New York 10021

Abstract

We present an NMR study on the structure of a DNA fragment of the human telomere containing three guanine-tracts, d(GGGTTAGGGTTAGGGT). This sequence forms in Na^+solution a unique asymmetric dimeric quadruplex, in which the G-tetrad core involves all three G-tracts of one strand and only the last 3′-end G-tract of the other strand. We show that a three-repeat human telomeric sequence can also associate with a single-repeat human telomeric sequence into a structure with the same topology that we name (3 + 1) quadruplex assembly. In this G-quadruplex assembly, there are one syn·syn·syn·anti and two anti· anti·anti·syn G-tetrads, two edgewise loops, three G-tracts oriented in one direction and the fourth oriented in the opposite direction. We discuss the possible implications of the new folding topology for understanding the structure of telomeric DNA, including t-loop formation, and for targeting G-quadruplexes in the telomeres.

Introduction

The terminal regions of eukaryotic chromosomes called telomeres are essential for stable chromosome maintenance.[1] Telomeric DNA contains highly repeated G-rich sequences that consists of $(GGGTTA)_n$ in human, $(GGGGTT)_n$ in *Tetrahymena*,and $(GGGGTTTT)_n$ in *Oxytricha*.[2] In the presence of monova-lent cations, such as Na^+ or K^+, G-rich sequences can form G-quadruplex structures built from the stacking of multiple planar (G·G·G·G) tetrads.[3] These structures have recently become an attractive anticancer target.[4]

The structures of G-quadruplexes reported for various DNA G-rich sequences reveal a polymorphism regarding (i) the syn/anti distribution of guanine glycosidic bonds within each tetrad, (ii) the orientation of G-tracts in the G-tetrad core, and (iii) the type of loops that connect them.[3] The stoichiometry of G-quadruplexes depends on the concentration and the sequence of DNA fragments studied. Thus, oligonucleotides containing one, two, or four G-tracts can form tetrameric, dimeric, or monomeric G-quadruplexes, respectively.[3]

*phantuan@mskcc.org.

Supporting Information Available: Table S1, a list of site-specific analogue-substituted sequences used for resonance assignments; Table S2, a list of three-repeat and single-repeat human telomeric sequences used for strand mixing experiments; Figure S1, the NOESY spectrum of *htel3I* at 50 ms; Figure S2, models of (3 + 1) G-quadruplexes in the human telomere. This material is available free of charge via the Internet at http://pubs.acs.org.

Sequences containing one human telomeric repeat d(TAGGGT) or d(TTAGGG) have been shown to form in K^+ solution a tetrameric parallel-stranded G-quadruplex, with all guanines adopting anti conformations.[5] The two-repeat human telomeric $d(TAGGGT)_2$ sequence has been reported to form a dimeric parallel-stranded G-quadruplex in a K^+-containing crystal[6] and to interconvert between this G-quadruplex and another dimeric antiparallel-stranded G-quadruplex in K^+ solution.[7] While in the first structure all guanines are anti and loops are double-chain-reversal, in the second structure guanines around each tetrad are anti·syn·anti·syn and loops are edgewise. Crystal structure of the nearly four-repeat human telomeric $d(AG_3(TTAGGG)_3)$ sequence in the presence of K^+ represents a monomeric (intramolecular) parallel-stranded G-quadruplex with anti guanines and double-chain-reversal loops.[6] In contrast, the same sequence forms in Na^+ solution a different intramolecular G-quadruplex,[8] where guanines around each tetrad are syn·syn·anti·anti, loops are successively edgewise-diagonal-edgewise and each G-tract has both a parallel and an antiparallel adjacent strands (Figure 1b). Interestingly, the *Tetrahymena* telomeric sequences, which differ from the human counterparts only by one G-for-A replacement in each repeat, form totally different G-quadruplex topologies in Na^+ solution.[9,10] In particular, the four-repeat *Tetrahymena* $d(T_2G_4)_4$ sequence forms an intramolecular G-quadruplex topology,[10] which contains three G-tetrads adopting either syn·syn·syn·anti or anti·anti·anti·syn alignments; the successive loops are edgewise, edgewise, and double-chain-reversal; three strands are oriented in one direction and the fourth in the opposite direction (Figure 1c).

Here we present the first structure of a G-quadruplex formed by three human telomeric repeats in Na^+ solution. This is a unique asymmetric dimeric quadruplex, in which the G-tetrad core involves all three G-tracts of one strand and only a single G-tract of the other strand (Figure 1d). Furthermore, we have shown that a three-repeat human telomeric sequence can associate with a single-repeat human telomeric sequence into a qua-druplex with the same topology (Figure 1e). This new folding topology may have important implications for understanding the structure of telomeric DNA and targeting G-quadruplexes in the telomeres.

Results

NMR Spectra of Three-Repeat Human Telomeric Sequences

The proton NMR spectrum of the three-repeat human telomeric d(GGGTTAGGGTTAGGGTTA) sequence (*htel3*) in Na^+ solution (Figure 2a) indicated formation of G-quadruplexes. However, its quality was not suitable for more detailed structural analysis. A cleaner NMR spectrum (Figure 2b) was obtained for the d(GGGTTAGGGTTAGGGT) sequence (*htel3*Δ), which lacks two 3′-end residues. Furthermore, the I2 analogue (*htel3I*), d(GIGTTAGGGTTAGGGT), gave well-resolved spectra with spectral characteristics similar to those of *htel3*Δ (Figure 2c). The similarity between NOE patterns of *htel3*Δ and *htel3I* (data not shown) suggested that these sequences adopt the same overall folding topology. Eleven guanine imino protons at 10-12 ppm and one inosine imino proton at 13.4 ppm for *htel3I* (Figure 2c) exchange slowly with water and are characteristic of a G-quadruplex structure.

Sequence-Specific and Strand-Specific Resonance Assignments of *htel3I*

The number of major nonexchangeable protons observed for *htel3I* indicated the presence of two different strand conformations in solution. The NOEs observed between these protons established that these two strands belong to the same structure, suggesting formation of an asymmetric dimer. Most of the guanine imino and H8 protons were unam-biguously assigned to their positions in the sequence by the site-specific low-enrichment approach[11] using samples that were 2.5% ^{15}N-labeled (Figure 3a,b). The resonance assignments for A, T, and G residues were completed (or confirmed) by using A-to-^{8Br}A, T-to-U, and G-to-^{8Br}G substituted sequences (Table S1 of the Supporting Information). Spectral assignments were also assisted by through-bond (COSY and TOCSY) and through-space (NOESY) correlations between protons.

The strand-specific assignments of nonexchangeable protons were obtained by using H8/6-H1′NOE sequential connectivities (Figure 4a) and other NOEs between base and sugar protons (not shown). This procedure established resonances belonging to each strand and sorted all resonances into two groups (black and red, Figure 4). The imino-H8 through-bond correlation spectra (Figure 3c) independently confirmed the assignments of guanine imino and H8 protons and also provided the strand-specific assignments for imino protons. Among 12 sharp guanine and inosine imino protons, nine belong to one strand (black) and three (G13, G14, and G15) belong the other strand (red, underlined). Thus, the ^{15}N-filtered spectrum of sequences that are specifically labeled at G13, G14, and G15 shows two intense peaks, while the ^{15}N-filtered spectrum of sequences that are specifically labeled at other guanines shows only one intense peak (Figure 3a).

htel3I Forms Asymmetric Dimeric G-Quadruplex

The G-tetrad arrangements can be defined on the basis of the specific imino-H8 NOE pattern within a G-tetrad (Figure 1a). Examination of NOE connectivities between imino and H8 protons of *htel3I* (Figure 5a) revealed the formation of three G-tetrads, each consists of three guanines on one strand (black) and one guanine on the other strand (red, underlined): (G13·G13·G1·G9), (I2·G14·G14·G8), and (G3·G15·G15·G7).

The folding topology of *htel3I* (Figure 1d) was subsequently drawn from G-tetrad arrangements. Three G-tracts are oriented in one direction and the fourth is oriented in the opposite direction. The (I2·G14·G14·G8) and (G3·G15·G15·G7) tetrads adopt anti·anti·anti·syn alignments with the same donor-to-acceptor hydrogen-bond directionality (clockwise when looking from the bottom), whereas the third G-tetrad, (G13·G13·G1·G9), adopts syn·syn·syn·anti alignment with the anticlockwise hydrogen-bond directionality. The syn glycosidic conformation of G1, G7, G8, G13, and G13 is supported by the observation of strong H8-H1′NOEs at short mixing times for these residues (Figure S1 of the Supporting Information); the anti glycosidic conformation of other guanines are consistent with weaker intraresidue H8-H1′NOEs. The topology presented in Figure 1d is also consistent with NOE cross-peaks observed in other spectral regions such as imino-imino, aromatic-aromatic, and aromatic-sugar proton regions (data not shown). The T4-T5-A6 and T10-T11-A12 fragments that connect G-tracts of the black strands form edgewise loops on the top and the bottom of the G-tetrad core, respectively (Figure 1d). The Watson-Crick base-pair between

T11 and A12 is supported by the observation of the imino proton of T11 at 13.7 ppm and NOE between this proton and the H2 proton of A12 (Figure 5a).

Guanines from the first two 5′-end G-tracts of the red strand do not participate in the G-tetrad core formation, consistent with their imino protons not being observed (Figures 3a and 5a) and with the fact that resonances from these G-tracts are less well dispersed and give weaker NOE connectivities compared to those from the black strand and from the last 3′-end G-tract of the red strand (Figure 4a).

The imino proton spectrum of the *htel3I* quadruplex after dissolving in D_2O (Figure 6b) indicated that the guanine and inosine imino protons of the central G-tetrad (I2, G14, G14, and G8) are among the most protected peaks, supporting the folding topology identified above (Figure 1d). Their exchange time, as measured at 30 °C by NMR, is about 5 min. The imino protons of G9 and G13 also exchange slowly with water (Figure 6b), indicating the protection of the (G13·G13·G1·G9) tetrad by the A12·T11 base pair and other bases at the bottom of the quadruplex.

In summary, the black strand forms a compact structure in which each guanine and inosine is engaged in hydrogen bonds within the G-tetrad core, while only the last 3′-end G-tract of the red strand participates in the G-tetrad core formation. This topology is further supported by the observation of the heterodimeric complex between a three-repeat and a single-repeat human telomeric sequences (see below). It seems that both ends (top and bottom) of the G-tetrad core contribute to the stability of the structure, and in particular, the interaction at the top end may allow the black strand to specifically recognize the 3′-end G-tract rather than other two G-tracts of the red strand.

Association between Three-Repeat and Single-Repeat Human Telomeric DNA Sequences into Asymmetric Dimeric G-Quadruplex

Thus, in the previous section we have established that the three-repeat black strand associates with only the last 3′-end repeat of the red strand in the G-tetrad core of the *htel3I* dimeric quadruplex. To find out whether the same black strand structure can associate with a single-repeat human telomeric sequence in the same G-quadruplex topology, we added increasing amounts of *htel3I* to the solution of the d(TAGGGU) single-repeat human telomeric sequence,[12] designated *htel1U*. Figure 7a-c shows the proton NMR spectra of *htel1U* in the absence and presence of different concentrations of *htel3I*. Under our experimental conditions, the *htel1U* sequence alone is essentially unstructured except for a small fraction, presumably forming a G-quadruplex. The peaks characteristic for the major unstructured and minor G-quadruplex forms are marked with red and green arrows, respectively (Figure 7a). In the presence of 20% of *htel3I*, while the intensity of peaks for the *htel1U* unstructured form decreased, a new set of peaks appeared (blue stars, Figure 7b). This new set of peaks, which includes guanine imino protons at 10-12 ppm, is similar (but not identical) to that of the *htel3I* quadruplex (Figure 7b and d), suggesting formation of a new G-quadruplex. In the presence of 100% of *htel3I*, peaks for the unstructured (red arrows) as well as the “self“ G-quadruplex (green arrows) of *htel1U* almost disappeared, while the intensity of peaks for the new G-quadruplex increased (Figure 7c). We conclude that *htel3I* associates with *htel1U* to form a new G-quadruplex. The NOESY spectra of the

htel3I/htel1U quadruplex (Figures 4 and 5) indicated formation of the same G-tetrad core (Figure 1e) as that of the *htel3I* dimeric quadruplex (Figure 1d). Residues of *htel1U* in the *htel3I/htel1U* quadruplex were numbered from 17 to 22 (Figure 1e). We refer to the *htel3I/htel1U* quadruplex as a (3 + 1) quadruplex[13] assembly, because each G-tetrad contains three guanines from one strand and one guanine from the other strand. It should be noted that compared to the NMR spectra of the *htel3I* quadruplex the spectra of the *htel3I/htel1U* quadruplex showed the same number guanine and inosine imino protons, but fewer nonexchangeable protons (Figures 4, 5, and 7), consistent with the lack of an overhang tail in the *htel3I/htel1U* quadruplex (Figure 1d,e). In particular, there were only six methyl protons (at 1-2 ppm) for the *htel3I/htel1U* quadruplex instead of 10 for the *htel3I* quadruplex (Figure 7).

The observation of only the *htel3I/htel1U* dimeric quadruplex for the equimolar mixture of the two strands indicates that this dimeric quadruplex (Figure 1e) is more stable than the *htel3I* dimeric quadruplex (Figure 1d). This finding is consistent with the results of hydrogen exchange experiments, which showed that the imino proton exchange time of the central G-tetrad is almost 2 times longer for the *htel3I*/*htel1U* quadruplex than for the *htel3I* quadruplex.

A reverse titration by adding increasing amounts of *htel1U* into *htel3I* was consistent with formation of the same *htel3I/htel1U* quadruplex (data not shown). We have performed the strand-mixing experiments using different three-repeat and single-repeat human telomeric sequences (Table S2). The results also suggested formation of similar heterodimeric G-quadruplexes.

Solution Structure of *htel3I/htel1U* Quadruplex

The structure of the *htel3I/htel1U* quadruplex was calculated on the basis of NMR restraints (Table 1) by using the X-PLOR program[14] (see Methods). Ten superpositioned lowest-energy intensity-refined structures of the *htel3I/htel1U* quadruplex are shown in Figure 8a. Stick, slab, and surface views of a representative refined structure of the *htel3I/htel1U* quadruplex are shown in Figure 8b-d.

The *htel1U* strand (orange backbone) intimately associates with the *htel3I* strand (gray backbone) (Figure 8a,b) mainly by participating in formation of the core of three G-tetrads: (G13·G19·G1·G9), (I2·G20·G14·G8), and (G3·G21·G15·G7). We also observed an intermolecular Watson-Crick base pair between A18 of the *htel1U* strand and T11 of the *htel3I* strand. This base pair is well-stacked under G1 and G9 (Figure 9c). The stacking patterns between three G-tetrads are shown in Figure 9a,b. There are two edgewise loops in the structure formed by the T4-T5-A6 and T10-T11-A12 fragments of the *htel3I* strand. Each loop links two adjacent antiparallel G-tracts and connects an *anti*-guanine (G3 and G9) to a *syn*-guanine (G7 and G13). However, the G3·G7 and G9·G13 hydrogen-bond directionalities are different, with donor guanines *syn*-G7 and *anti*-G9 pointing toward the groove, respectively. As a result, the T4-T5-A6 loop spans a narrow groove, while the T10-T11-A12 loop spans a wide groove (Figure 1e). The T4-T5-A6 and T10-T11-A12 loops are also different in that the former loop is close to the 3′-ends of both strands (T16 and U22), while the latter loop is close to the 5′-ends, with T11 being paired to A18. Other bases at

both ends of the G-tetrad core are less well defined due to the lack of NMR restraints, but there are possibilities for additional base pairs or triads[15a] to be formed and stabilized by stacking on G-tetrads.[15b]

The *htel3I/htel1U* quadruplex contains four grooves, which differ mainly by the orientation of strands that define them. There are one narrow, one wide, and two medium grooves (Figure 1e and 9). As shown in a surface view (Figure 8d), the shapes of the grooves are quite different and could serve as potential targets for small-molecule ligands.

Discussion

Unique (3 + 1) G-Quadruplex Assembly

We have shown that the G-rich three-repeat human telomeric sequence forms an asymmetric dimeric quadruplex in Na^+ solution. This structure is unique in that its G-tetrad core involves all three G-tracts of one strand and only the last 3′-end G-tract of the other strand (Figure 1d). Furthermore, we have established that the same topology can be obtained in the assembly between a three-repeat and a single-repeat human telomeric sequences (Figure 1e). This structure differs from previously reported telomeric G-quadruplex structures[5-10,16] in that it contains a three G-repeat telomeric sequence instead of one-, two-, or four-repeat sequences. It is also the first G-quadruplex structure that involves the heterodimeric association of two strands of such different lengths.

Interestingly, the folding topology of the G-quadruplex presented here is different from that of the one-, two- or four- repeat human telomeric sequences, even though the G-tracts are separated by the same TTA linkers. On the other hand, there is a striking similarity between the G-tetrad core of the present structure (Figure 1e) with that of the four-repeat *Tetrahymena* telomeric $d(T_2G_4)_4$ sequence (Figure 1c): three G-tracts are oriented to the same direction, while the fourth G-tract is oriented in the opposite direction; there are two anti·anti·anti· syn and one syn·syn·syn·anti G-tetrads. The folding topology of the *htel3I* strand in the *htel3I/htel1U* quadruplex (Figure 1e) is however different from that of the first three 5′-end repeats of the $d(T_2G_4)_4$ sequence (Figure 1c). When the two structures are oriented so that their cores are identical (Figure 1c, e), the 5′-end of *htel3I* starts from the left rear corner (Figure 1e), while that of the $d(T_2G_4)_4$ sequence starts from the right front corner (Figure 1c). As a result, there is a difference in the relative hydrogen-bond directionality between G-tracts of these two sequences. For instance, as regards hydrogen bonds between the first and the second G-tracts, guanines of the first G-tract of *htel3I* are donor-acceptor-acceptor (Figure 1e), while those of the $d(T_2G_4)_4$ sequence are acceptor-donor-donor (Figure 1c).

Biological and Pharmacological Implications

Human telomeres contain a 3′-end G-rich strand overhang that is about 200 nt long.[17] There is an enzyme called telomerase that is able to elongate this strand.[18] Telomerase activity is not detected in normal human somatic cells, and telomeres progressively decrease in length after each round of cell division. In contrast, telomerase is activated in 80-85% cancer cells and helps to maintain the length of telomeres in these cells.[19] It has been shown that

formation of G-quadruplexes by the telomeric G-rich strand inhibits the activity of telomerase.[20] This inhibition is most likely due to the trapping the 3′-end guanines within the G-tetrad core. Use of small molecules for stabilizing telomeric G-quadruplex has been proposed as one of the anti-telomerase strategies for the development of anticancer drugs.[4]

The new folding topology of the *htel3I* quadruplex presented here may have important implications for understanding the structure of telomeric DNA. This quadruplex structure suggests a way of how a segment of three G-tracts could bind to a remote G-tract. Interestingly, the observation of only the 3′-end G-tract (not other two G-tracts) of the *htel3I* red strand being trapped by the *htel3I* black strand suggests that the three-G-tract segment would prefer binding to the 3′-end G-tract rather than other internal ones. Such a structure could occur within the 3′-end overhang alone or even when this strand invades the preceding double-stranded part of the telomere to form a so-called t-loop[21] (Figure S2 of the Supporting Information).

The structure of the *htel3I* and *htel3I/htel1U* quadru-plexes may also have pharmacological implications. First, this G-quadruplex structure could be a very good target for antitelomerase drugs, since it represents a structure where specifically the 3′-end is captured in a G-tetrad core. Second, this structure suggests that a G-rich strand can be targeted by trapping it into a G-quadruplex structure with a linear guanine-containing molecule. This approach parallels the proposal of Armitage's laboratory to use G-rich PNA to target the G-rich DNA strand,[22] as well as another approach from our laboratory to target the grooves of the G-quadruplex by forming hydrogen bonds with the G-tetrad edges.[23] In this regard, this approach is different from most other approaches reported so far that use planar ligands to target G-quadruplexes.[4]

Methods

Sample Preparation

The unlabeled and the site-specific low-enrichment (2.5% ^{15}N-labeled) oligonucleotides were synthesized on an Applied Biosystems 392 DNA synthesizer using solid-phase β-cyanoethyl phosphoamidite chemistry and purified by HPLC as previous decribed.[8,11] The DNA samples (~7 mL) were first dialyzed against water, then against 150 mM NaCl, 10 mM sodium phosphate buffer (pH 6.8), and finally against 15 mM NaCl, 0.2 mM sodium phosphate buffer (pH 6.6) and lyophilized. Unless otherwise stated, the strand concentration of the NMR samples was typically 0.5-5 mM; the solutions contained 100 mM NaCl and 10 mM sodium phosphate (pH 6.8). The samples were heated at 90 °C for several minutes and then slowly annealed to room temperature.

NMR Spectroscopy

NMR experiments were performed on 600 MHz Varian and 800 MHz Bruker spectrometers at 30 °C, unless otherwise specified. Experiments in H_2O used the jump-and-return (JR) water suppression for detection.[24] Proton assignments were based on homonuclear NOESY, correlation spectroscopy (COSY), total correlated spectroscopy (TOCSY), site-specific low-

enrichment labeling,[11] and through-bond correlations at natural abundance.[25] Data were processed and analyzed using the FELIX program (Molecular Simulations).

Structure Calculation

The distances between nonexchangeable protons were estimated from the buildup curves of cross-peak intensities in NOESY spectra at five different mixing times (50, 100, 150, 200, and 250 ms) in D_2O and given bounds of ±20% with distances referenced relative to the uracil H5-H6 distance of 2.47 Å. Exchangeable proton restraints are based on NOESY data sets at two mixing times (60 and 200 ms) in H_2O. Cross-peaks involving exchangeable protons were classified as strong (strong intensity at 60 ms), medium (weak intensity at 60 ms), and weak (observed only at 200 ms), and distances between protons were then restrained to 3.0((0.9) Å, 4.0((1.2) Å and 6.0((1.8) Å, respectively.

The structure of the *htel3I/htel1U* quadruplex was determined by molecular dynamics simulated annealing computations driven by NOE distance, dihedral angle, and hydrogen bonding restraints using the X-PLOR program[14] as previously described.[13] The calculation protocol started with initial structures that were generated as sets of two strands (16-nt *htel3I* and 6-nt *htel1U*), randomized for all dihedral angles, and separated by space intervals of 50 Å. A total of 100 structures were embedded and optimized using the distance geometry simulated annealing protocol. The 10 best structures were selected and subjected to distance-restrained molecular dynamics calculations and then refined against the intensities of NOE cross-peaks at various mixing times between nonexchangeable protons.

Coordinates Deposition

The coordinates for the structure of the *htel3I/htel1U* quadruplex have been deposited in the Protein Data Bank (accession code 2AQY).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This research was supported by NIH grant GM34504. D.J.P. is a member of the New York Structural Biology Center supported by NIH grant GM66354.

REFERENCES

(1). Blackburn EH. FEBS Lett. 2005; 579:859–862. [PubMed: 15680963]

2. (a) Henderson, E. Telomeres. Blackburn, EH.; Greider, CW., editors. Cold Spring Harbor Laboratory Press; 1995. p. 11-34.(b) Wellinger RJ, Sen D. Eur. J. Cancer. 1997; 33:735–749. [PubMed: 9282112]

3. (a) Williamson JR. Curr. Opin. Struct. Biol. 1993; 3:357–362.(b) Patel, DJ.; Bouaziz, S.; Kettani, A.; Wang, Y. In Oxford Handbook of Nucleic Acid Structures. Neidle, S., editor. Oxford University Press; Oxford: 1999. p. 389-453.(c) Gilbert DE, Feigon J. Curr. Opin. Struct. Biol. 1999; 9:305–314. [PubMed: 10361092] (d) Keniry MA. Biopolymers. 2000-2001; 56:123–146. [PubMed: 11745109] (e) Shafer RH, Smirnov I. Biopolymers. 2000-2001; 56:209–227. [PubMed: 11745112] (f) Simonsson T. Biol. Chem. 2001; 382:621–628. [PubMed: 11405224] (g) Arthanari H, Bolton PH. Chem. Biol. 2001; 8:221–230. [PubMed: 11306347] (h) Neidle S, Parkinson GN. Curr. Opin.

Struct. Biol. 2003; 13:275–283. [PubMed: 12831878] (i) Davis JT. Angew. Chem., Int. Ed. Engl. 2004; 43:668–698. [PubMed: 14755695]

4. (a) Neidle S, Parkinson G. Nat. ReV. Drug. Des. 2002; 1:383–393.(b) Hurley LH. Nat. ReV. Cancer. 2002; 2:188–200. [PubMed: 11990855] (c) Mergny JL, Heéleène C. Nat. Med. 1998; 4:1366–1367. [PubMed: 9846570] (d) Kerwin SM. Curr. Pharm. Des. 2000; 6:441–478. [PubMed: 10788591] (e) Perry PJ, Jenkins TC. Mini. Rev. Med. Chem. 2001; 1:31–41. [PubMed: 12369989] (f) Riou JF. Curr. Med. Chem. Anti-Cancer Agents. 2004; 4:439–443. [PubMed: 15379700] (g) Phan AT, Kuryavyi V, Gaw HY, Patel DJ. Nat. Chem. Biol. 2005; 1:167–173. [PubMed: 16408022]
5. (a) Wang Y, Patel DJ. Biochemistry. 1992; 31:8112–8119. [PubMed: 1525153] (b) Patel PK, Koti AS, Hosur RV. Nucleic Acids Res. 1999; 27:3836–3843. [PubMed: 10481022] (c) Gavathiotis E, Searle MS. Org. Biomol. Chem. 2003; 1:1650–1656. [PubMed: 12926351]
6. Parkinson GN, Lee MPH, Neidle S. Nature. 2002; 417:876–880. [PubMed: 12050675]
7. Phan AT, Patel DJ. J. Am. Chem. Soc. 2003; 125:15021–15027. [PubMed: 14653736]
8. Wang Y, Patel DJ. Structure. 1993; 1:263–282. [PubMed: 8081740]
9. Phan AT, Modi YS, Patel DJ. J. Mol. Biol. 2004; 338:93–102. [PubMed: 15050825]
10. Wang Y, Patel DJ. Structure. 1994; 2:1141–1156. [PubMed: 7704525]
11. Phan AT, Patel DJ. J. Am. Chem. Soc. 2002; 124:1160–1161. [PubMed: 11841271]
12. The U-for-T substitution was used for the spectral assignment purpose and did not change the outcome of the strand-mixing experiments
13. Phan AT, Kuryavyi V, Ma JB, Faure A, Andreéola ML, Patel DJ. Proc. Natl. Acad. Sci. U.S.A. 2005; 102:634–639. [PubMed: 15637158]
14. Bruünger, AT. X-PLOR: A system for X-ray crystallography and NMR. Yale University Press; New Haven, CT: 1992.
15. (a) Kuryavyi, VV.; Jovin, TM. Biological Structure and Dynamics. Sarma, RH.; Sarma, MH., editors. Vol. 2. Adenine Press; New York; 1996. p. 91-103.(b) Kettani A, Basu G, Gorin A, Majumdar A, Skripkin E, Patel DJ. J. Mol. Biol. 2000; 301:129–146. [PubMed: 10926497]
16. (a) Smith FW, Feigon J. Nature. 1992; 356:164–168. [PubMed: 1545871] (b) Smith FW, Schultze P, Feigon J. Structure. 1995; 3:997–1008. [PubMed: 8590010] (c) Wang Y, Patel DJ. J. Mol. Biol. 1995; 251:76–94. [PubMed: 7643391] (d) Kettani A, Bouaziz S, Wang W, Jones RA, Patel DJ. Nat. Struct. Biol. 1997; 4:382–389. [PubMed: 9145109] (e) Haider S, Parkinson GN, Neidle S. J. Mol. Biol. 2002; 320:189–200. [PubMed: 12079378]
17. Makarov VL, Hirose Y, Langmore JP. Cell. 1997; 88:657–666. [PubMed: 9054505]
18. Greider CW, Blackburn EH. Cell. 1985; 43:405–413. [PubMed: 3907856]
19. Kim NW, Piatyszek MA, Prowse KR, Harley CB, West MD, Ho PL, Coviello GM, Wright WE, Weinrich SL, Shay JW. Science. 1994; 266:2011–2015. [PubMed: 7605428]
20. Zahler AM, Williamson JR, Cech TR, Prescott DM. Nature. 1991; 350:718–720. [PubMed: 2023635]
21. Griffith JD, Comeau L, Rosenfield S, Stansel RM, Bianchi A, Moss H, de Lange T. Cell. 1999; 97:503–514. [PubMed: 10338214]
22. (a) Datta B, Schmitt C, Armitage BA. J. Am. Chem. Soc. 2003; 125:4111–4118. [PubMed: 12670232] (b) Datta B, Bier ME, Roy S, Armitage BA. J. Am. Chem. Soc. 2005; 127:4199–4207. [PubMed: 15783201]
23. (a) Kettani A, Gorin A, Majumdar A, Hermann T, Skripkin E, Zhao H, Jones R, Patel DJ. J. Mol. Biol. 2000; 297:627–644. [PubMed: 10731417] (b) Zhang N, Gorin A, Majumdar A, Kettani A, Chernichenko N, Skripkin E, Patel DJ. J. Mol. Biol. 2001; 311:1063–1079. [PubMed: 11531340]
24. Plateau P, Gueéron M. J. Am. Chem. Soc. 1982; 104:7310–7311.
25. (a) Phan AT. J. Biomol. NMR. 2000; 16:175–178. [PubMed: 10723997] (b) Phan AT, Gueéron M, Leroy JL. Methods Enzymol. 2001; 338:341–371. [PubMed: 11460557]

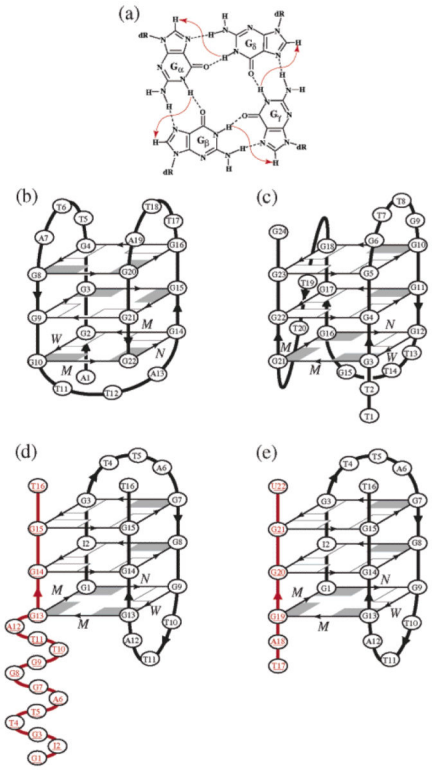

Figure 1.

(a) (G_α·G_β·Gγ·Gφ) tetrad with characteristic imino-H8 NOE pattern indicated with arrows. Schematic structure of intramolecular G-quadruplexes formed by (b) the human telomeric d($AG_3(TTAGGG)_3$)8 and (c) *Tetrahymena* telomeric d$(T_2G_4)^{10}$ sequences. Schematic structure of dimeric asymmetric G-quadruplexes formed by (d) two strands of the human telomeric d(GIGTTAGGGTTAGGGT) sequence and (e) the association of two different human telomeric d(GIGTTAGGGTTAGGGT) and d(TAGGGU) sequences. All these G-quadruplexes were observed in Na^+ solution. The backbone is shown as thick lines with the strand directionality indicated by large arrows. Hydrogen-bond directionality (donor-to-acceptor) between guanines within a G-tetrad is shown with small arrows. *syn*-Guanines are gray; *anti*-guanines are white. W, M, and N represent wide, medium, and narrow groove, respectively.

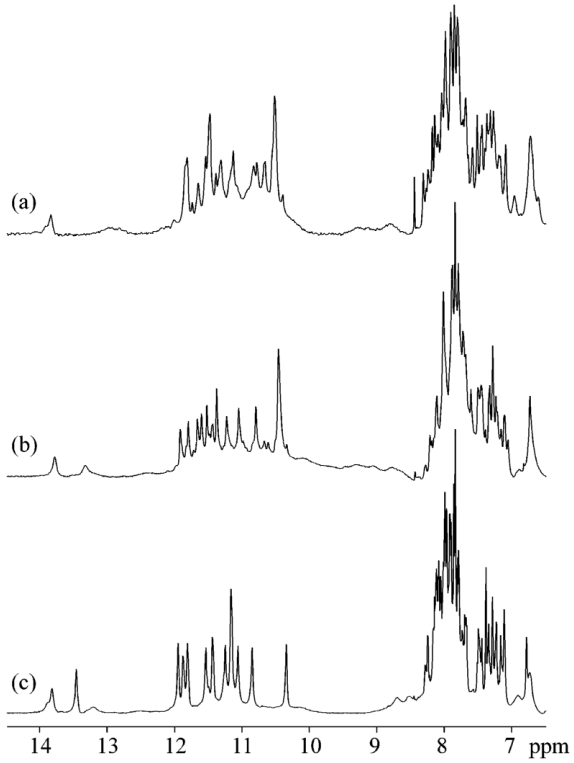

Figure 2.

The 600-MHz proton NMR spectra of human telomeric (a) d(GGGTTA)$_3$, (b) d(GGGTTAGGGTTAGGGT), and (c) d(GIGTTAGGGTTAGGGT) sequences.

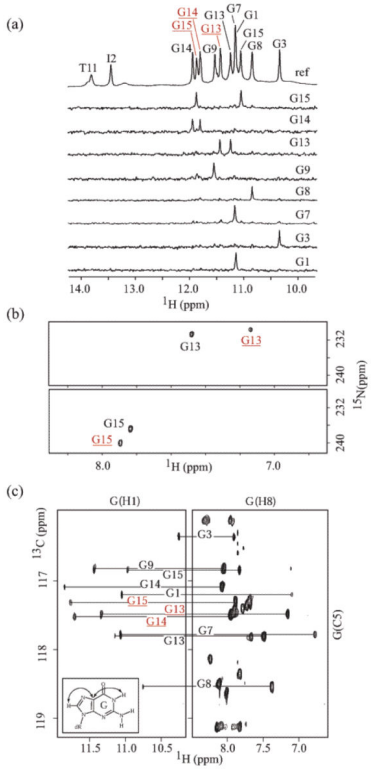

Figure 3.

(a) Imino proton spectra with assignments indicated over the reference spectrum (ref). Guanine imino protons were assigned in ^{15}N-filtered spectra of samples, 2.5% ^{15}N-labeled at the indicated positions. (b) Examples of guanine H8 proton assignments in [^{1}H-^{15}N] long range correlation experiments using 2.5% ^{15}N-labeled at G13 and G15. (c) Through-bond correlations between imino and H8 protons via ^{13}C5 at natural abundance, using long-range *J*-couplings shown in the inset.

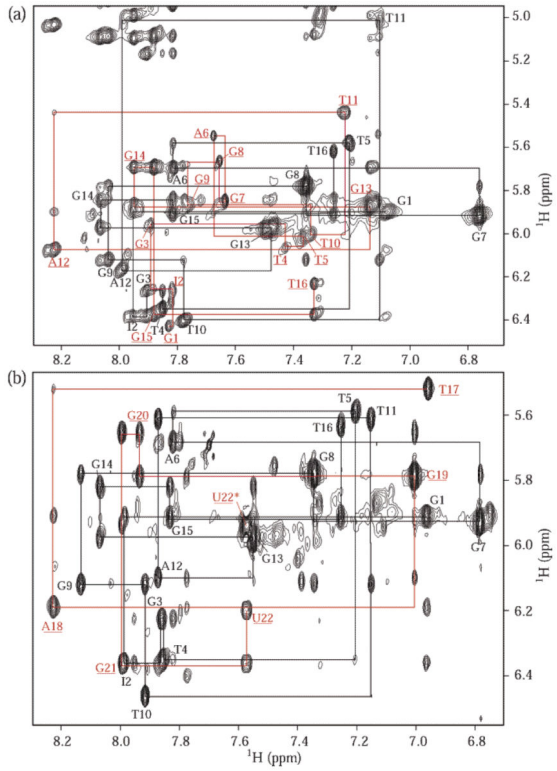

Figure 4.

The H8/6-H1′proton region of NOESY spectra (mixing time, 200 ms) of the *htel3I* dimeric quadruplex (a) and *htel3I/htel1U* dimeric quadruplex (b). The assignments and H8/6-H1′NOE sequential connectivities are shown. The peak between U22(H5) and U22(H6) is labeled with a star.

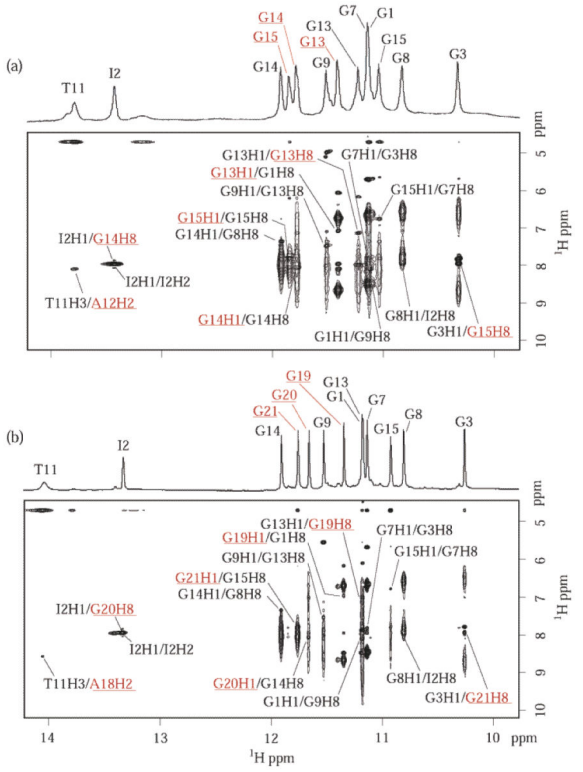

Figure 5.

The imino-H8 proton region of NOESY spectra (mixing time, 200 ms) of the *htel3I* dimeric quadruplex (a) and *htel3I/htel1U* dimeric quadruplex (b). One-dimensional imino proton projections are shown. Characteristic imino-H8 cross-peaks around G-tetrads are labeled.

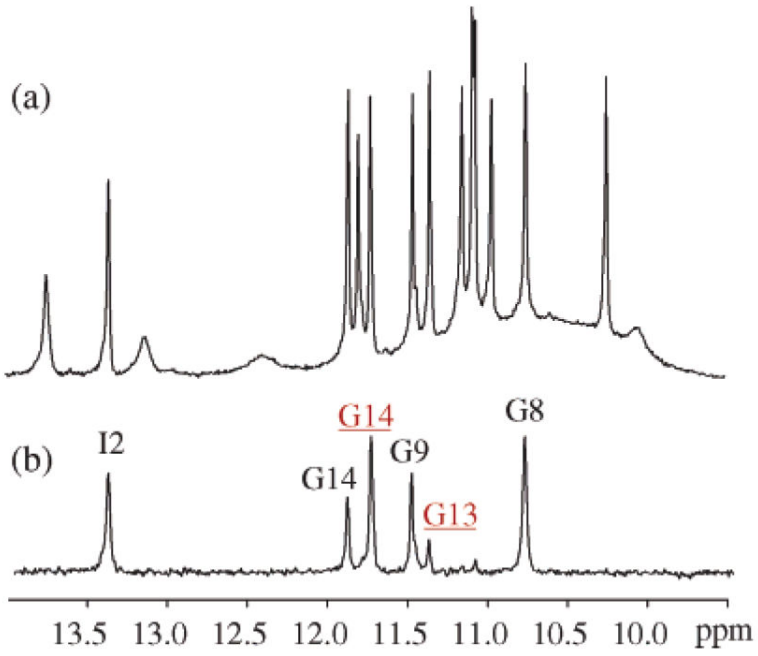

Figure 6.

Imino proton spectra of the human telomeric d(GIGTTAGGGTTAGGGT) sequence in H_2O (a) and after 5 min in D_2O (b).

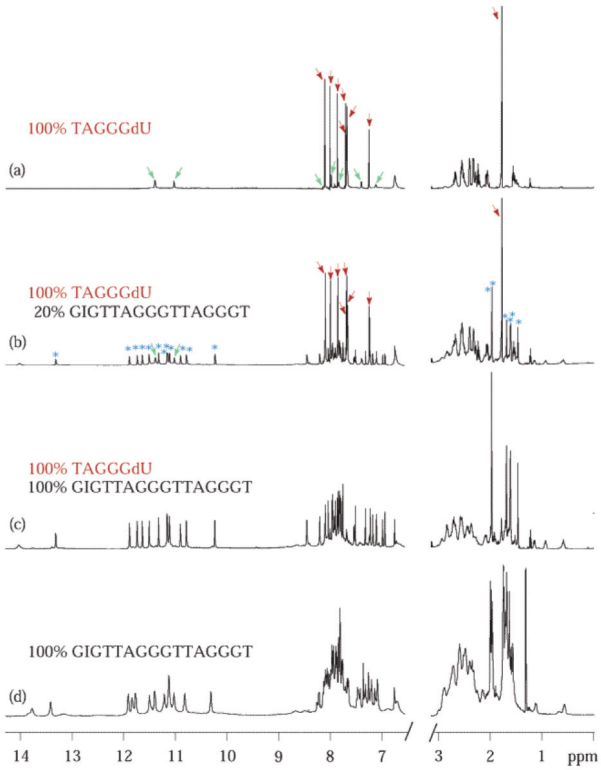

Figure 7.

Strand-mixing experiment: the titration of various amounts of *htel3I* into *htel1U* as monitored by proton NMR spectra. (a) Spectrum of 100% single-repeat human telomeric d(TAGGGU) sequence. Seven aromatic H2/6/8 peaks at 6.5-8.5 ppm and one methyl peak at 1.8 ppm from the major unfolded form are labeled with red arrows; peaks from the minor G-quadruplex form are labeled with green arrows. (b) Spectrum of 100% d(TAGGGU) in the presence of 20% d(GIGTTAGGGTTAGGGT). Eleven guanine imino protons at 10-12 ppm, one inosine imino proton at 13.4 ppm, and six methyl protons at 1-2 ppm from the newly formed quadruplex are labeled with blue stars. (c) Spectrum of 100% d(TAGGGU) titrated with 100% d(GIGTTAGGGTTAGGGT). (d) Spectrum of 100% d(GIGTTAGGGTTAGGGT). The strand concentration for 100% d(TAGGGU) and d(GIGTTAGGGTTAGGGT) is 5 mM.

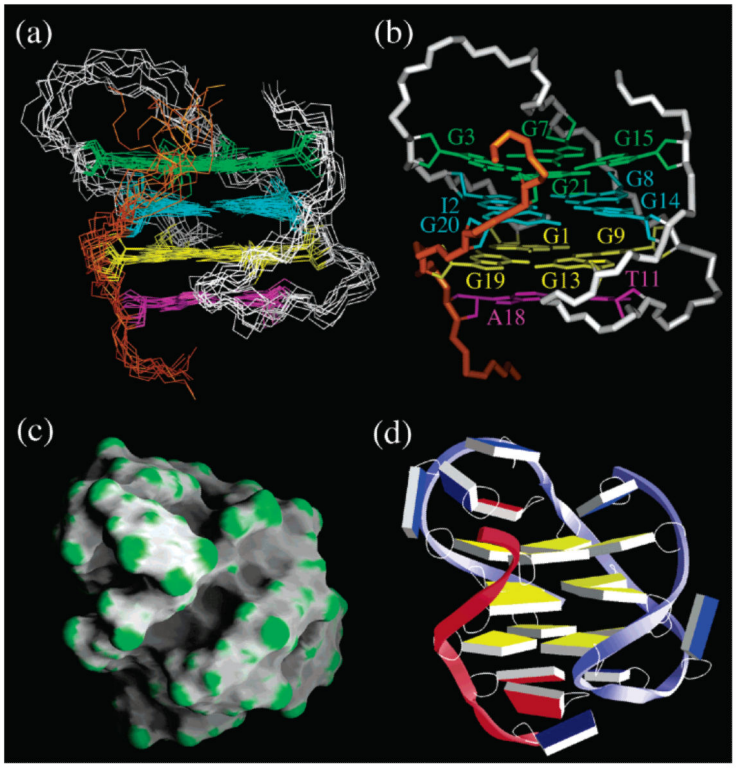

Figure 8.

Refined structures of the *htel3I/htel1U* quadruplex. Ten superpositioned lowest-energy intensity-refined structures (a) and one representative structure (b) of the *htel3I/htel1U* quadruplex. The backbone of *htel1U* and *htel3I* is colored in orange and gray, respectively. The Watson-Crick A18·T11 pair is colored magenta, the (G13·G19·G1·G9) tetrad is colored yellow, the (I2·G20·G14·G8) tetrad is colored cyan, and the (G3·G21·G15·G7) tetrad is colored green. Underdefined bases in the loops and terminal bases are not shown for clarity. Slab (c) and surface (d) views of a representative refined structure of the *htel3I/htel1U* quadruplex. In (d) the backbone of *htel1U* and *htel3I* is colored in red and blue, respectively; guanine and adenine bases are colored yellow and red, respectively; thymine and uracil bases are colored blue.

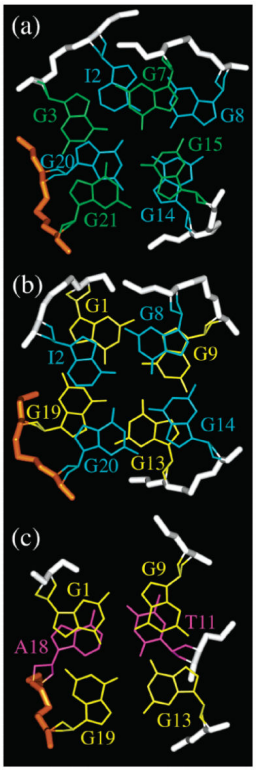

Figure 9.

Base-stacking overlap patterns in a representative refined structure of the *htel3I/htel1U* G-quadruplex. (a) Stacking between (G3·G21·G15· G7) tetrad (in green) and (I2·G20·G14·G8) tetrad (in cyan). (b) Stacking between (I2·*G20*·G14·G8) tetrad (in cyan) and (G13·G19·G1·G9) tetrad (in yellow). (c) Stacking between (G13·G19·G1·G9) tetrad (in yellow) and A18·T11 Watson-Crick base pair (in magenta).

Table 1

Statistics of the Computed Structures of the *htel3I/htel1U* Quadruplex

A. NMR Restraints	
total number of DNA distance restraints	517
exchangeable distance restraints	76
nonexchangeable distance restraints	390
hydrogen bond restraints	51
torsion angle restraints	22
B. Structural Statistics	
NOE violations	
number > 0.2 Å	0.5 ± 0.1
maximum violations (Å)	0.24 ± 0.04
rms deviation of violations (Å)	0.022 ± 0.003
deviation from the ideal covalent geometry	
bond length (Å)	0.013 ± 0.001
bond angle (deg)	2.67 ± 0.04
impropers (deg)	0.29 ± 0.05
pairwise rms deviation (Å) (10 refined structures)	
all heavy atoms	1.85 ± 0.27