Published online 10 July 2007 *Nucleic Acids Research, 2007, Vol. 35, No. 15* ***4927–4940***
doi:10.1093/nar/gkm522

Structure of the Hybrid-2 type intramolecular human telomeric G-quadruplex in K+ solution: insights into structure polymorphism of the human telomeric sequence

Jixun Dai[1], Megan Carver[1], Chandanamali Punchihewa[1], Roger A. Jones[2] and Danzhou Yang[1,3,4,*]

[1]College of Pharmacy, The University of Arizona, 1703 E. Mabel St, Tucson, AZ 85721, [2]Department of Chemistry and Chemical Biology, Rutgers University, 610 Taylor Road, Piscataway, NJ 08854, [3]Arizona Cancer Center, 1515 N. Campbell Avenue, Tucson, AZ 85724 and [4]BIO5 Institute, The University of Arizona, 1140 E. South Campus Dr, Tucson, AZ 85721, USA

Received May 28, 2007; Revised June 19, 2007; Accepted June 21, 2007

ABSTRACT

Formation of the G-quadruplex in the human telomeric sequence can inhibit the activity of telomerase, thus the intramolecular telomeric G-quadruplexes have been considered as an attractive anticancer target. Information of intramolecular telomeric G-quadruplex structures formed under physiological conditions is important for structure-based drug design. Here, we report the first structure of the major intramolecular G-quadruplex formed in a native, non-modified human telomeric sequence in K^+ solution. This is a hybrid-type mixed parallel/antiparallel-G-stranded G-quadruplex, one end of which is covered by a novel T:A:T triple capping structure. This structure (Hybrid-2) and the previously reported Hybrid-1 structure differ in their loop arrangements, strand orientations and capping structures. The distinct capping structures appear to be crucial for the favored formation of the specific hybrid-type intramolecular telomeric G-quadruplexes, and may provide specific binding sites for drug targeting. Our study also shows that while the hybrid-type G-quadruplexes appear to be the major conformations in K^+ solution, human telomeric sequences are always in equilibrium between Hybrid-1 and Hybrid-2 structures, which is largely determined by the 3′-flanking sequence. Furthermore, both hybrid-type G-quadruplexes suggest a straightforward means for multimer formation with effective packing in the human telomeric sequence and provide important implications for drug targeting of G-quadruplexes in human telomeres.

INTRODUCTION

Telomeres are specialized DNA sequences that cap the ends of chromosomes and play an important role in cancer (1–4), aging (5,6) and genetic stability (7–9). Human telomeric DNA consists of tandem repeats of hexanucleotide d(TTAGGG)$_n$ for 5–8 kb in length, terminating in a single-stranded 3′-overhang of 100–200 bases in length (10–14). In normal somatic cells, each cell replication results in a 50–200 base loss of the telomere, and after reaching a critical shortening of the telomeric DNA, the cell undergoes apoptosis (15). In contrast, telomeres of cancer cells do not shorten on replication, due to the activation of a reverse transcriptase telomerase that extends the telomeric sequence at the chromosome ends (16). Telomerase has been shown to be activated in 80–85% of human cancer cells (17), and has been suggested to play a key role in maintaining the malignant phenotype by stabilizing telomere length and integrity (18). The G-rich telomeric sequence can form DNA G-quadruplex structures consisting of stacked G-tetrad planes with Hoogsteen-type hydrogen bonds and stabilized by monovalent cations such as Na^+ and K^+. Formation of DNA G-quadruplex in the human telomeric sequence has been shown to inhibit the activity of telomerase, thus the intramolecular telomeric G-quadruplex has been considered to be an attractive target for cancer therapeutic intervention (1,3,4,19–21).

*To whom correspondence should be addressed. Tel: +1 520 626 5969; Fax: +1 520 626 6988; Email: yangd@pharmacy.arizona.edu

Structural information of the intramolecular human telomeric G-quadruplex formed under physiologically relevant conditions is necessary for structure-based rational drug design. Because the K^+ structure is considered to be biologically more relevant due to the higher intracellular concentration of K^+, it has been the subject of intense investigation (22–36), although the Na^+ structure was reported more than a decade ago (37). The minimal requirement for an intramolecular telomeric G-quadruplex is a four-G-tract human telomeric sequence (Figure 1A and B). We have recently determined the folding and solution structure of a hybrid-type mixed parallel/antiparallel-stranded intramolecular human telomeric G-quadruplex (Hybrid-1) in K^+ solution (Figure 1C, left) using a 26-nt sequence that contains the wild-type 22-nt four-G-tract human telomeric core sequence with modified flanking sequences (Tel26, Figure 1A) (38,39). An adenine triple capping structure was found to form in the Hybrid-1 telomeric G-quadruplex. The Hybrid-1 folding has also been independently reported by others (40,41).

Here we report the NMR solution structure of the major intramolecular G-quadruplex formed in a wild-type 26-nt four-G-tract human telomeric sequence wtTel26 in K^+ solution. This is the first intramolecular G-quadruplex structure of a biologically native, unmodified human telomeric sequence. This structure is also a hybrid-type mixed parallel/antiparallel-stranded G-quadruplex (Hybrid-2) (Figure 1C, right) but differs from the Hybrid-1 structure in its loop arrangements, strand orientations and capping structures. The Hybrid-2 folding topology has also been independently reported (42). In our NMR structure of the Hybrid-2 telomeric G-quadruplex, a novel T:A:T triple capping structure is found to form, which covers one end of the Hybrid-2 telomeric G-quadruplex and appears to play an important role in the selective stabilization of the Hybrid-2 structure. The structural information obtained from this study and our previous study (39) allowed us to understand the favored formation of a specific hybrid-type telomeric G-quadruplex, which appears to be largely determined by the specific capping structures. The distinct capping structures of each hybrid-type intramolecular telomeric G-quadruplex may provide specific binding sites for drug targeting. To gain insight into the structure polymorphism and equilibrium of human telomeric DNA, we have systematically examined the four-G-tract human telomeric sequences. We found that the hybrid-type G-quadruplexes appear to be the major conformations formed in these human telomeric sequences in K^+ solution and are always in equilibrium between Hybrid-1 and Hybrid-2 structures, which is largely determined by the 3′-flanking sequence. The energy barrier between the two hybrid forms appears to be rather small and can be readily shifted by minor changes. Intriguingly, both hybrid-type G-quadruplex structures suggest a straightforward means for multimer formation with effective packing in the human telomeric sequence (Figure 1D) and provide important implications for specific drug targeting of G-quadruplexes in human telomeres.

MATERIALS AND METHODS

Sample preparation

The DNA oligonucleotides were synthesized using β-cyanoethylphosphoramidite solid-phase chemistry on an Expedite™ 8909 Nucleic Acid Synthesis System (Applied Biosystem, Inc.) in DMT-on mode, and were purified using C18 reverse-phase HPLC chromatography, as described previously (38,39,43–45). Samples in D_2O were prepared by repeated lyophilization and final dissolution in 99.96% D_2O. Samples in water were prepared in 10%/90% D_2O/H_2O solution. The final NMR samples contained 0.1–2.5 mM DNA in 25 mM K-phosphate buffer (pH 7.0) and 70 mM KCl.

NMR experiments

NMR experiments were performed on a Bruker DRX-600 spectrometer as described previously (38,39,43,46). Standard 2D NMR experiments, including NOESY, TOCSY and DQF-COSY, were collected at 1, 5, 10, 15, 20, 25 and 30°C to obtain the complete proton resonance assignment. Non-exchangeable protons were estimated based on the NOE cross-peak volumes at 50–300 ms mixing times, with the upper and lower boundaries assigned to ±20% of the estimated distances. Distances between exchangeable protons were assigned with looser boundaries of ±25%. The methyl base proton Me-H6 distance (2.99 Å) was used as a reference. The distances involving the unresolved protons, e.g. methyl protons, were assigned using pseudo-atom notation in X-PLOR.

Distance geometry and simulated annealing (DGSA) calculations

Metric matrix distance geometry (MMDG) and simulated annealing calculations were carried out in X-PLOR (47) to embed and optimize 100 initial structures based on an arbitrary extended conformation for the single-stranded wtTel26 sequence, as described previously (39,43). The experimentally obtained distance restraints and G-tetrad hydrogen-bonding distance restraints were included during the calculation.

NOE-distance restrained molecular dynamics calculations

All of the 100 molecules obtained from the DGSA calculations were subjected to NOE-restrained simulated annealing refinement in XPLOR (47) with a distance-dependent dielectric constant, as described previously (39,43). The force constants were scaled at 10–30 and 80–100 $kcal\,mol^{-1}\,Å^{-2}$ for NOE and hydrogen bond distance restraints, respectively. A total of 727 NOE distance restraints, of which 324 are from inter-residue NOEs, were incorporated into the NOE-restrained structure calculation.

Dihedral angle restraints were used to restrict the glycosidic torsion angle (χ) for the experimentally assigned *syn* and *anti* conformations. A dihedral angle restraint of 60(±35)° was applied to the *syn* G-tetrad guanines, and a dihedral angle restraint of 240(±40)° was applied to the *anti* G-tetrad guanines. The force constants of dihedral angle restraints were 10 $kcal \cdot mol^{-1} \cdot rad^{-2}$.

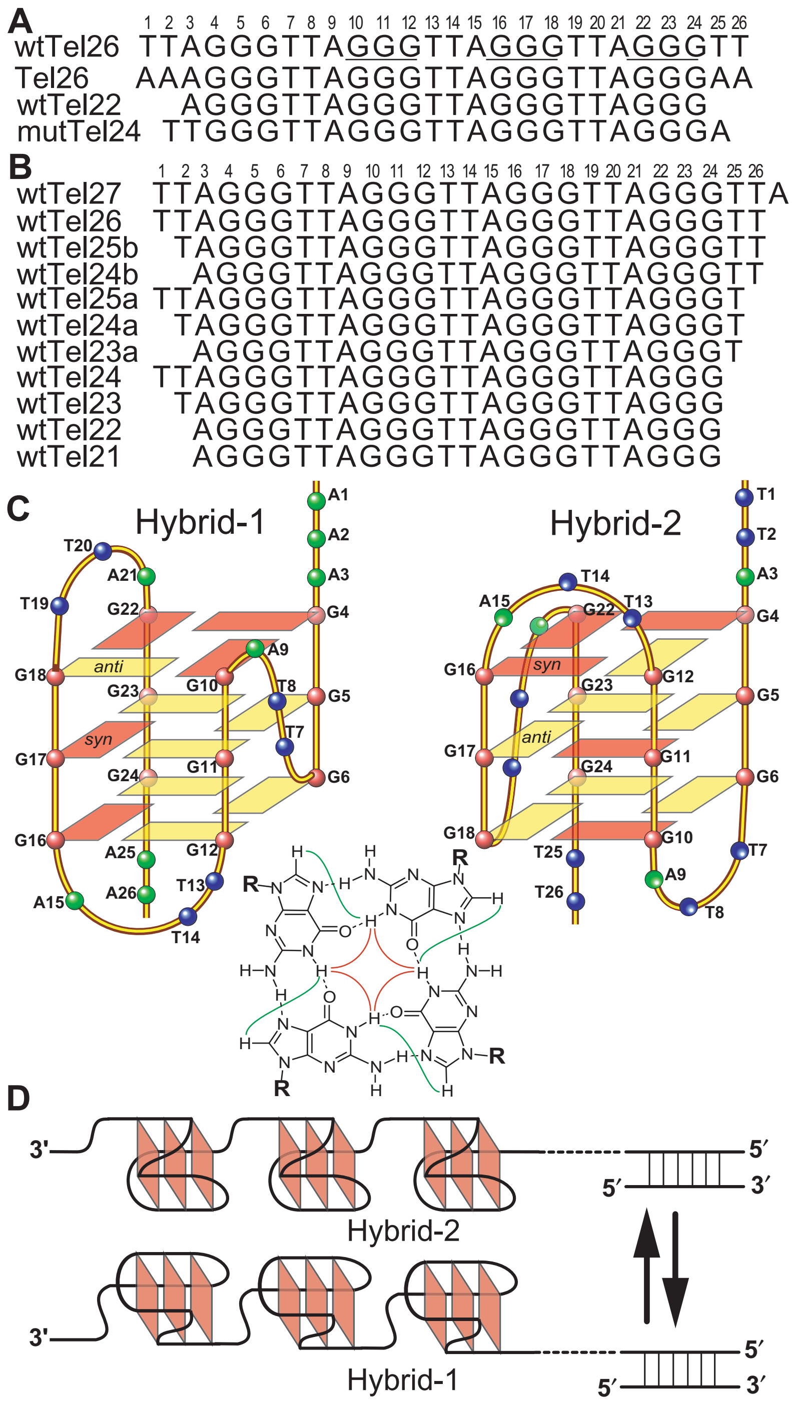

Figure 1. (**A**) Four-G-tract human telomeric DNA sequences that have been used for structure determination. The numbering system is shown above wtTel26. (**B**) Four-G-tract native human telomeric sequences with different flanking sequences. The numbering system is shown above wtTel27. (**C**) Schematic drawing of the folding topologies of the Hybrid-1 (major conformation in Tel26) and Hybrid-2 (major conformation in wtTel26) intramolecular telomeric G-quadruplexes in K^+ solution. A G-tetrad with H1–H1 and H1–H8 connectivity patterns detectable in NOESY experiments is also shown in the inset. Yellow box = (*anti*) guanine, red box = (*syn*) guanine; red ball = guanine, green ball = adenine, blue ball = thymine. (**D**) A schematic model of DNA secondary structure in human telomeres. The hybrid-type telomeric G-quadruplex structures can be readily folded and stacked end to end to form compact-stacking structures for multimers in the elongated telomeric DNA.

NOE-restrained simulated annealing refinement calculations were performed as described previously (39,43). The time steps for all processes of heating, cooling and equilibration were set to 1 fs. The 10 best molecules were selected based both on their minimal energy terms and number of NOE violations and have been deposited in the Protein Data Bank (PDB ID 2JPZ).

RESULTS

The wild-type 26-nt human telomeric sequence forms a major intramolecular G-quadruplex structure in K^+ solution

We have very recently reported a hybrid-type intramolecular telomeric G-quadruplex structure formed in potassium solution using a 26-nt sequence, which contains the wild-type 22-nt four-G-tract human telomeric sequence that was used for the previous NMR and X-ray structural studies (22,37) with a flanking AA at each end (Tel26, Figure 1A). During this work, we found that the Tel26 sequence forms two stable G-quadruplex conformations, with well-defined guanine imino-peaks, when examined right after it was prepared in K^+ solution (Figure S1). The second conformation (~40%) slowly converts to the first or final conformation after overnight, while the complete conversion takes about 1 day (Figure S1). The CD spectrum of Tel 26 is very similar to that of wtTel26, indicating that the two conformations are similar in nature. We decided to go back and re-examine the wild-type 26-nt telomeric sequence wtTel26, $(TTAGGG)_4TT$ (Figure 1A).

The 1D ^1H NMR spectrum of the wtTel26 sequence in K^+ solution is shown in Figure 2 (top). It shows a major species with twelve resolved imino proton resonances between 10.5 and 12 p.p.m., characteristic of a G-quadruplex structure. The NMR resonances of this major species have sharp line widths (5–6 Hz at 25°C), suggesting a unimolecular G-quadruplex structure. Minor conformations are also present, as indicated by the presence of weak and broader resonances, which account for ~20–25% of the total population. The melting temperature of this major conformation is around 53°C (Figure S2). The melting temperature is concentration independent as shown by both NMR and CD, indicating that the major G-quadruplex structure is unimolecular.

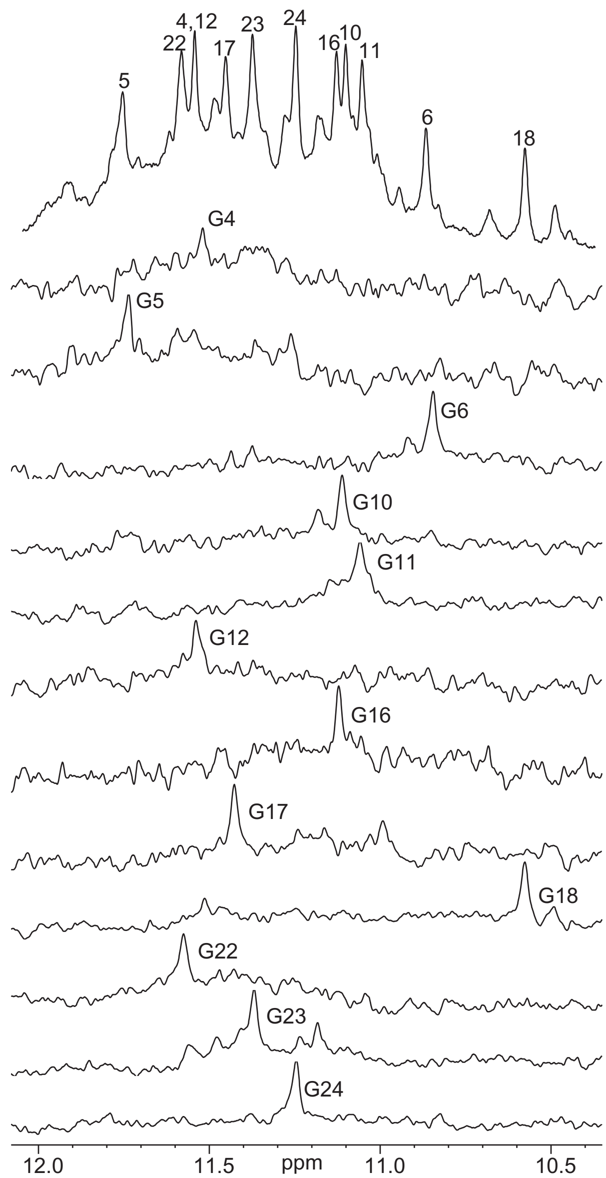

Figure 2. The imino proton region with assignment of the 1D ^1H NMR spectrum of wtTel26 in K^+ solution (top), and imino proton assignments of wtTel26 using 1D ^{15}N-filtered experiments on site-specific labeled oligonucleotides. Conditions: 25 mM K-PO_4, 70 mM KCl, pH 7.0, 25°C, 0.5–0.6 mM DNA.

Complete proton assignment was accomplished for wtTel26 in K^+

The presence of twelve imino peaks indicates that the major G-quadruplex structure contains three G-tetrads. Using ^{15}N-filtered experiments as previously reported (38,39,43,45,46), the imino (one-bond connection to N1) and base aromatic H8 (two-bond connection to N7) protons of tetrad guanines were unambiguously assigned by site-specific low-enrichment (6%) of 1, 2, 7-^{15}N-labeled guanine nucleoside at each guanine position one at a time. The assignment of each imino proton H1 of the twelve guanines involved in the three G-tetrads is shown in Figure 2. The base H6 and methyl proton resonances of thymines have been unambiguously assigned by substituting deoxyuridine (dU) for each thymine one at a time (Figure S3). We used multiple 2D NMR experiments, including 2D-NOESY, TOCSY and COSY, at various temperatures, to assign the proton resonances of wtTel26 in K^+. All proton resonances, except some of the H5′/H5″ protons, were unambiguously assigned (Table 1).

It needs to be noted that it is a challenging process to get complete resonance assignment for wtTel26, as there are clearly minor species. To get an unambiguous assignment, we have collected data at various temperatures, 1, 5, 10, 15, 20, 25 and 30°C, with multiple mixing times, in both D_2O and H_2O, as the chemical shifts are not only temperature dependent but also solvent dependent. The complete spectral assignment was achieved by examining all of these spectra at different conditions. We made sure that the same NOE assignment pathway for the major conformation can be followed for all experimental conditions, i.e., the connectivities with the adjacent residues can be followed and are consistent at different conditions. For weak NOE peaks, only those that could be consistently observed at all different conditions and that follow the same connectivities with

Table 1. Proton chemical shifts for the wtTel26 at 15°C[a]

	H6/H8	H2/H5/Me	H1′	H2′/H2″	H3′	H4′	H5′/H5″	NH1/NH3
T1	7.25	1.53	5.74	1.82,2.05	4.30	3.70	3.36,3.40	
T2	6.92	1.18	5.49	1.4,1.93	4.24	3.77	3.49	
A3	8.18		6.00	2.10,2.52	4.65	4.02	3.21,3.62	
G4[c]	7.39		5.93	2.90,3.43	4.76	4.29	4.04,4.18	11.47
G5	7.83		5.64	2.32,2.58	4.90	4.17	4.09,4.23	11.69
G6	7.64		5.85	2.21,2.24	4.86	4.26	4.08	10.78
T7	7.57	1.72	6.11	2.05,2.40	4.60	4.21	3.94,4.10	10.27[b]
T8	7.32	1.64	5.66	1.82,2.05	4.46	3.88	3.78,3.82	10.26[b]
A9	7.68	7.48	5.67	1.85,2.34	4.53	4.01	2.83,3.42	
G10	7.20		5.94	3.61,3.05	4.72	4.29	3.97	11.02
G11	7.27		5.65	2.44,2.33	4.88	4.09	4.17,4.20	10.96
G12	7.62		5.79	2.45,2.25	4.93	4.24	4.03	11.47
T13	7.60	1.78	5.97	2.15,2.45	4.74	4.27	3.99,4.09	
T14	7.28	1.55	5.81	1.99,1.59	4.63	4.08	3.87	
A15	8.16		6.12	2.83,2.68	4.90	4.26	3.75,3.89	
G16	7.00		5.85	2.84,3.33	4.85	4.35	4.17	11.06
G17	7.52		5.72	2.57,2.24	4.80	4.11	4.24	11.34
G18	7.52		6.05	2.45,2.36	4.86	3.92	3.93,3.77	10.49
T19	7.67	1.82	6.27	2.39,2.28	4.62	4.32	3.87,4.08	
T20	7.71	1.93	6.24	2.26,2.45	4.81	4.19	4.01,4.07	
A21	8.29		6.32	2.67	4.91	4.33	4.06,4.02	
G22	7.03		5.74	2.77,2.95	4.72	4.28	4.22	11.48
G23	7.80		5.66	2.55,2.42	4.90	4.21	4.13,4.06	11.30
G24	7.69		6.04	2.59,2.39	4.79	4.34	4.07,4.12	11.18
T25	7.24	1.30	5.96	1.98,2.29	4.63	4.11	4.10,3.98	
T26	7.14	1.38	5.55	1.85	4.25	3.68	3.93,3.81	

[a]The chemical shifts are measured in 25 mM K-phosphate, 70 mM KCl, pH 7.0 referenced to DSS.
[b]The chemical shifts are measured at 5 °C.
[c]Guanines in *syn* glycosidic conformation are in bold.

the flanking residues were used for structure determination and calculation.

wtTel26 adopts a second hybrid-type (Hybrid-2) G-quadruplex folding structure in K^+

The assignment of the imino and base H8 protons of guanines leads to the direct determination of the folding topology of the major G-quadruplex structure formed in wtTel26 in K^+ solution (Figure 1C, right). In a G-tetrad plane with a Hoogsteen-type H-bond network, the imino proton NH1 of a guanine is in close spatial vicinity to the NH1s of the two adjacent guanines, and to the base H8 of one of the adjacent guanines (Figure 1C). Three G-tetrad planes were determined based on the NOE connectivities of exchangeable protons (Figure 3A and B). For example, the GH1/GH1 NOE interactions, including G6H1/G10H1, G10H1/G18H1, G18H1/G24H1 and G24H1/G6H1 (Figure 3A); and the GH1/GH8 NOE interactions, including G6H1/G10H8, G10H1/G18H8, G18H1/G24H8 and G24H1/G6H8 (Figure 3B), define a G-tetrad plane of G6-G10-G18-G24 (Figure 1C, right). The overall G-quadruplex alignment is further defined based on the inter-tetrad NOE connections from residues that are positioned far apart in the DNA sequence. For example, the strong NOE interactions of G5H1/G12H1, G11H1/G16H1 and G17H1/G22H1 (Figure 3A) connect the top two G-tetrads and define their reversed G-arrangements (Figure 1C, right). The sequential inter-tetrad NOE interactions, including G5H1/G6H1, G17H1/G18H1 and G23H1/G24H1 (Figure 3A) indicate the same G-conformations of the bottom two G-tetrads (Figure 1C, right), while the inter-residue NOEs between the two G-tetrads, e.g. G6H1/G11H8, G10H1/G17H8, G18H1/G23H8 and G24H1/G5H8 (Figure 3B), reflect the right-handed twist of the DNA backbone.

An expanded region for base and sugar H1′ protons of non-exchangeable proton NOESY is shown in Figure 3C. Five guanine residues are in *syn*-conformation, including G4, G10, G11, G16 and G22, as indicated by the very strong H8–H1′ NOE intensities (Figure 3C). The regular sequential NOE connectivities are either missing or very weak at the N(i)− *syn*G(i + 1) steps, i.e. A3–G4, A9–G10, G10–G11, A15–G16 and A21–G22. The characteristic *syn*G(i)H8/G(i + 1)H1′ NOEs are observed, including those of G4–G5, G10–G11, G11–G12, G16–G17 and G22–G23 (Figure 3C). A characteristic downfield shift is observed for the H2′/H2″ sugar protons of the *syn*-guanines (Table 1). For the N(i)− *anti*G(i + 1) steps, the sequential NOE crosspeak connectivities of the base H8 protons to the 5′-flanking residue sugar H1′/H2′/H2″ protons, typical for right-handed DNA twist, are clearly observed (Figure 3C).

The major G-quadruplex in the wtTel26 also adopts a hybrid-type folding structure, but is different from that found in the Tel26 sequence, so we named this new folding the Hybrid-2 type and the first one the Hybrid-1 type (Figure 1C). Hybrid-2 folding differs from Hybrid-1 in its arrangement of the three TTA loops, which is lateral-lateral-side versus side-lateral-lateral. Thus the first, third and fourth G-tracts of Hybrid-2 folding are

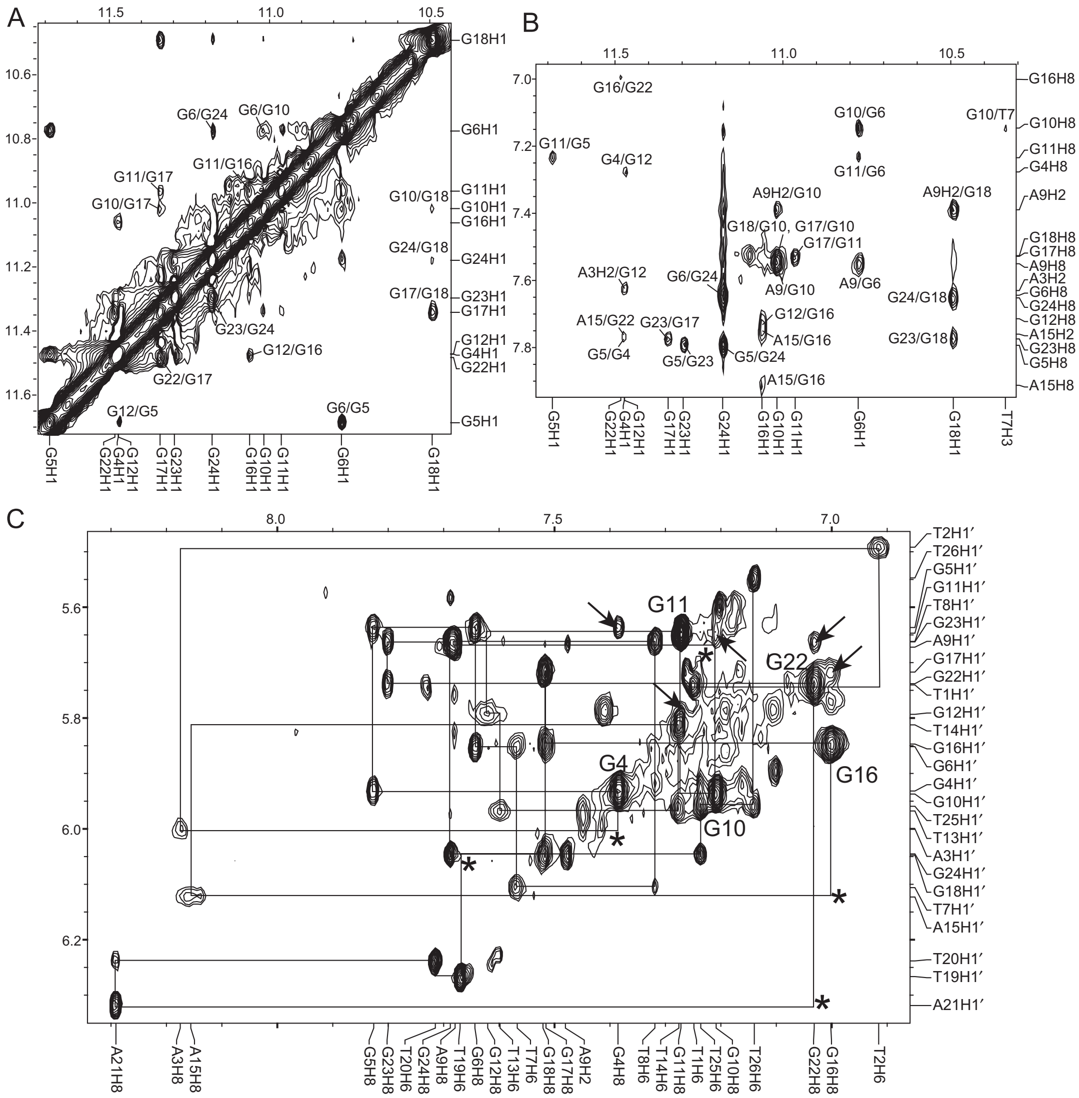

Figure 3. The expanded H1–H1 (**A**) and H1-H8/H2/H6 (**B**) regions, with labeling, of the exchangeable proton 2D JR-NOESY spectrum of wtTel26 in K^+ solution. In (B), only the aromatic H2 protons are specified in the labels. (**C**) The expanded H8/H6–H1′ region of the non-exchangeable 2D-NOESY spectrum 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. A3 and A15 show much broader peaks, likely related with the dynamic motion. 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 pH 7.0 K-phosphate, 70 mM KCl, 2.5 mM DNA.

parallel while the second G-tract is antiparallel, whereas in Hybrid-1 folding the third G-tract is the antiparallel one. The same number of *syn* guanines was observed in the Hybrid-1 folding in the Tel26 sequence (38,39), with the only difference being the *syn*-G11 in Hybrid-2 versus *syn*-G17 in Hybrid-1.

NOE-restrained structure calculation

Many inter-residue NOE interactions are observed in 2D-NOESY of wtTel26 in K^+. Critical inter-residue NOE interactions are schematically summarized in Figure 4, which immediately define the overall structure of the Hybrid-2 G-quadruplex formed in wtTel26 in K^+. Except for residues A21 and T26 whose H2′ and H2″ protons are overlapping, all the other residues have resolved H2′ and H2″ resonances, and the H1′–H2″ NOE is stronger than the H1′–H2′ NOE while the H1′–H4′ is considerably weaker, indicating the C2′-endo sugar pucker conformations. All residues in the loop regions are in the *anti* conformation except T19, which shows strong H8–H1′ NOEs (Figure 3C) even at a short mixing time. Solution structures of the Hybrid-2 type human telomeric G-quadruplex were calculated using a NOE-restrained distance geometry (DGSA) and molecular dynamics (RMD) approach (Figure 5, PDB ID 2JPZ), starting from an arbitrary extended single-stranded DNA. A total of 727 NOE distance restraints, of which 324 are from inter-residue NOE interactions, were incorporated into the NOE-restrained structure calculation (Table 2). Very few NOEs associated with H5′ or H5″ were used for structure calculation, due to possible ambiguities

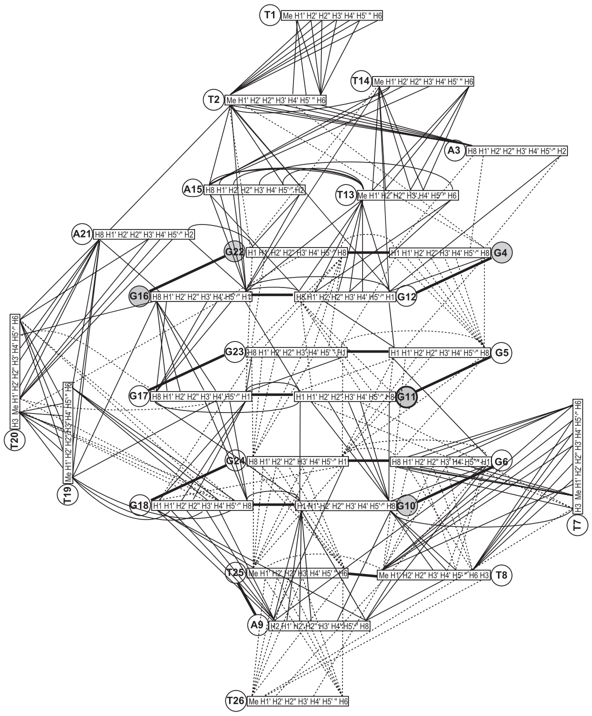

Figure 4. Schematic diagram of inter-residue NOE connectivities of wtTel26 G-quadruplex formed in K^+ solution. The guanines in *syn* conformation are represented using gray circles. The NOE connectivities clearly define the G-quadruplex conformation and provide distance restraints for structure calculation.

associated with H5′ or H5″ protons. Dihedral angle restraints were used for the glycosidic torsion angle (χ) based on the experimentally determined *syn* and *anti* conformations. The structure statistics are listed in Table 2. Remarkably, like the Hybrid-1 human telomeric G-quadruplex structure (39), the Hybrid-2 telomeric G-quadruplex structure is very well defined (Figure 5A). For the 10 best NMR refined structures, the RMSD is 1.38 Å for all residues except the 5′-terminal T1 (Table 2).

The molecular structure of wtTel26 in K^+: a very well-defined T:A:T capping structure is formed at the 3′-end of the Hybrid-2 human telomeric G-quadruplex

A representative model of the Hybrid-2 type human telomeric G-quadruplex structure in K^+ is shown in two different views in Figures 5B. The guanine distribution of the three parallel G-strands (first, third and fourth) is (5′-*syn-anti-anti*), while that of the antiparallel G-strand (second) is (5′-*syn-syn-anti*). The widths of the four grooves are 12.98 Å (groove I), 8.62 Å (groove II), 13.56 Å (groove III) and 9.88 Å (groove IV), as measured by the closest P-P distance across groove, with groove III occupied by a double-chain-reversal loop.

A very well-defined T:A:T capping structure is formed at the bottom end of the Hybrid-2 telomeric G-quadruplex structure in K^+ (Figure 6A). It appears that the 3′-flanking-TT of wtTel26 (Figure 1A) plays an important role in the favored formation of the Hybrid-2 structure (Figure 1C, right). T8 and A9 of the first TTA lateral loop, as well as T25 of the 3′-flanking segment, appear to form a T:A:T triple capping structure, with potential H-bonds formed between T8 and T25, and between T25 and A9 (Figure 6B). Both imino H3 protons of T8 and T25 were observed in NMR even at 15°C. Clear NOEs were observed between T8H3 and T26H1′ as well as T25H1′. The 3′-terminal T26 appears to stack above the T:A:T triple. It is interesting to note that, in the modified Tel26 sequence (Figure 1A), the A25 of the 3′-flanking-segment was involved in the formation of a stable A:T base pair with T14 and appears to be important for the stabilization of the Hybrid-1 structure (Figure 1C, left) (38,39) (see Discussion).

In accord with the T:A:T triple capping structure, T8, A9 and T25 all appear to stack very well with the bottom tetrad (Figures 5 and 6A). A9 is positioned largely underneath the G18 base. A9H8 has strong NOEs with both G10H1 and G6H1, while A9H2 has strong NOEs with G10H1 and G18H1 (Figure 3B). A9H2 shows clear NOEs with sugar protons of G18 (Figure 4), including H1′ (medium-strong), H2′/″ (strong) and H3′ (medium), indicating the H2 end of A9 is quite close to the sugar moiety of G18. An unusually large number of NOEs were observed between T8 and G6, including T8Me/G6H8, as well as NOEs between the base protons Me&H6 of T8 and G6 sugar protons, e.g. G6H1′, H2′/″ and H3′ (Figure 4). Such NOEs are indicative of two stacking bases. T8Me shows clear NOEs with T26H1′/H4′, as well as T25 H1′ (Figure 4), consistent with the T:A:T triple capping conformation (Figure 6A). Clear NOEs were observed between the sugar protons of T8 and G6H1 as well as G10H8 (Figure 4), indicating the T8 sugar is stacked below G6 and G10 of the bottom tetrad. In fact, most base and sugar protons of T8 and A9 are significantly upfield shifted (Table 1), indicating the two residues are covered by the bottom G-tetrad, whose ring-current effect likely induces the resonance upfield shifting of T8 and A9. A9H8 showed unusual medium-strong NOEs with T8H1′/H4′, but not H2′/″ and H3′, indicating a left-handed conformation at the T8-A9 step with the A9 positioned at the H1′/H4′ side of T8. T7, which is adjacent to G6 of the bottom G-tetrad, appears to fall into groove I at the H1′/H4′ side of the G6 sugar, as supported by clear NOEs between G6H4′/H1′ and T7Me/H6 combined with weak sequential NOEs between T7H6/Me and G6H2′/″ and H3′ (Figure 4). Potential H-bonds could be formed between T7O2 and G6NH21, and between T7O4 and G5NH21. The H3 imino proton of T7 was also observed in NMR at low temperatures. The 3′-terminal G24-T25-T26 segment is very well stacked (Figures 5 and 6A), as supported by clear sequential NOE connectivities observed in this region (Figure 3C), as well as the NOEs between base protons, such as G24H8 with T25H6/Me T25H6 with T26H6/Me, and T25Me with G18H1 (strong) and G24H1 (medium) (Figure 4). Consistent with the stacking conformation, the base and H1′, H2′/″ sugar protons of T25 and T26 are markedly upfield shifted (Table 1).

As observed in the Hybrid-1 structure (Figure 1C, left) (39), A21 of the double-chain-reversal loop, formed by the third TTA linker in the Hybrid-2 structure (Figures 1C right and 6C), appears to position partially above G16

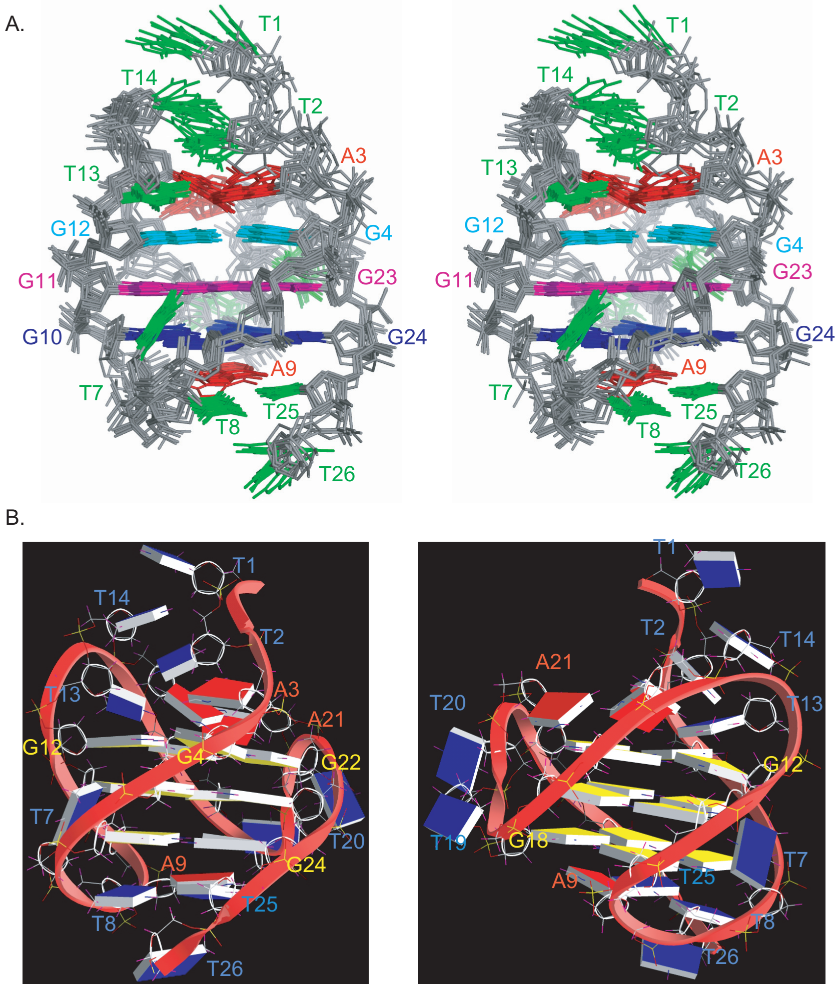

Figure 5. (**A**) Stereo view of the superimposed 10 NMR-refined structures of the wtTel26 G-quadruplex in K^+ solution. The loop adenines are in red, while the loop thymines are in green. The top G-tetrad (as in Figure 1C) is in cyan, the middle G-tetrad is in magenta, and the bottom G-tetrad is in blue. (**B**) A representative model of the NMR-refined wtTel26 G-quadruplex structure from two different views, prepared using GRASP (56) (guanine = yellow, adenine = red, thymine = blue).

and G22 of the top G-tetrad with its H2 end pointing toward the G-tetrad center, as indicated by NOEs between A21 and G16/G22 (Figure 4), similar to A9 of the first TTA loop in the Hybrid-1 structure. The base and sugar protons of A21 are downfield shifted (Table 1), as also observed for A9 in the Hybrid-1 structure, likely due to their partially stacked position with the top G-tetrad. T20 and T19 appear to position in groove III (Figure 6C), with T19 positioned closer to the core G-tetrads as indicated by a number of NOEs between T19 and the third G-tract, G16-G17-G18 (Figure 4).

The loop and flanking residues above the top end of the Hybrid-2 human telomeric G-quadruplex are not as well structured as those covering the bottom end (Figure 5). A3 of the 5′-flanking segment, T13 and A15 of the second TTA lateral loop, and A21 of the third TTA double-chain-reversal loop, all appear to position right above the top G-tetrad, forming the first layer of capping at the top end; however, no H-bond formations were detected for this layer. A15 appears to stack well with the top tetrad, as indicated by NOEs between A15H2 and G22H1/G16H1, and between A15H8 and G16H1/G12H8 (Figure 4). T13 appears to stack over G12, as indicated by NOEs between T13H6/Me and the sugar protons of G12, and between T13Me and G12H8/G16H1 (Figure 4). The methyl end of T13 appears to be close to A15, as a number of NOEs were observed between T13Me and A15H8 as well as A15 sugar protons (Figure 4). A3 is positioned above the G4 and G12 of the top G-tetrad, with its H2 end pointing toward G12, as indicated by NOE interactions of A3H8/G4H1 and A3H2/G12H1 (Figure 4). T2 of the 5′-flanking segment and T14 of the second TTA loop appear to position above the first layer, while the 5′-terminal T1 is not very well defined.

Mutational analysis of wtTel26 supports the T:A:T triple capping structure of Hybrid-2 human telomeric G-quadruplex and provides insights into versatile loop interactions

WtTel26 contains four TTA segments, including the 5′-flanking TTA and three TTA loops. Adenine residues

Table 2. Structural statistics for the wtTel26[a]

NMR distance and dihedral constraints	
Distance restraints	
Total NOE	727
Intraresidue	403
Interresidue	
Sequential($\|i-j\| = 1$)	225
Non-sequential($\|i-j\| > 1$)	99
Hydrogen bonds	25
Total dihedral angle restraints	25
Structure statistics	
NOE violations	
Number (>0.2 Å)	1 ± 0.9
r.m.s.d. of violations(Å)	0.036 ± 0.003
Deviations from idealized geometry	
Bond length (Å)	0.008 ± 0.0001
Bond angle (°)	1.49 ± 0.02
Impropers (°)	1.07 ± 0.03
Pairwise r.m.s.d. of heavy atoms (Å)	
G-tetrads + 3′-capping(residue 8,9,25)	1.03 ± 0.09
Residues A2–T26	1.38 ± 0.16
All residues	1.50 ± 0.21

[a]The ensemble of 10 structures is selected based both on the minimal energy terms and number of NOE violations.

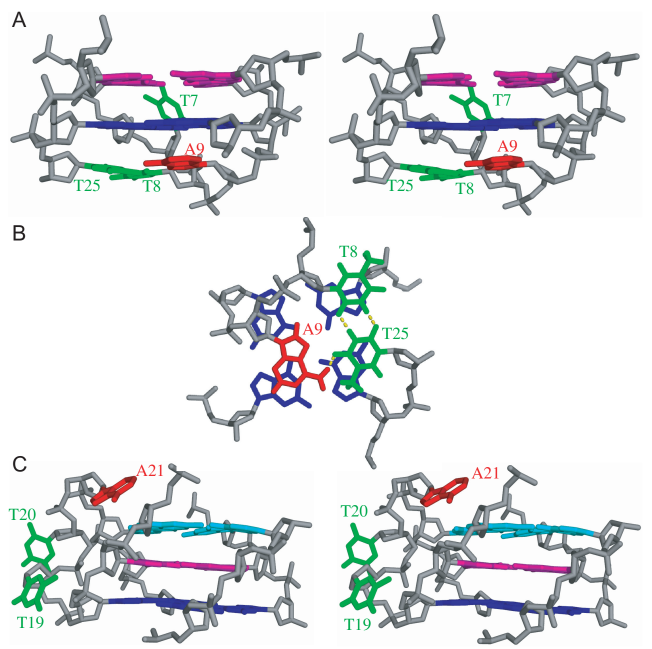

Figure 6. Loop conformations in the Hybrid-2 wtTel26 quadruplex in K^+ solution. (**A**) Stereo views of the T:A:T triple capping structure formed by T8, A9 and T25. (**B**) The bottom view of the T:A:T triple capping the bottom G-tetrad (blue). The potential hydrogen bonds are shown, with the distances of: 1.74 Å for T8H3–T25O2, 1.81 Å for T25H3–T8O2 and 1.6 Å for A9H61–T25O4. (**C**) Stereo views of the third TTA loop (T19–T20–A21) which adopts a double-chain-reversal loop conformation. The color code is the same as that in Figure 5A.

have been shown to be largely involved in the formation of capping structures, which is an important factor for the stability and thus for the preferred formation of a particular G-quadruplex structure. We have carried out systematic mutational analysis of wtTel26 to determine the functional role of adenines in the stability of the Hybrid-2 telomeric G-quadruplex structure, using A-to-T mutation one at a time for each adenine, including A3, A9, A15 and A21. The 1D NMR spectra of each A-to-T mutant sample are shown in Figure 7. In accord with the important role of A9 in the T:A:T capping structure (Figure 6B), mutation of A9T clearly destabilizes the Hybrid-2 structure. Interestingly, it appears that the major imino peaks of wtTel26-A9T resemble those of Tel26 (**Figure S1**), which forms a Hybrid-1 structure (**Figure 1C, left**), implying the A9T mutation may favor the Hybrid-1 structure. A possible explanation is that while A9T destabilizes the T:A:T triple capping structure in the Hybrid-2 structure, the thymine of the A9T mutant may form a stable H-bonded base pair with A21 and thus favor the Hybrid-1 structure (**Figure 1C, left**). In addition, while wtTel26 is a mixture of multiple conformations, with the major conformation (~75%) being the Hybrid-2 type telomeric G-quadruplex structure, the Hybrid-2 type structure only accounts for ~50% in the T25U sequence with a single T-to-U substitution at position 25 (**Figure S3**). This result implies the importance of T25 for the favored formation of Hybrid-2 structure; and that T8:A9:U25 triple is not as stable as the T8:A9:T25 triple capping structure, as the methyl group of a thymine (T25 here) is likely to be important for the stacking interactions. The important role of T25 was also reflected by the observation that four-G-tract telomeric sequences lacking the 3′-flanking segment (wtTel24, wtTel23) (**Figure 1B**) appear to form a Hybrid-1 type major structure (42).

In contrast, A-to-T mutations at position 3 or 15 markedly improve the NMR spectra (Figure 7), indicating the formation of a more predominant G-quadruplex structure. The folding structures of both the mutant sequences, A15T and A3T, has been shown to be Hybrid-2 type by unambiguous assignment of the tetrad guanine imino protons using site-specific incorporation of ^{15}N-labeled guanine (**Figure S4**). It is likely that a stable base-paired capping structure can be formed at the top end by the two mutations, e.g. (A15T):A3 or A15:(A3T), respectively, which further stabilizes the Hybrid-2 structure. In addition, the A21T mutation appears to destabilize the Hybrid-2 conformation (Figure 7).

Major intramolecular G-quadruplex structures formed in four-G-tract wild-type human telomeric sequences with different flanking sequences

As we found that human telomeric sequences can adopt both Hybrid-1 and Hybrid-2 types intramolecular G-quadruplexes, we have carried out systematic analysis to obtain insights into the major G-quadruplex conformations of different wild-type telomeric sequences in K^+ solution. We found that the 3′-flanking segment is important for the formation of a specific G-quadruplex conformation, hence we systematically prepared the wild-type four-G-tract human telomeric sequences with various 3′-flanking segments of null, T, TT and TTA, respectively (Figure 1B). The 1D ^{1}H NMR spectra of these sequences in K^+ solution are shown in Figure 8. It appears that all sequences form a mixture of multiple

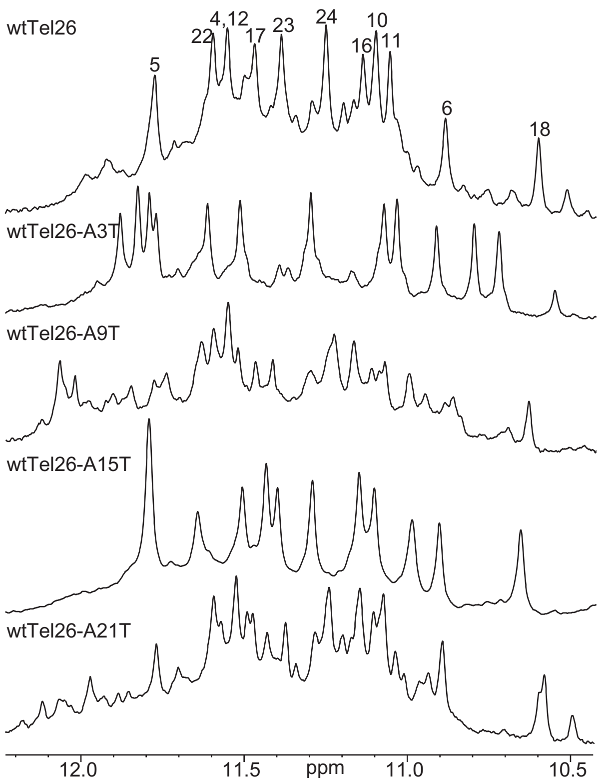

Figure 7. The imino proton region of the 1D ^{1}H NMR spectra of wtTel26 (with assignment) and modified wtTel26 sequences with a single A-to-T mutation, in K^{+} solution at 25°C.

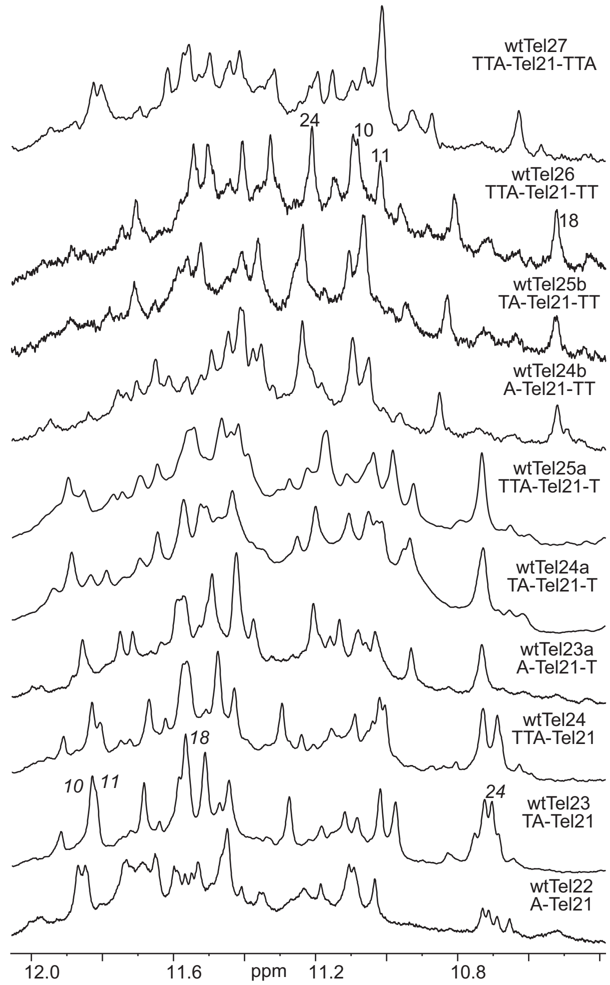

Figure 8. The imino proton region of the 1D ^{1}H NMR spectra of four-G-tract wild-type human telomeric sequences (Figure 1B) in K^{+} solution at 25°C. The characteristic imino peaks for Hybrid-2 structure (wtTel26) and Hybrid-1 structure (wtTel23, Italic) are labeled.

conformations, with a major conformation existing in most cases. It is clear that the 3′-flanking segment has a determinant role in the major G-quadruplex structure formation in human telomeric sequences, as the four-G-tract sequences with the same 3′-flanking segment give rise to very similar 1D NMR spectra (Figure 8).

Human telomeric sequences with the 3′-flanking segment of TT (wtTel26, wtTel25b, wtTel24b) (Figure 1B) all appear to form the major Hybrid-2 structure as observed for wtTel26 (Figure 8). This observation is in good agreement with the NMR structure of the Hybrid-2 type intramolecular telomeric G-quadruplex, in which the 3′-flanking TT is important for the formation of a stable T:A:T triple capping structure that selectively stabilizes the Hybrid-2 type structure. For the human telomeric sequence wtTel27 with a 3′-flanking TTA segment (Figure 1B), the Hybrid-2 structure still appears to be the major conformation (Figures 8 and S5).

Interestingly, while it was shown in our structure that T25 is very important for the Hybrid-2 structure, human telomeric sequences with 3′-flanking-T (wtTel25a, wtTel24a, wtTel23a) gave rise to 1D NMR spectra not very similar to that of wtTel26 and other telomeric sequences with 3′-flanking TT. In fact, those spectra look more like that of wtTel23, which has been shown to form a major structure of Hybrid-1 type (42). In order to understand the effect of the 3′-flanking sequences, we decided to determine the major conformation in these four-G-tract telomeric sequences, using site-specific incorporation of ^{15}N-labeled guanines for unambiguous assignment of tetrad guanine imino protons for each sequence. For wtTel25a with a 3′-flanking T (T25) (Figure 1B), the Hybrid-2 type structure also appears to be the major conformation (Figures 9 and S6). These results are in good agreement with the important role of the 3′-flanking segment, especially T25, in selective stabilization of Hybrid-2 structure due to the stable formation of a T:A:T triple capping structure. However, while wtTel24a has the same 3′-flanking T as wtTel25a (Figure 1B), it appears that wtTel24a forms a mixture of two conformations with similar populations (~50–50 distribution) (Figures 9 and S6), as a second set of imino peaks was detected for each ^{15}N-G labeling. This second set of guanine imino protons appears to be consistent with the Hybrid-1 type structure, and was also detected in the wtTel25a sequence, with a much smaller population (~25%) (Figure 9).

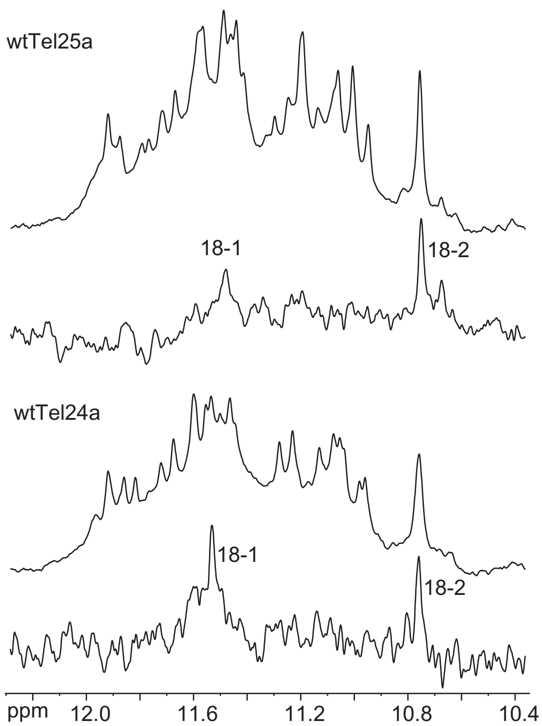

Figure 9. The imino proton region of the 1D ^{1}H NMR spectrum of wtTel25a (Figure1B) and the 1D ^{15}N-filtered spectrum of wtTel25a with site-specific labeling at G18 (top), and of wtTel24a (Figure1B) and its site-specific labeling at G18 (bottom). The G18 imino peaks for Hybrid-1 and Hybrid-2 structures are labeled with 18-1 and 18-2, respectively.

DISCUSSION

Capping structures determine the selective formation of Hybrid-1 or Hybrid-2 telomeric G-quadruplex structure

Here we have determined the NMR-refined molecular structure of the major intramolecular telomeric G-quadruplex formed in the native 26-nt human telomeric sequence wtTel26 (Figure 1A) in K^+ solution, which was a Hybrid-2 type telomeric G-quadruplex (Figure 1C, right). We have very recently determined the NMR structure of the intramolecular G-quadruplex formed in a modified 26-nt telomeric sequence Tel26 (Figure 1A) in K^+, which was a Hybrid-1 type telomeric G-quadruplex (Figure 1C, left) (38,39). The two sequences contain the same 22-nt core with four G-tracts and differ only in the flanking segments, with Tel26 containing the modified flanking-AA at both ends instead of the native TT (Figure 1A). It thus appears that while human telomeric sequences form predominantly hybrid-type G-quadruplex structures in K^+ solution, the flanking sequences determine the specific structure, namely, Hybrid-1 or Hybrid-2, formed.

While both telomeric intramolecular structures are hybrid type with mixed parallel/antiparallel strands, they differ in loop arrangements, strand orientations and capping structures (Figure 1C). The Hybrid-1 structure has sequential side-lateral-lateral loops with the first TTA loop adopting the double-chain-reversal conformation, whereas the Hybrid-2 structure has lateral-lateral-side loops with the last TTA loop adopting the double-chain-reversal conformation. Consequently, in the Hybrid-1 intramolecular telomeric G-quadruplex structure, the top end was covered by the 5′-flanking segment and the third TTA lateral loop, while the bottom end was covered by the second TTA lateral loop and the 3′-flanking segment. In contrast, the top end of the Hybrid-2 intramolecular telomeric G-quadruplex was covered by the 5′-flanking segment and the second TTA lateral loop, while the bottom end was covered by the first TTA lateral loop and the 3′-flanking segment (Figure 1C).

Our study indicated that the specific hybrid-type intramolecular telomeric G-quadruplex structure was selectively stabilized by specific capping structures. In the Hybrid-2 structure of the native wtTel26 sequence, a novel stable T:A:T triple capping structure was formed at the bottom end by T8 and A9 of the first TTA lateral loop and T25 of the 3′-flanking segment (Figures 5 and 6A and B). This T8:A9:T25 capping structure appears to play an important role for the stabilization of the Hybrid-2 structure, as demonstrated by mutational analysis and sequence screening. The T8:A9:T25 triple structure is specific to the Hybrid-2 folding structure (Figure 1C, right) and thus selectively favors the Hybrid-2 structure. This T:A:T triple is not able to form in the modified Tel26 sequence, as T25 was mutated to A25 (Figure 1A and C). In fact, the mutated A25 in the Tel26 sequence was shown to be involved in the formation of a different capping structure, the A25:T14 base pair, which is specific to the Hybrid-1 structure (Figure 1C, left) (39). It has been found in our previous study that the mutated A25 is important for the stable formation of the Hybrid-1 structure (38). In addition, an adenine triple capping structure was found to form in the Hybrid-1 structure (39), which provides additional stabilization specific to the Hybrid-1 structure.

The adenine of the TTA double-chain-reversal loop adopts similar conformations in Hybrid-1 and Hybrid-2 telomeric G-quadruplex structures

The third TTA linker adopts the double-chain-reversal loop conformation in the Hybrid-2 telomeric G-quadruplex structure, while the first TTA linker in the Hybrid-1 structure adopts the double-chain-reversal loop conformation (Figure 1C) (38). Similar to the observation in our previous Hybrid-1 structure that A9 of the double-chain-reversal loop is positioned partially above the top tetrad (39), A21 of the double-chain-reversal loop in the Hybrid-2 structure is also found to position partially above the top G-tetrad (Figure 6C). The double-chain-reversal loop conformations have been shown to favor short loop sizes of 1 and 2 nt (45,46,48–51), while a loop longer than 2 nt is in general not as favored for the formation of such a loop conformation, likely due to the lack of stacking interactions of loop residues in the groove region. The specific conformation observed in both Hybrid-1 and Hybrid-2 structures makes the TTA double-chain-reversal loop equivalent to a two-nt sized loop, as only the two thymine residues are positioned in the groove region. This may explain why a 3-nt double-chain-reversal loop can be stably present in the hybrid-type human telomeric G-quadruplexes in K^+.

Human telomeric sequences form a mixture of Hybrid-1 and Hybrid-2 structures, the population of each conformation largely depends on the 3′-flanking sequences

A four-G-tract human telomeric sequence is required for the formation of an intramolecular telomeric G-quadruplex. To gain insight into structure polymorphism of human telomeric DNA, we have systematically examined the four-G-tract human telomeric sequences (Figure 1A and B). We found that the G-quadruplexes formed in these human telomeric sequences in K^+ solution are always in equilibrium between Hybrid-1 and Hybrid-2 conformations. The equilibrium of Hybrid-1 and Hybrid-2 structures appears to be largely determined by the 3′-flanking sequence.

In the wtTel26 sequence with a 3′-flanking-TT (Figure 1B), the Hybrid-2 structure is the major form in K^+ solution (~75%) (Figure 2, top). The same major conformation, the Hybrid-2 structure, appears to form in wtTel25b and wtTel24b (Figure 1B) with the same 3′-flanking sequence but different 5′-flanking sequences (TA and A, respectively) (Figure 8). The Hybrid-2 structure also appears to be the major conformation in wtTel27 (Figure 1B) with a 3′-flanking-TTA segment that contains an additional 3′-terminal adenine (Figures 8 and S5). In contrast, telomeric sequences wtTel23 and wtTel24 (Figure 1B), which lack the 3′-flanking segment, appear to form the Hybrid-1 structure as the major conformation (Figure 8) (42). This result indicates that the 3′-flanking-T (T25 in our structure), which is necessary for the formation of the T:A:T triple capping structure (Figure 6A and B), is important for the selective formation of Hybrid-2 type intramolecular telomeric G-quadruplex structure. The telomeric sequences lacking this thymine are unable to form the T:A:T triple capping structure and thus a stable Hybrid-2 structure.

While the 1D NMR spectrum of wtTel25a with a 3′-flanking-T (Figure 1B) appears to be different from that of wtTel26 (Figure 8), we found that the Hybrid-2 form is still the major conformation in wtTel25a, although the Hybrid-1 structure can also be detected as a minor conformation (Figures 9 and S6). The T:A:T triple capping needed for the Hybrid-2 structure can be formed in wtTel25a; however, it appears that the second thymine (T26) of the 3′-flanking-TT in wtTel26 (Figure 1B) may provide additional stabilization for the T:A:T capping as well as the Hybrid-2 structure (Figure 5). Interestingly, we found that wtTel24a (Figure 1B) forms a mixture of Hybrid-1 and Hybrid-2 structures with comparable populations (Figure 9), although the 1D spectrum of WtTel24a looks quite similar to that of wtTel25a (Figure 8). WTel24a differs from wtTel25a only in its 5′-flanking segment, which is a TA instead of TTA (Figure 1B). Whereas wtTel24a contains the 3′-flanking-T that can form the T:A:T triple capping to stabilize the Hybrid-2 structure, it has been shown that the 5′-terminal thymine of the 5′-flanking TA can fold back and form an base-paired capping structure with A21 to stabilize the Hybrid-1 conformation in a modified mutTel24 sequence (Figure 1A) (41). The observation of the Hybrid-2 structure being clearly favored over Hybrid-1 in wtTel25a which contains an additional 5′-T (Figure 1B) suggests that this A:T base-paired capping structure involving the terminal thymine (41) is unlikely to form in an extended telomeric sequence. It also needs to be noted that the prediction of G-quadruplex folding topology based on 1D NMR spectral patterns should be made with caution.

Insight into structure polymorphism of G-quadruplexes formed in human telomere in K^+ solution

It appears that Hybrid-1 and Hybrid-2 conformations coexist in four-G-tract human telomeric sequences in K^+ solution. The energy difference between the two hybrid-type core G-quadruplex conformations appears to be rather small, as the equilibrium of Hybrid-1 and Hybrid-2 structures can be readily shifted by minor changes, such as the lengths or the minor modifications of the flanking sequences. However, the kinetics of the interconversion between the two conformations appears to be slow on the NMR timescale, as very few exchange peaks were observed in NOESY experiments. It may be noted that while the Hybrid-2 form appears to be the major conformation in the elongated sequences (such as wtTel26 and wtTel27), it accounts for only ~70% in wtTel26 and less than 65% in wtTel27, and the Hybrid-1 form can be detected to some extent in both sequences. It is thus suggested that both forms can form and coexist in the extended human telomeric sequence. Due to the low energy barrier between the two forms, the *in vivo* equilibrium of the two forms may be affected and readily shifted by different factors, such as body temperatures, ion concentrations and protein bindings.

The structure polymorphism and dynamic equilibrium of the human telomeric sequence appear to be intrinsic to its sequence and may be important for the biology of human telomeres. Unlike the G-rich sequences in gene promoter regions (45,46,48–55), the telomeric DNA sequences are unique in that they contain the same tandem repeats with the same linker segments. Thus in a telomeric sequence, any four-G-tract region has the same potential of forming an intramolecular G-quadruplex structure. The linker segment in the human (and vertebrates) telomeric sequence is TTA, whereas the linker segments in the telomeric sequences of most lower organisms only contain thymines. The presence of adenine in the TTA loops adds an asymmetry in human telomeric sequences, and as a consequence, provides a more selective basis for different capping structures and thus different G-quadruplex conformations. For example, the T:A:T triple capping with T8, A9 and T25 can only be formed in the Hybrid-2 structure, while the adenine triple capping with A3, A9 and A21 can only be formed in the Hybrid-1 structure (Figure 1C). It is interesting to note that the hybrid-type G-quadruplex is an asymmetric structure; and it is this asymmetry that determines the possibility of forming the two very closely related but yet distinct intramolecular G-quadruplexes, the Hybrid-1 and Hybrid-2 structures.

Significantly, as discussed in our previous report (38,39), both hybrid-type G-quadruplex structures provide an efficient scaffold for a compact-stacking structure of multimers in human telomeric DNA. The 5′- and 3′-ends of the hybrid-type G-quadruplex structure point in opposite directions, allowing the hybrid-type G-quadruplex to be readily folded and stacked end to end in the elongated linear telomeric DNA strand (Figure 1D). Intriguingly, the capping structures observed, such as the adenine triple in the Hybrid-1 structure and the T:A:T triple in Hybrid-2 structure, can provide not only stacking interactions between the two adjacent telomeric G-quadruplexes, but also specific binding sites for small molecule ligands. As the energy barrier between the two hybrid-type structures is rather small, it can be easily surpassed by a specific ligand binding, which can change the G-quadruplex conformation. Consequently, the different G-quadruplex structure may disrupt the existing protein interactions and introduce new protein recognitions. The human telomeric sequence has been shown to be highly conserved. Nature may have chosen this specific sequence with its asymmetry and the low energy barrier between different forms, which may provide a means to affect the protein recognition and to control the biology of human telomeres.

In summary, we have determined the molecular structure of the major intramolecular G-quadruplex formed in the native, non-modified 26-nt human telomeric sequence wtTel26 in physiological K^+ solution. Our study revealed a hybrid-type intramolecular G-quadruplex structure (Hybrid-2), which is different from the hybrid-type (Hybrid-1) structure that has been recently determined in our lab (38,39), in loop arrangements, strand orientations and capping structures. The distinct capping structures appear to determine the favored formation of a specific hybrid-type telomeric G-quadruplex structure, and may provide specific binding sites for drug targeting. The hybrid-type G-quadruplex structures appear to be the major conformations formed in four-G-tract human telomeric sequences in K^+ solution, with a dynamic equilibrium between Hybrid-1 and Hybrid-2 conformations that appears to be largely determined by the 3′-flanking sequence. Both hybrid-type G-quadruplex structures suggest a straightforward means for multimer formation with effective packing in the human telomeric sequence and provide important implications for drug targeting of G-quadruplexes in human telomeres.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

This result was presented at the "First International Meeting on Quadruplex DNA" held in Louisville, KY, USA on April 21–24, 2007 (http://pubs.acs.org/cen/coverstory/85/8522aboutcover.html). We thank Dr Laurence Hurley for insightful comments and suggestions. This research was supported by the National Institutes of Health (1K01CA83886 and 1S10 RR16659). Funding to pay the Open Access publication charges for this article was waived by Oxford University Press.

Conflict of interest statement. None declared.

REFERENCES

1. Mergny,J.L. and Helene,C. (1998) G-quadruplex DNA: a target for drug design. *Nat. Med.*, **4**, 1366–1367.
2. Sun,D.Y. and Hurley,L.H. (2001) In Chaires,J.B. and Waring,M.J. (eds), *Methods In Enzymology, Drug-Nucleic Acid Interactions*, Vol. 340, Academic Press Inc, San Diego, 573–592.
3. Hurley,L.H. (2002) DNA and its associated processes as targets for cancer therapy. *Nat. Rev. Cancer*, **2**, 188–200.
4. Neidle,S. and Parkinson,G. (2002) Telomere maintenance as a target for anticancer drug discovery. *Nat. Rev. Drug Discov.*, **1**, 383–393.
5. Bodnar,A.G., Ouellette,M., Frolkis,M., Holt,S.E., Chiu,C.P., Morin,G.B., Harley,C.B., Shay,J.W., Lichtsteiner,S. *et al.* (1998) Extension of life-span by introduction of telomerase into normal human cells. *Science*, **279**, 349–352.
6. Harley,C.B. (1991) Telomere loss: mitotic clock or genetic time bomb? *Mutat. Res.*, **256**, 271–282.
7. Sun,H., Karow,J.K., Hickson,I.D. and Maizels,N. (1998) The bloom's syndrome helicase unwinds G4 DNA. *J. Biol. Chem.*, **273**, 27587–27592.
8. Sun,H., Bennett,R.J. and Maizels,N. (1999) The Saccharomyces cerevisiae Sgs1 helicase efficiently unwinds G-G paired DNAs. *Nucleic Acids Res.*, **27**, 1978–1984.
9. Hackett,J.A., Feldser,D.M. and Greider,C.W. (2001) Telomere dysfunction increases mutation rate and genomic instability. *Cell*, **106**, 275–286.
10. Blackburn,E.H. and Gall,J.G. (1978) A tandemly repeated sequence at the termini of extrachromosomal ribosomal RNA genes in Tetrahymena. *J. Mol. Biol.*, **120**, 33–53.
11. Allshire,R.C., Dempster,M. and Hastie,N.D. (1989) Human telomeres contain at least 3 types of G-rich repeat distributed non-randomly. *Nucleic Acids Res.*, **17**, 4611–4627.
12. de Lange,T., Shiue,L., Myers,R.M., Cox,D.R., Naylor,S.L., Killery,A.M. and Varmus,H.E. (1990) Structure and variability of human chromosome ends. *Mol. Cell. Biol.*, **10**, 518–527.
13. Moyzis,R.K., Buckingham,J.M., Cram,L.S., Dani,M., Deaven,L.L., Jones,M.D., Meyne,J., Ratliff,R.L. and Wu,J.R. (1988) A highly conserved repetitive DNA sequence, $(TTAGGG)_n$, present at the telomeres of human chromosomes. *Proc. Natl Acad. Sci. USA*, **85**, 6622–6626.
14. Makarov,V.L., Hirose,Y. and Langmore,J.P. (1997) Long G tails at both ends of human chromosomes suggest a C strand degradation mechanism for telomere shortening. *Cell*, **88**, 657–666.
15. Harley,C.B., Futcher,A.B. and Greider,C.W. (1990) Telomeres shorten during ageing of human fibroblasts. *Nature*, **345**, 458–460.
16. Greider,C.W. and Blackburn,E.H. (1985) Identification of a specific telomere terminal transferase activity in Tetrahymena extracts. *Cell*, **43**, 405–413.
17. Kim,N.W., Piatyszek,M.A., Prowse,K.R., Harley,C.B., West,M.D., Ho,P.L., Coviello,G.M., Wright,W.E., Weinrich,S.L. *et al.* (1994) Specific association of human telomerase activity with immortal cells and cancer. *Science*, **266**, 2011–2015.
18. Hanahan,D. and Weinberg,R.A. (2000) The hallmarks of cancer. *Cell*, **100**, 57–70.
19. Hurley,L.H. (2001) Secondary DNA structures as molecular targets for cancer therapeutics. *Biochem. Soc. Trans.*, **29**, 692–696.
20. Hurley,L.H., Wheelhouse,R.T., Sun,D., Kerwin,S.M., Salazar,M., Fedoroff,O.Y., Han,F.X., Han,H.Y., Izbicka,E. *et al.* (2000) G-quadruplexes as targets for drug design. *Pharmacol. Ther.*, **85**, 141–158.
21. Neidle,S. and Read,M.A. (2000) G-quadruplexes as therapeutic targets. *Biopolymers*, **56**, 195–208.
22. Parkinson,G.N., Lee,M.P.H. and Neidle,S. (2002) Crystal structure of parallel quadruplexes from human telomeric DNA. *Nature*, **417**, 876–880.

23. Balagurumoorthy,P. and Brahmachari,S.K. (1994) Structure and stability of human telomeric sequence. *J. Biol. Chem.*, **269**, 21858–21869.
24. Ying,L.M., Green,J.J., Li,H.T., Klenerman,D. and Balasubramanian,S. (2003) Studies on the structure and dynamics of the human telomeric G quadruplex by single-molecule fluorescence resonance energy transfer. *Proc. Natl Acad. Sci. USA*, **100**, 14629–14634.
25. Redon,S., Bombard,S., Elizondo-Riojas,M.A. and Chottard,J.C. (2003) Platinum cross-linking of adenines and guanines on the quadruplex structures of the AG(3)(T(2)AG(3))(3) and (T(2)AG(3))(4) human telomere sequences in Na + and K + solutions. *Nucleic Acids Res.*, **31**, 1605–1613.
26. Phan,A.T. and Patel,D.J. (2003) Two-repeat human telomeric d(TAGGGTTAGGGT) sequence forms interconverting parallel and antiparallel G-quadruplexes in solution: distinct topologies, thermodynamic properties, and folding/unfolding kinetics. *J. Am. Chem. Soc.*, **125**, 15021–15027.
27. He,Y.J., Neumann,R.D. and Panyutin,I.G. (2004) Intramolecular quadruplex conformation of human telomeric DNA assessed with I-125-radioprobing. *Nucleic Acids Res.*, **32**, 5359–5367.
28. Rezler,E.M., Seenisamy,J., Bashyam,S., Kim,M.Y., White,E., Wilson,W.D. and Hurley,L.H. (2005) Telomestatin and diseleno sapphyrin bind selectively to two different forms of the human telomeric G-quadruplex structure. *J. Am. Chem. Soc.*, **127**, 9439–9447.
29. Li,J., Correia,J.J., Wang,L., Trent,J.O. and Chaires,J.B. (2005) Not so crystal clear: the structure of the human telomere G-quadruplex in solution differs from that present in a crystal. *Nucleic Acids Res.*, **33**, 4649–4659.
30. Wlodarczyk,A., Grzybowski,P., Patkowski,A. and Dobek,A. (2005) Effect of ions on the polymorphism, effective charge, and stability of human telomeric DNA. Photon correlation spectroscopy and circular dichroism studies. *J. Phys. Chem. B*, **109**, 3594–3605.
31. Qi,J.Y. and Shafer,R.H. (2005) Covalent ligation studies on the human telomere quadruplex. *Nucleic Acids Res.*, **33**, 3185–3192.
32. Vorlickova,M., Chladkova,J., Kejnovska,I., Fialova,M. and Kypr,J. (2005) Guanine tetraplex topology of human telomere DNA is governed by the number of (TTAGGG) repeats. *Nucleic Acids Res.*, **33**, 5851–5860.
33. Rujan,I.N., Meleney,J.C. and Bolton,P.H. (2005) Vertebrate telomere repeat DNAs favor external loop propeller quadruplex structures in the presence of high concentrations of potassium. *Nucleic Acids Res.*, **33**, 2022–2031.
34. Risitano,A. and Fox,K.R. (2005) Inosine substitutions demonstrate that intramolecular DNA quadruplexes adopt different conformations in the presence of sodium and potassium. *Bioorg. Med. Chem. Lett.*, **15**, 2047–2050.
35. Ourliac-Garnier,I., Elizondo-Riojas,M.A., Redon,S., Farrell,N.P. and Bombard,S. (2005) Cross-links of quadruplex structures from human telomeric DNA by dinuclear platinum complexes show the flexibility of both structures. *Biochemistry*, **44**, 10620–10634.
36. Lee,J.Y., Okumus,B., Kim,D.S. and Ha,T.J. (2005) Extreme conformational diversity in human telomeric DNA. *Proc. Natl Acad. Sci. USA*, **102**, 18938–18943.
37. Wang,Y. and Patel,D.J. (1993) Solution structure of the human telomeric repeat D[Ag(3)(T(2)Ag(3))3] G-tetraplex. *Structure*, **1**, 263–282.
38. Ambrus,A., Chen,D., Dai,J.X., Bialis,T., Jones,R.A. and Yang,D.Z. (2006) Human telomeric sequence forms a hybrid-type intramolecular G-quadruplex structure with mixed parallel/antiparallel strands in potassium solution. *Nucleic Acids Res.*, **34**, 2723–2735.
39. Dai,J., Punchihewa,C., Ambrus,A., Chen,D., Jones,R.A. and Yang,D.Z. (2007) Structure of the intramolecular human telomeric G-quadruplex in potassium solution: a novel adenine triple formation. *Nucleic Acids Res.*, **35**, 2440–2450.
40. Xu,Y., Noguchi,Y. and Sugiyama,H. (2006) The new models of the human telomere d[AGGG(TTAGGG)(3)] in K + solution. *Bioorg. Med. Chem.*, **14**, 5584–5591.
41. Luu,K.N., Phan,A.T., Kuryavyi,V., Lacroix,L. and Patel,D.J. (2006) Structure of the human telomere in K + solution: an intramolecular (3 + 1) G-quadruplex scaffold. *J. Am. Chem. Soc.*, **128**, 9963–9970.
42. Phan,A.T., Luu,K.N. and Patel,D.J. (2006) Different loop arrangements of intramolecular human telomeric (3 + 1) G-quadruplexes in K + solution. *Nucleic Acids Res.*, **34**, 5715–5719.
43. Dai,J.X., Chen,D., Jones,R.A., Hurley,L.H. and Yang,D.Z. (2006) NMR solution structure of the major G-quadruplex structure formed in the human BCL2 promoter region. *Nucleic Acids Res.*, **34**, 5133–5144.
44. Dai,J.X., Punchihewa,C., Mistry,P., Ooi,A.T. and Yang,D.Z. (2004) Novel DNA bis-intercalation by MLN944, a potent clinical bisphenazine anticancer drug. *J. Biol. Chem.*, **279**, 46096–46103.
45. Ambrus,A., Chen,D., Dai,J.X., Jones,R.A. and Yang,D.Z. (2005) Solution structure of the biologically relevant g-quadruplex element in the human c-MYC promoter. Implications for G-quadruplex stabilization. *Biochemistry*, **44**, 2048–2058.
46. Dai,J., Dexheimer,T.S., Chen,D., Carver,M., Ambrus,A., Jones,R.A. and Yang,D.Z. (2006) 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.*, **128**, 1096–1098.
47. Brünger,A.T. (1993) *X-PLOR Version 3.1: A system for X-ray crystallography and NMR*. Yale University Press, New Haven, CT.
48. Seenisamy,J., Rezler,E.M., Powell,T.J., Tye,D., Gokhale,V., Joshi,C.S., Siddiqui-Jain,A. and Hurley,L.H. (2004) The dynamic character of the G-quadruplex element in the c-MYC promoter and modification by TMPyP4. *J. Am. Chem. Soc.*, **126**, 8702–8709.
49. Phan,A.T., Modi,Y.S. and Patel,D.J. (2004) Propeller-type parallel-stranded G-quadruplexes in the human c-myc promoter. *J. Am. Chem. Soc.*, **126**, 8710–8716.
50. Sun,D.Y., Guo,K.X., Rusche,J.J. and Hurley,L.H. (2005) Facilitation of a structural transition in the polypurine/polypyrimidine tract within the proximal promoter region of the human VEGF gene by the presence of potassium and G-quadruplex-interactive agents. *Nucleic Acids Res.*, **33**, 6070–6080.
51. Phan,A.T., Kuryavyi,V., Burge,S., Neidle,S. and Patel,D.J. (2007) Structure of an unprecedented G-quadruplex scaffold in the human c-kit promoter. *J. Am. Chem. Soc.*, **129**, 4386–4392.
52. De Armond,R., Wood,S., Sun,D.Y., Hurley,L.H. and Ebbinghaus,S.W. (2005) Evidence for the presence of a guanine quadruplex forming region within a polypurine tract of the hypoxia inducible factor 1 alpha promoter. *Biochemistry*, **44**, 16341–16350.
53. Rankin,S., Reszka,A.P., Huppert,J., Zloh,M., Parkinson,G.N., Todd,A.K., Ladame,S., Balasubramanian,S. and Neidle,S. (2005) Putative DNA quadruplex formation within the human c-kit oncogene. *J. Am. Chem. Soc.*, **127**, 10584–10589.
54. Hurley,L.H., Von Hoff,D.D., Siddiqui-Jain,A. and Yang,D.Z. (2006) Drug targeting of the c-MYC promoter to repress gene expression via a G-quadruplex silencer element. *Semin. Oncol.*, **33**, 498–512.
55. Yang,D.Z. and Hurley,L.H. (2006) Structure of the biologically relevant G-quadruplex in the c-MYC promoter. *Nucleosides Nucleotides Nucleic Acids*, **25**, 951–968.
56. Nicholls,A., Sharp,K.A. and Honig,B. (1991) Protein folding and association – insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins:Structure Function and Genetics*, **11**, 281–296.