Published online 25 September 2007 *Nucleic Acids Research, 2007, Vol. 35, No. 19* **6517–6525**
doi:10.1093/nar/gkm706

Structure of two intramolecular G-quadruplexes formed by natural human telomere sequences in K^+ solution†

Anh Tuân Phan[1,2,*], Vitaly Kuryavyi[1], Kim Ngoc Luu[1] and Dinshaw J. Patel[1]

[1]Structural Biology Program, Memorial Sloan-Kettering Cancer Center, New York, NY 10021, USA and [2]Division of Physics and Applied Physics, School of Physical and Mathematical Sciences, Nanyang Technological University, Singapore 637551, Singapore

Received July 7, 2007; Revised and Accepted August 24, 2007 Accession Nos PDB2JSK, 2JSL, 2JSM, 2JSQ

ABSTRACT

Intramolecular G-quadruplexes formed by human telomere sequences are attractive anticancer targets. Recently, four-repeat human telomere sequences have been shown to form two different intramolecular (3 + 1) G-quadruplexes in K^+ solution (Form 1 and Form 2). Here we report on the solution structures of both Form 1 and Form 2 adopted by natural human telomere sequences. Both structures contain the (3 + 1) G-tetrad core with one double-chain-reversal and two edgewise loops, but differ in the successive order of loop arrangements within the G-quadruplex scaffold. Our results provide the structural details at the two ends of the G-tetrad core in the context of natural sequences and information on different loop conformations. This structural information might be important for our understanding of telomere G-quadruplex structures and for anticancer drug design targeted to such scaffolds.

INTRODUCTION

Guanine-rich DNA sequences can form G-quadruplex structures *in vitro* through stacking of planar G·G·G·G tetrads (1–5). DNA at the ends (telomeres) of eukaryotic chromosomes consists of tandem repeats of G-rich sequences, such as $(GGGTTA)_n$ in humans (6). The *in vitro* (7,8) and *in vivo* (9–11) observations of G-quadruplex formation in telomeric sequences and telomeres, respectively, show the biological importance of this DNA scaffold. G-quadruplexes formed by human telomere sequences are promising anticancer targets (12–16), because formation of such structures inhibits the activity of telomerase (17–19), an enzyme (20) that is required for the proliferation of 80%–85% of cancer cells (21).

In 1993, our group characterized the first NMR-based solution structure of a four-repeat human telomere sequence, $d[AGGG(TTAGGG)_3]$ in Na^+ solution (7). This sequence forms an intramolecular G-quadruplex involving three stacked G-tetrads with *anti·anti·syn·syn* glycosidic conformations around each tetrad. Three connecting TTA loops adopt successive edgewise, diagonal and edgewise alignments, such that each strand has both parallel and antiparallel adjacent strands (Figure 1A). In 2002, a very different G-quadruplex structure of the same sequence was observed in a K^+-containing crystal by the Stephen Neidle group (8). In this structure, all four strands are parallel, the connecting TTA loops are double-chain-reversal, and all guanines adopt *anti* glycosidic conformations (Figure 1B). Subsequent studies from many laboratories indicated the presence of a mixture of multiple G-quadruplex forms for human telomere sequences in physiological K^+ solution conditions (22–40).

In 2005, our group showed that three-repeat human telomere sequences form a bimolecular (3 + 1) quadruplex in Na^+ solution, whose core contains three strands oriented in one direction and the fourth in the opposite direction (41). This type of (3 + 1) G-quadruplex core was first reported by our group in 1994 for a *Tetrahymena* telomere G-quadruplex (42). In 2006, our group showed that four-repeat human telomere sequences form at least two intramolecular G-quadruplexes of the (3 + 1)-type (Form 1 and Form 2) in K^+ solution (43,44). Form 1 (Figure 1C) and Form 2 (Figure 1D) are the major conformations (~70%) of the natural human telomere 23-nt $d[TAGGG(TTAGGG)_3]$ and 25-nt $d[TAGGG(TTAGGG)_3TT]$ sequences in K^+ solution, respectively (43,44). Both forms contain one double-chain-reversal and two

*To whom correspondence should be addressed. Tel: +65 6514 1915; Fax: +65 6794 1325; Email: phantuan@ntu.edu.sg
Correspondence may also be addressed to Dinshaw J. Patel. Tel: +1 212 639 7207; Fax: +1 212 717 3066; Email: pateld@mskcc.org

The authors wish to be known that, in their opinion, the first two authors should be regarded as joint First Authors
†Much of this work was presented at the First International Quadruplex DNA Meeting, Louisville, KY, USA; April 2007

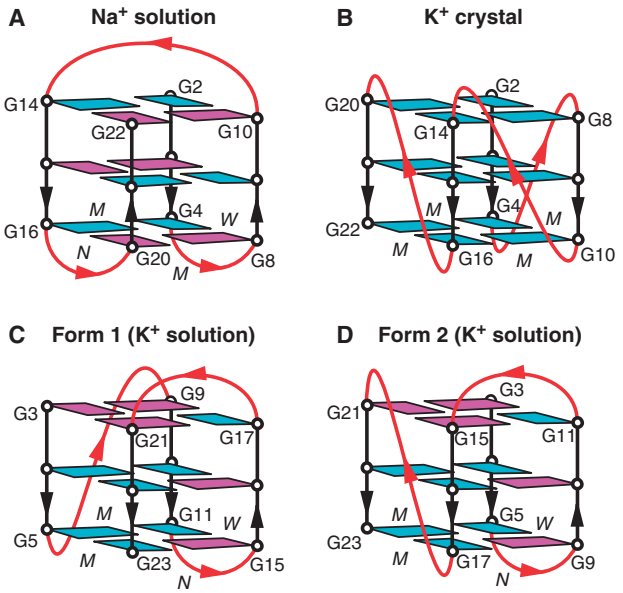

Figure 1. Schematic structures of intramolecular G-quadruplexes formed by the human telomeric sequences: (**A**) d[AGGG(TTAGGG)$_3$] in Na^+ solution; (**B**) d[AGGG(TTAGGG)$_3$] in a K^+-containing crystal; (**C**) d[TAGGG(TTAGGG)$_3$] in K^+ solution (natural-Form 1) and (**D**) d[TAGGG(TTAGGG)$_3$TT] in K^+ solution (natural-Form 2). Loops are colored red; *anti* and *syn* guanines are colored cyan and magenta, respectively. *W*, *M* and *N* denote wide, medium and narrow groove, respectively.

edgewise loops, but they differ in the successive order of loop arrangements: the double-chain-reversal loop is formed by the third TTA linker in Form 2 (Figure 1D) instead of the first TTA linker in Form 1 (Figure 1C). Formation of Form 1 human telomere quadruplex was independently proposed by two other research groups using end-modified (45) and multiple 8-bromoguanine-substituted (46) sequences, respectively. It appeared that some modifications at terminal residues favor certain loop conformations of the quadruplex due to their interaction with the loop residues (43–45), while multiple guanine-to-8-bromoguanine (G $\rightarrow$ ^{Br}G) substitutions would favor a quadruplex form, in which the substituted guanines are forced to adopt *syn* conformations (46–48).

By slightly modifying some flanking terminal residues, we could favor Form 1 (to ~95%) and determine its solution structure (43). Although this structure provided the first approximation to the 3D structure of an intramolecular human telomere G-quadruplex in K^+ solution (49), some structural information could be altered by incorporated modified bases at the ends of the molecule (43). Very recently, structures of Form 1 were also reported by Dai *et al.* (50) for a different end-modified sequence and by Matsugami *et al.* (51) for a sequence containing five ^{Br}G substitutions.

Here we report on the NMR-based solution structures of both Form 1 and Form 2 quadruplexes adopted by natural human telomere sequences. The structure determination was assisted by the study of sequences each containing judiciously chosen single ^{Br}G substitution (see below). Our results provide the structural details at the two ends of the G-tetrad core in the context of natural sequences and information on the range of conformations accessible to the TAA loops. This structural information might be important for the design of anticancer drugs targeted to human telomeric DNA.

METHODS

Sample preparation

The unlabeled and the site-specific low-enrichment (2% ^{15}N-labeled) oligonucleotides were synthesized and purified as described previously (52,53). Unless otherwise stated, the strand concentration of the NMR samples was typically 0.5–5 mM; the solutions contained 70 mM of KCl and 20 mM of potassium phosphate (pH 7). Sequences used in this work are shown below:

Name	Sequence
Natural-Form 1	d[TAGGGTTAGGGTTAGGGTTAGGG]
^{Br}G16-Form 1	d[TAGGGTTAGGGTTAG(^{Br}G)GTTAGGG]
Natural-Form 2	d[TAGGGTTAGGGTTAGGGTTAGGGTT]
^{Br}G15-Form 2	d[TAGGGTTAGGGTTA(^{Br}G)GGTTAGGGTT]
Two forms	d[TAGGGTTAGGGTTAGGGTTAGGGT]

NMR spectroscopy

Experiments were performed on 600 MHz spectrometers at 25°C, unless otherwise specified. Resonances for G residues were assigned unambiguously by using site-specific low-enrichment labeling and through-bond correlations at natural abundance (52–55). Resonances for T residues were assigned following systematic T-to-U replacements. Assignments of resonances for A residues were obtained from NOE connectivities with neighboring T and G residues in the fold. NMR spectral assignments were completed by through-bond (COSY, TOCSY) and through-space (NOESY) correlation experiments as described previously (54). Interproton distances were measured by using NOESY experiments at different mixing times.

Structure calculation

The structures of Form 1 and Form 2 human telomere quadruplexes were calculated using the X-PLOR program (56). NMR-restrained molecular dynamics computations were performed as described previously (43). The structures were first calculated for ^{Br}G-substituted sequences. Inclusion of K^+ ions within the top and bottom caps of Form 2 (57) resulted in well-converged structures, which are consistent with experimental data. The ensembles of structures for natural sequences were computed from those for ^{Br}G-substituted sequences by refining them against the experimental restraints obtained for natural sequences.

Data deposition

The coordinates for four quadruplex structures formed by the 23-nt and 25-nt natural and ^{Br}G-substituted human telomere sequences have been deposited in the Protein Data

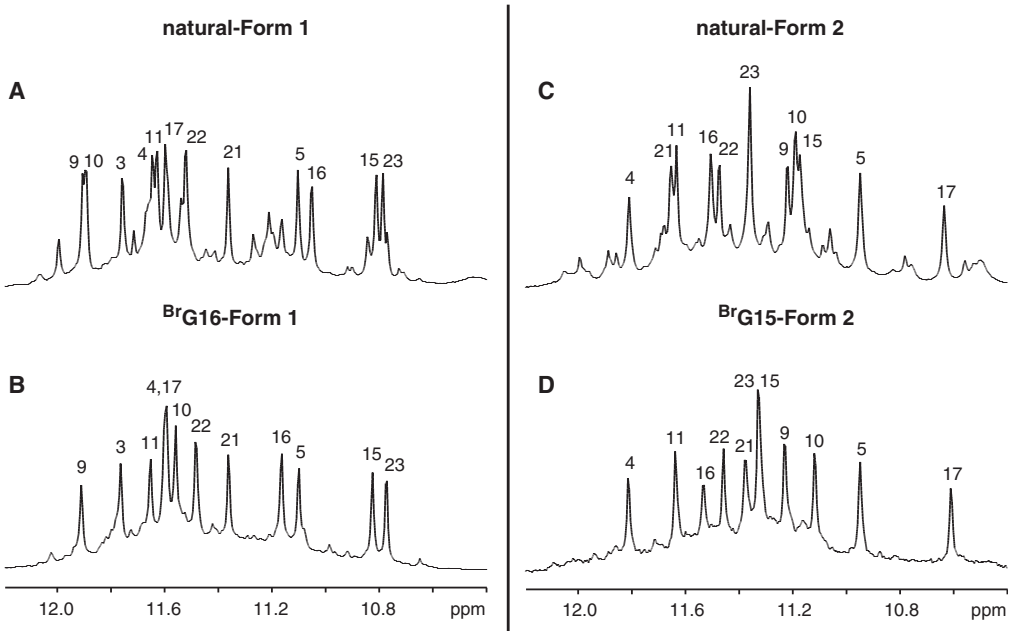

Figure 2. Imino proton spectra of (**A**) d[TAGGG(TTAGGG)$_3$] (natural-Form 1), (**B**) ^{Br}G16-Form 1, (**C**) d[TAGGG(TTAGGG)$_3$TT] (natural-Form 2) and (**D**) ^{Br}G15-Form 2, in K^+ solution with assignments listed over the spectra.

Bank (accession codes natural-Form 1: 2JSM; ^{Br}G16-Form 1: 2JSK; natural-Form 2: 2JSL; ^{Br}G15-Form 2: 2JSQ).

RESULTS AND DISCUSSION

Effects of DNA sequences on relative quadruplex populations and quality of NMR spectra

Previously, we systematically examined human telomere sequences containing four G-tracts and showed that small changes to flanking sequences can perturb the equilibrium between different coexisting G-quadruplex forms (44). In K^+ solution, d[TAGGG(TTAGGG)$_3$] (a natural human telomere sequence) forms up to 70% of Form 1 (43), while d[TAGGG(TTAGGG)$_3$TT] (also a natural human telomere sequence) with two Ts at 3′-end forms up to 70% of Form 2 (44). Here we will call these sequences natural-Form 1 and natural-Form 2, respectively. It should be noted that two G-quadruplex conformers coexist in slow exchange at comparable proportions in d[TAGGG(TTA GGG)$_3$T] (another natural human telomere sequence) having one T at 3′-end (Figure S1, Supplementary Data).

Substitution of proton by bromine at position C8 of a guanine has been shown to favor *syn* glycosidic conformation of the nucleotide (47). When all five guanines that adopted *syn* conformations in Form 1 were substituted by 8-bromoguanines, Form 1 predominated (46,51). However, each G-to-^{Br}G substitution removes a proton (useful in NMR studies). For the natural-Form 1 and natural-Form 2 sequences, we could identify single G-to-^{Br}G substitutions (at position G16 and G15, respectively) that further favored the corresponding major form, thereby significantly improving NMR spectra (Figure 2). They will be called ^{Br}G16-Form 1 and ^{Br}G15-Form 2, respectively.

NMR spectral assignments

We have previously unambiguously assigned imino and H8 protons of guanines in natural sequences (43,44). Corresponding assignments for single ^{Br}G-substituted sequences could be obtained by comparing spectral patterns of the modified and natural sequences (Figures 2 and 3). These assignments were also independently confirmed by some low-enrichment site-specific labeling and natural abundance through-bond correlation experiments (Figure 4) (52–55). Assignments of resonances for T residues were obtained from T-to-U substitution samples (54). NMR spectral assignments were completed by through-bond (COSY, TOCSY) and through-space (NOESY) correlation experiments as described previously (54).

We first assigned most peaks in NOESY spectra of ^{Br}G-substituted sequences. These NOE assignments helped us to assign NOEs for the natural sequences, which would have been very difficult to complete due to the presence of a significant amount of minor conformation(s).

Analysis of NOE patterns (Figure 3) suggested that the major forms of ^{Br}G-substituted sequences and those of the corresponding natural sequences are of the same general folds (Figure 1). The Form 1 and Form 2 folds for ^{Br}G-substituted sequences were supported by proton exchange data, which showed that imino protons of the central G-tetrad are the most protected from exchange with water (Figure 4).

Overall solution structure

The structures of Form 1 (Figures 5, 6 and S2, Table 1) and Form 2 (Figures 7, 8 and S3, Table 2) human telomere quadruplexes adopted by ^{Br}G-substituted and natural sequences were calculated on the basis of NMR restraints

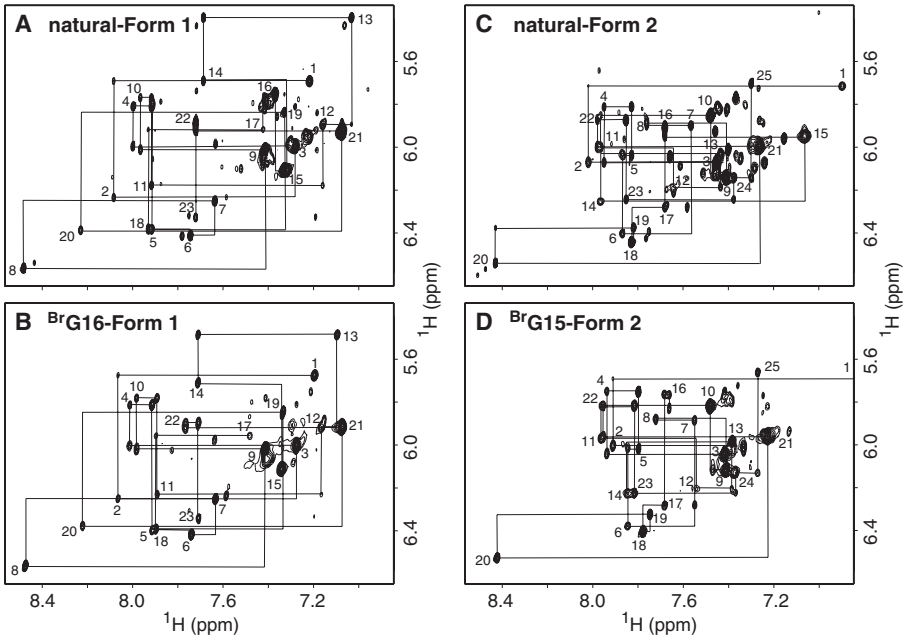

Figure 3. The H8/6-H1′ proton region of NOESY spectra (mixing time, 300 ms) of (**A**) d[TAGGG(TTAGGG)$_3$] (natural-Form 1), (**B**) ^{Br}G16-Form 1, (**C**) d[TAGGG(TTAGGG)$_3$TT] (natural-Form 2) and (**D**) ^{Br}G15-Form 2 in K^+ solution. The assignments and H8/6-H1′ NOE sequential connectivities are shown.

using the X-PLOR program (56). In all cases, the structure of the G-tetrad cores is better defined than that of the loops (Figures 5 and 7). The structures formed by the ^{Br}G-substituted and the corresponding natural sequences are quite similar. However, the structures calculated for the ^{Br}G-substituted sequences (Figures 5A and 7A) are slightly better defined than those calculated for the natural sequences (Figures 5C and 7C) thanks to larger numbers of defined NOE peaks associated with cleaner NMR spectra.

Both Form 1 and Form 2 contain the (3 + 1) G-quadruplex core, which is identified by one narrow, one wide and two medium grooves (Figure 1C and D) (41–43). The groove widths are defined mainly by the relative orientations of strands, but are also somewhat affected by the structures of the closing loops. For example, the narrow groove seems narrower in Form 2 than in Form 1 (Figures 5B and 7B, respectively). These grooves could serve as potential targets for small-molecule ligands.

Structure of loops and caps

In both Form 1 and Form 2, there are one double-chain-reversal and two edgewise TTA loops. The double-chain-reversal loop is situated in a medium groove. The edgewise loops always connect an *anti* guanine to a *syn* guanine, across narrow or wide grooves; they cap the top and the bottom of the G-tetrad core, respectively (Figure 1). The detailed structures of these elements in Form 1 and Form 2 are shown in Figures 6 and 8, respectively. Figures S2 and S3 show the distribution of sequential and long-range NOEs used to derive the loop structures in these forms.

In Form 1, the T18–T19–A20 loop (Figure 6A) closes a narrow groove and caps the top of the G-tetrad core. Residue A20 from this loop is aligned with residues T1 and A2 from the 5′-end to form the (T1–A2)·A20 triad platform (Figure 6A and D). Residue T19 is stacked on top of this platform, while T18 is projected aside. An adenine triple was observed by Dai *et al.* (50) in the top cap of Form 1 for an end-modified sequence, in which the natural residue T1 was replaced by an A. This base triple involved adenines which are equivalent to A2, A8 and A20 in our sequence. Such an adenine triple does not form in our structure of the natural human telomere sequence. The T12–T13–A14 loop (Figure 6C) closes a wide groove and caps the bottom of the G-tetrad core. Hoogsteen base pair A14–T12 was observed among computed structures and stacks over the terminal G-tetrad (Figures 5 and 6C). This configuration of the bottom loop is clearly different from that observed for Form 1 with modified bases at the 3′-end (43,50). The structure of the T6–T7–A8 double-chain-reversal loop shows the stacking between T7 and A8 (Figure 6B). We observe differences in the conformations of both edgewise TTA loops between our Form 1 solution structure (Figures 6A and C) and the corresponding solution structure reported by Matsugami *et al.* (51). In our case, there is much greater stacking of the loop residues over the terminal G-tetrads.

In Form 2, the T12–T13–A14 loop (Figure 8A) closes a narrow groove and caps the top of the G-tetrad core. The residues in this loop interact intimately with the 5′-end residues T1–A2 (Figure 8A and D). NOEs detected within this region are consistent with a structure where T1, T12 and A14 residues, together with the carbonyl groups of the top G-tetrad, can potentially coordinate a K^+ ion within the loop (Figure S4). This structure would explain the upfield chemical shifts of many protons of T1

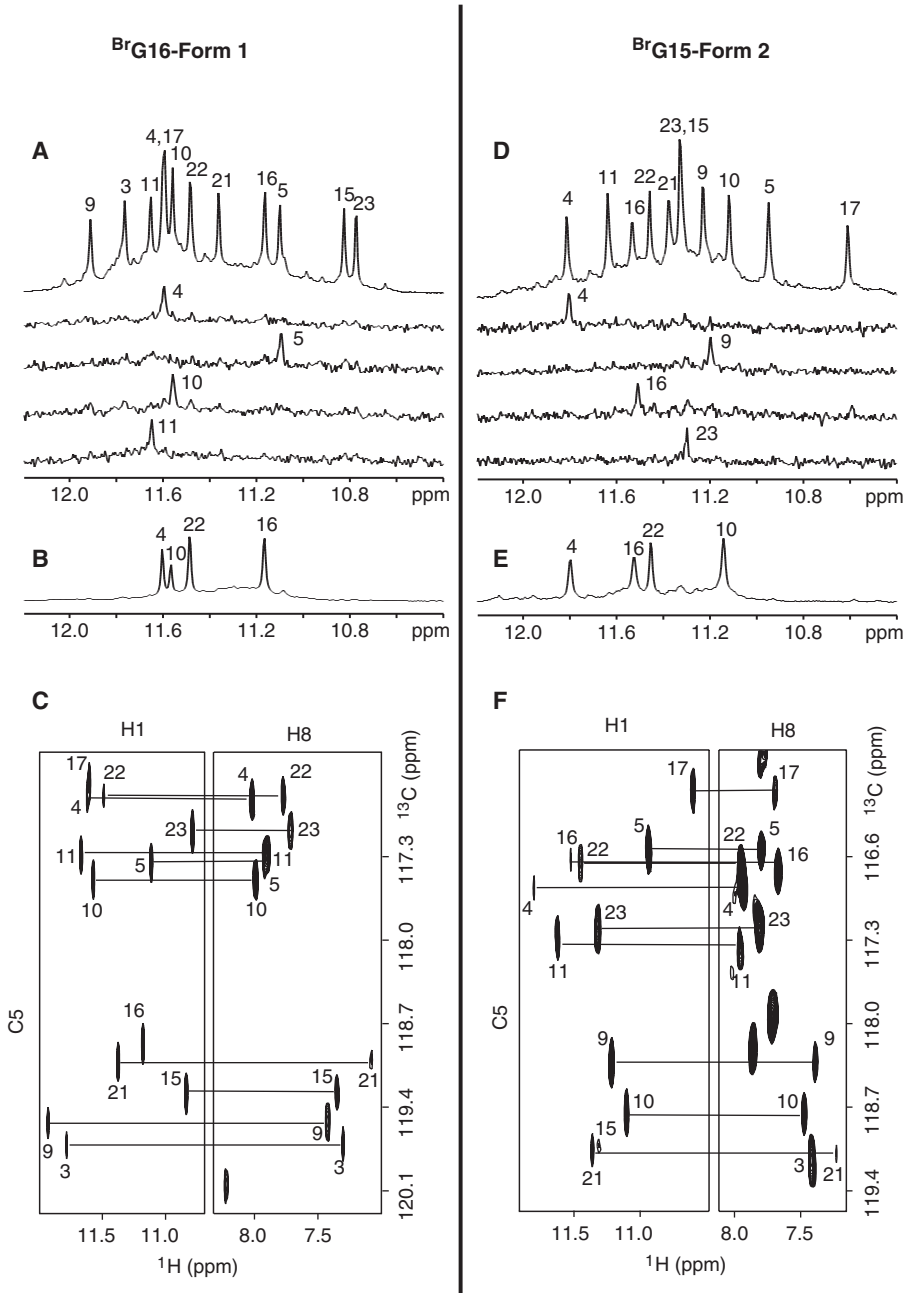

Figure 4. Imino proton spectra and assignments of (**A** and **B**) the ^{Br}G16-Form 1 and (**D** and **E**) ^{Br}G15-Form 2 human telomere sequences in K^+ solution. (A and D) Reference guanine imino proton spectra (reference) and some examples of imino protons assignments by ^{15}N-filtered spectra recorded for samples, 2% ^{15}N-labeled at the indicated positions. (B and E) Imino proton spectra after 1 h in D_2O at 25°C. (**C** and **F**) H8 proton assignments of (C) the ^{Br}G16-Form 1 and (F) ^{Br}G15-Form 2 human telomere sequences by through-bond correlations between imino and H8 protons via ^{13}C5 at natural abundance.

(Figure 3 and Table S2). However, other structures might also be possible for this region, as manifested by the broadening of some resonances at 25°C (44). For example, in some stages of computation we observed the interaction of A20 from the double-chain-reversal loop with residues in the top cap. The structure of the T18–T19–A20 double-chain-reversal loop shows some stacking between T18 and T19 bases (Figure 8B). The T6–T7–A8 loop (Figure 8C) closes a wide groove and caps the bottom of the G-tetrad core. Residues T7 and A8 interact with T24 from the 3′-end to form a (T7–A8)·T24 triad platform. Residue T25 is stacked below this platform. The configuration of the bottom cap in Form 2 involving the interaction between the loop and the 3′-end residues is quite different from that observed in Form 1 [see above and Refs. (43,50,51)]. A K^+ ion can also be potentially coordinated between the bottom G-tetrad and the cap (Figure S4).

In general, K^+ can be coordinated within all edgewise loops presented here, and the equilibrium between base pairings and K^+ coordination is probably the best

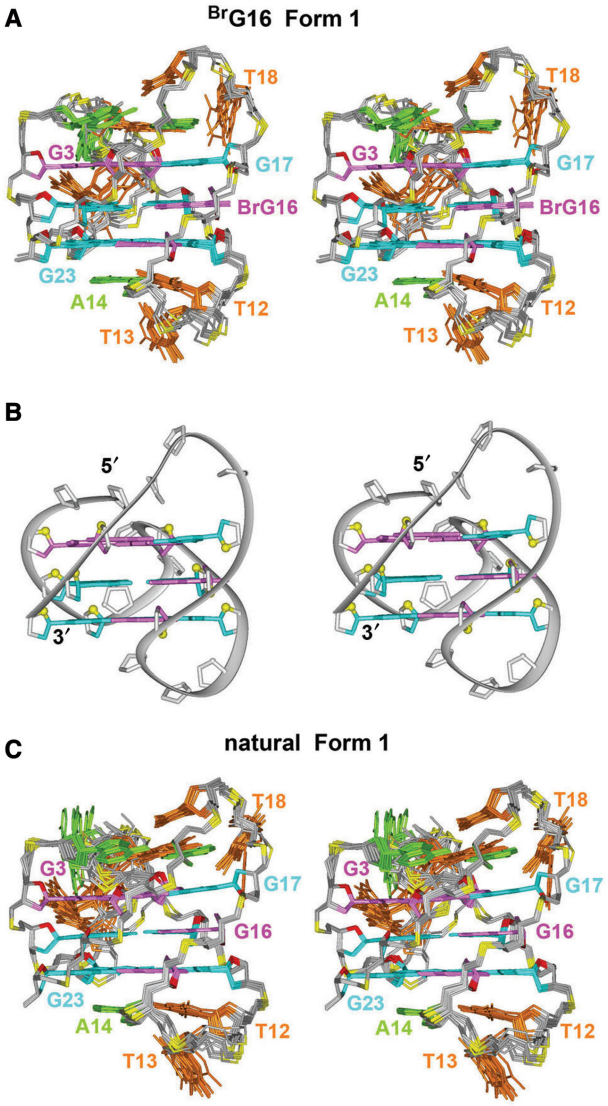

Figure 5. Stereo views of Form 1 structure. (**A**) Ten superpositioned refined structures of ^{Br}G16-Form 1. (**B**) Ribbon view of a representative structure. (**C**) Ten superpositioned refined structures of natural-Form 1. *Anti* and *syn* guanines are colored cyan and magenta, respectively; in (A) and (C), adenines are colored green; thymines, orange; backbone, gray; O4′ atoms, red; phosphorus atoms, yellow. In (B), O4′ atoms are colored yellow.

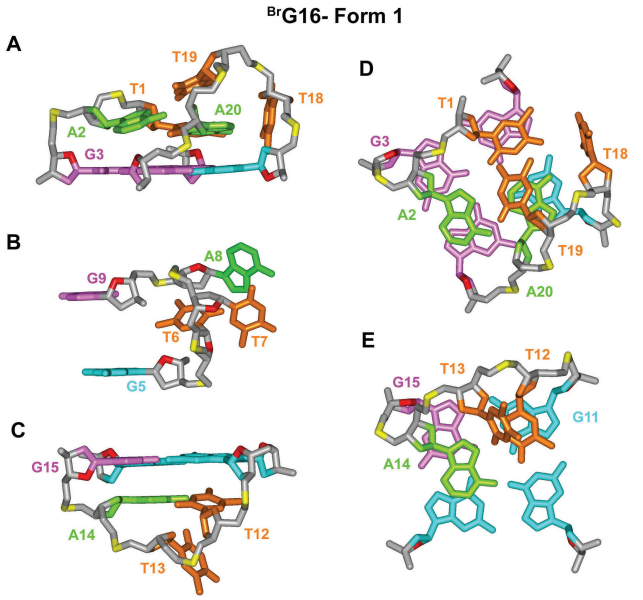

Figure 6. Detailed loop structure of ^{Br}G16-Form 1 (**A**) 5′-end top cap (side view). (**B**) Double-chain-reversal loop (**C**) 3′-end bottom cap (side view). (**D**) 5′-end top cap (top view). (**E**) 3′-end bottom cap (bottom view). Color coded as in Figure 5A.

description of their structure in solution. These observations reinforce a previous proposal by our group of K^+ cation coordination within edgewise loops in G-quadruplexes (57).

Structure of TTA loops in telomere quadruplexes

Four-repeat human telomere sequences can form at least two different intramolecular (3 + 1) G-quadruplexes in K^+ solution. These structures can coexist and be in dynamic equilibrium with other G-quadruplex forms. Such an equilibrium between conformers is reminiscent of what was reported previously by our group for two-repeat human telomere (22) and two-repeat *Tetrahymena* telomere (53) G-quadruplexes in solution. As all possible

Table 1. Statistics of the computed structures of Form 1

A. NMR restraints	^{Br}G16		Natural	
Distance restraints	D_2O	H_2O	D_2O	H_2O
Intra-residue distance restraints	200	0	210	0
Sequential (i, i + 1) distance restraints	71	4	68	11
Long-range (i, $\geq i + 2$) distance restraints	19	15	19	21
Other restraints				
Hydrogen bonding restraints		60		
Torsion angle restraints		53		
Intensity restraints				
Non-exchangeable protons (each of four mixing times)	282			–

B. Structure statistics for 10 molecules following intensity (distance) refinement	^{Br}G16 (intensity)	Natural (distance)
NOE violations		
Number (>0.2Å)	0.10 ± 0.32	0.00 ± 0.00
Maximum violation (Å)	0.21 ± 0.07	0.00 ± 0.00
RMSD of violations	0.02 ± 0.00	0.03 ± 0.00
Deviations from the ideal covalent geometry		
Bond lengths (Å)	0.004 ± 0.000	0.004 ± 0.000
Bond angles (deg)	0.93 ± 0.02	0.92 ± 0.01
Impropers (deg)	0.32 ± 0.02	0.36 ± 0.02
NMR R-factor ($R_{1/6}$)	0.02 ± 0.00	–
Pairwise all heavy atom RMSD values (Å)		
All heavy atoms except T6, T7, A8	0.53 ± 0.19	0.59 ± 0.15
All heavy atoms	0.67 ± 0.26	0.82 ± 0.22

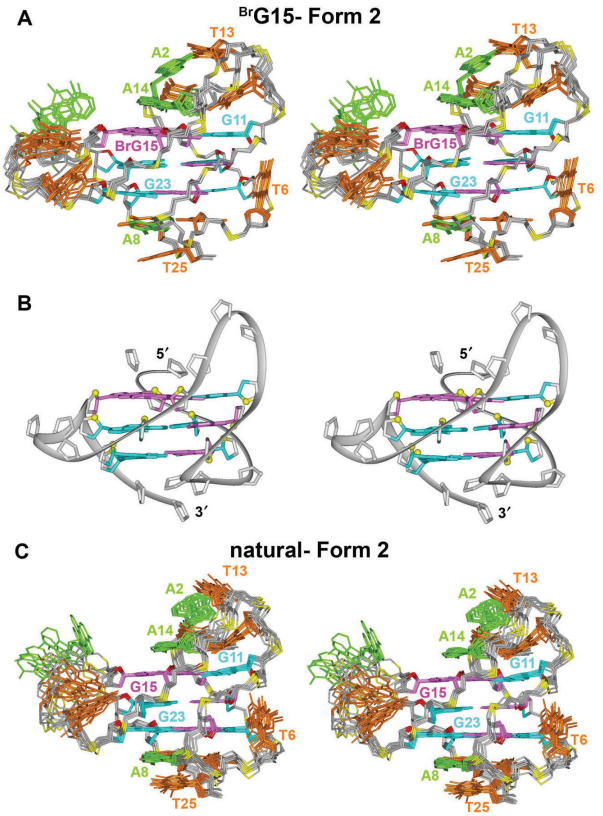

Figure 7. Stereo views of Form 2 structure. (**A**) Ten superpositioned refined structures of ^{Br}G15-Form 2. (**B**) Ribbon view of a representative structure. (**C**) Ten superpositioned refined structures of natural-Form 2. Color coded as in Figure 5A.

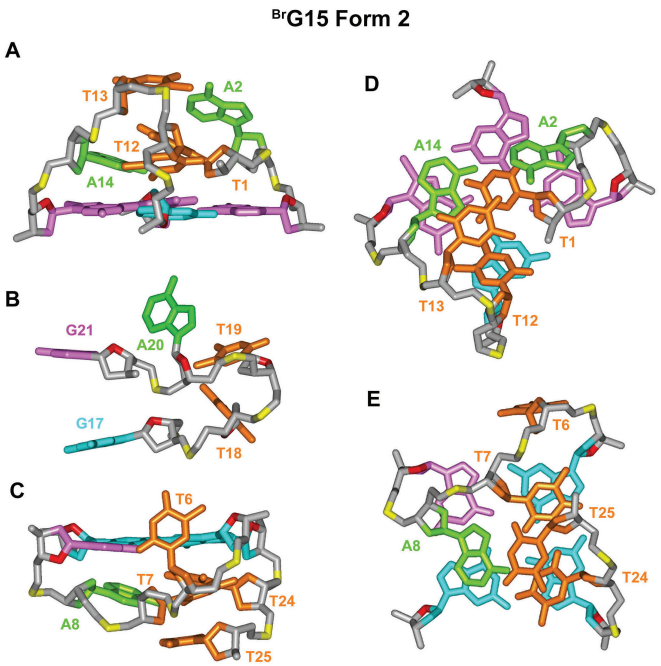

Figure 8. Detailed loop structure of ^{Br}G15-Form 2 (**A**) 5′-end top cap (side view). (**B**) Double-chain-reversal loop (**C**) 3′-end bottom cap (side view); (**D**) 5′-end top cap (top view); (**E**) 3′-end bottom cap (bottom view). Color coded as in Figure 5A.

human telomere G-quadruplexes might contain TTA loops, it is important to gather structural patterns of different TTA loop conformations. We have obtained from this work several different configurations of TTA loops in the context of G-quadruplexes formed by natural sequences. Generally, the bases of edgewise loops tend to maximize pairing by forming non-canonical pairs, triples and triads, which in turn stack over the terminal G-tetrads. In favorable cases, edgewise loop residues could also align to potentially coordinate a K^+ cation, embedded within the loop turn. A range of topologies have been observed for TTA double-chain-reversal loops, with a common theme that two out of the three bases generally tend to stack on each other. Note that even though the present structures were solved for natural human telomere sequences, sequence extension towards either 5′- or 3′-ends can affect the structure of the loops and caps (44). It should also be noted that the presence of a ligand may push the equilibrium towards one particular structure, which may or may not be represented by the free native sequence (58).

CONCLUSION

We have determined the structures of Form 1 and Form 2 intramolecular (3 + 1) G-quadruplexes adopted by natural human telomere sequences in K^+ solution in the

Table 2. Statistics of the computed structures of Form 2

A. NMR restraints	^{Br}G15		Natural	
Distance restraints	D_2O	H_2O	D_2O	H_2O
Intra-residue distance restraints	237	0	177	0
Sequential (i, $i+1$) distance restraints	107	6	85	9
Long-range (i, $\geq i+2$) distance restraints	42	14	22	14
Other restraints				
Hydrogen bonding restraints	60			
Torsion angle restraints	57			
Intensity restraints				
Non-exchangeable protons (each of four mixing times)	347		–	

B. Structure statistics for 10 molecules following intensity (distance) refinement	^{Br}G15 (intensity)	Natural (distance)
NOE violations		
Number (>0.2Å)	0.20 ± 0.42	0.30 ± 0.48
Maximum violation (Å)	0.25 ± 0.02	0.23 ± 0.02
RMSD of violations	0.02 ± 0.00	0.03 ± 0.00
Deviations from the ideal covalent geometry		
Bond lengths (Å)	0.005 ± 0.000	0.004 ± 0.000
Bond angles (deg)	0.95 ± 0.02	0.87 ± 0.01
Impropers (deg)	0.44 ± 0.04	0.37 ± 0.02
NMR R-factor ($R_{1/6}$)	0.02 ± 0.01	–
Pairwise all heavy atom RMSD values (Å)		
All heavy atoms except T18, T19, A20	0.49 ± 0.13	0.79 ± 0.13
All heavy atoms	0.80 ± 0.25	1.46 ± 0.36

presence of up to 30% of minor conformations. Both structures contain the (3 + 1) G-tetrad core with one double-chain-reversal and two edgewise loops, but differ in the successive order of loop appearance within the G-quadruplex scaffold. Our results provide the structural details at the two ends of the G-tetrad core in the context of natural sequences and context-dependent information on edgewise and double-chain-reversal loop conformations. Comparison between different TTA loop conformations has revealed structural patterns, which are likely to recur within the family of G-quadruplex structures adopted by human telomere sequences.

NOTE

A paper on the NMR-based solution structure of Form 2 human telomere G-quadruplex for the sequence d[TTAGGGTTAGGGTTAGGGTTAGGGTT] appeared online (coordinates are currently on hold) during the review of our paper. This sequence contains an additional T at the 5′-end compared to our Form 2 sequence (Dai,J., Carver,M., Punchihewa,C., Jones,R.A. and Yang,D. (2007) Structure of the hybrid-2 type intramolecular human telomeric G-quadruplex in K^+ solution: insights into structure polymorphism of the human telomeric sequence. *Nucleic Acids Res*. online).

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

This research was supported by US National Institutes of Health Grant GM34504 to D.J.P. and Singapore Ministry of Education Grants SUG05/06 and RG138/06 to A.T.P. D.J.P. is a member of the New York Structural Biology Center supported by US National Institutes of Health Grant GM66354. Funding to pay the Open Access publication charges for this article was provided by US National Institutes of Health Grant GM34504.

Conflict of interest statement. None declared.

REFERENCES

1. Gellert,M.N., Lipsett,M.N. and Davies,D.R. (1962) Helix formation by guanylic acid. *Proc. Natl Acad. Sci. USA*, **48**, 2013–2018.
2. Simonsson,T. (2001) G-quadruplex DNA structures–variations on a theme. *Biol. Chem.*, **382**, 621–628.
3. Davis,J.T. (2004) G-quartets 40 years later: from 5′-GMP to molecular biology and supramolecular chemistry. *Angew. Chem. Int. Ed. Engl.*, **43**, 668–698.
4. Phan,A.T., Kuryavyi,V. and Patel,D.J. (2006) DNA architecture: from G to Z. *Curr. Opin. Struct. Biol.*, **16**, 288–298.
5. Burge,S., Parkinson,G.N., Hazel,P., Todd,A.K. and Neidle,S. (2006) Quadruplex DNA: sequence, topology and structure. *Nucleic Acids Res.*, **34**, 5402–5415.
6. 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.
7. Wang,Y. and Patel,D.J. (1993) Solution structure of the human telomeric repeat $d[AG_3(T_2AG_3)_3]$ G-tetraplex. *Structure*, **1**, 263–282.
8. Parkinson,G.N., Lee,M.P.H. and Neidle,S. (2002) Crystal structure of parallel quadruplexes from human telomeric DNA. *Nature*, **417**, 876–880.
9. Schaffitzel,D.L., Berger,I., Postberg,J., Hanes,J., Lipps,H.J. and Plucthun,A. (2001) *In vitro* generated antibodies specific for telomeric guanine-quadruplex DNA react with *Stylonychia lemnae* macronuclei. *Proc. Natl Acad. Sci. USA*, **98**, 8572–8577.
10. Paeschke,K., Simonsson,T., Postberg,J., Rhodes,D. and Lipps,H.J. (2005) Telomere end-binding proteins control the formation of G-quadruplex DNA structures *in vivo*. *Nat. Struct. Mol. Biol.*, **12**, 847–854.
11. Maizels,N. (2006) Dynamic roles for G4 DNA in the biology of eukaryotic cells. *Nat. Struct. Mol. Biol.*, **13**, 1055–1059.
12. Neidle,S. and Parkinson,G. (2002) Telomere maintenance as a target for anticancer drug discovery. *Nat. Rev. Drug Discov.*, **1**, 383–393.
13. Hurley,L.H. (2002) DNA and its associated processes as targets for cancer therapy. *Nat. Rev. Cancer*, **2**, 188–200.
14. Mergny,J.L., Riou,J.F., Mailliet,P., Teulade-Fichou,M.P. and Gilson,E. (2002) Natural and pharmacological regulation of telomerase. *Nucleic Acids Res.*, **30**, 839–865.
15. Chang,C.C., Kuo,I.C., Ling,I.F., Chen,C.T., Chen,H.C., Lou,P.J., Lin,J.J. and Chang,T.C. (2004) Detection of quadruplex DNA structures in human telomeres by a fluorescent carbazole derivative. *Anal. Chem.*, **76**, 4490–4494.
16. Gomez,D., O'Donohue,M.F., Wenner,T., Douarre,C., Macadre,J., Koebel,P., Giraud-Panis,M.J., Kaplan,H., Kolkes,A. *et al.* (2006) The G-quadruplex ligand telomestatin inhibits POT1 binding to telomeric sequences *in vitro* and induces GFP-POT1 dissociation from telomeres in human cells. *Cancer Res.*, **66**, 6908–6912.
17. Zahler,A.M., Williamson,J.R., Cech,T.R. and Prescott,D.M. (1991) Inhibition of telomerase by G-quartet DNA structures. *Nature*, **350**, 718–720.
18. Zaug,A.J., Podell,E.R. and Cech,T.R. (2005) Human POT1 disrupts telomeric G-quadruplexes allowing telomerase extension *in vitro*. *Proc. Natl Acad. Sci. USA*, **102**, 10864–10869.
19. Oganesian,L., Moon,I.K., Bryan,T.M. and Jarstfer,M.B. (2006) Extension of G-quadruplex DNA by ciliate telomerase. *EMBO J.*, **25**, 1148–1159.
20. Greider,C.W. and Blackburn,E.H. (1985) Identification of a specific telomere terminal transferase activity in *Tetrahymena* extracts. *Cell*, **43**, 405–413.
21. 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.
22. 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.
23. Ying,L., Green,J.J., Li,H., 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.
24. 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 $AG_3(T2AG_3)_3$ and $(T_2AG_3)_4$ human telomere sequences in Na^+ and K^+ solutions. *Nucleic Acids Res.*, **31**, 1605–1613.
25. He,Y., Neumann,R.D. and Panyutin,I.G. (2004) Intramolecular quadruplex conformation of human telomeric DNA assessed with ^{125}I-radioprobing. *Nucleic Acids Res.*, **32**, 5359–5367.
26. Xu,Y. and Sugiyama,H. (2004) Highly efficient photochemical 2′-deoxyribonolactone formation at the diagonal loop of a 5-iodouracil-containing antiparallel G-quartet. *J. Am. Chem. Soc.*, **126**, 6274–6279.
27. D'Isa,G., Galeone,A., Oliviero,G., Piccialli,G., Varra,M. and Mayol,L. (2004) Effect of gamma-hydroxypropano deoxyguanosine, the major acrolein-derived adduct, on monomolecular quadruplex structure of telomeric repeat $d(TTAGGG)_4$. *Bioorg. Med. Chem. Lett.*, **14**, 5417–5421.

28. Hazel,P., Huppert,J., Balasubramanian,S. and Neidle,S. (2004) Loop-length-dependent folding of G-quadruplexes. *J. Am. Chem. Soc.*, **126**, 16405–16415.
29. 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.
30. 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.
31. 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.
32. 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.
33. Qi,J. and Shafer,R.H. (2005) Covalent ligation studies on the human telomere quadruplex. *Nucleic Acids Res.*, **33**, 3185–3192.
34. 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.
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. 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.
37. Lee,J.Y., Okumus,B., Kim,D.S. and Ha,T. (2005) Extreme conformational diversity in human telomeric DNA. *Proc. Natl Acad. Sci. USA*, **102**, 18938–18943.
38. Jaumot,J., Eritja,R., Tauler,R. and Gargallo,R. (2006) Resolution of a structural competition involving dimeric G-quadruplex and its C-rich complementary strand. *Nucleic Acids Res.*, **34**, 206–216.
39. Kan,Z.Y., Yao,Y., Wang,P., Li,X.H., Hao,Y.H. and Tan,Z. (2006) Molecular crowding induces telomere G-quadruplex formation under salt-deficient conditions and enhances its competition with duplex formation. *Angew. Chem. Int. Ed. Engl.*, **45**, 1629–1632.
40. Yu,H.Q., Miyoshi,D. and Sugimoto,N. (2006) Characterization of structure and stability of long telomeric DNA G-quadruplexes. *J. Am. Chem. Soc.*, **128**, 15461–15468.
41. Zhang,N., Phan,A.T. and Patel,D.J. (2005) (3 + 1) Assembly of three human telomeric repeats into an asymmetric dimeric G-quadruplex. *J. Am. Chem. Soc.*, **127**, 17277–17285.
42. Wang,Y. and Patel,D.J. (1994) Solution structure of the *Tetrahymena* telomeric repeat $d(T_2G_4)_4$ G-tetraplex. *Structure*, **2**, 1141–1156.
43. 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.
44. 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.
45. Ambrus,A., Chen,D., Dai,J., Bialis,T., Jones,R.A. and Yang,D. (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.
46. 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.
47. Dias,E., Battiste,J.L. and Williamson,J.R. (1994) Chemical probe for glycosidic conformation in telomeric DNAs. *J. Am. Chem. Soc.*, **116**, 4479–4480.
48. Esposito,V., Randazzo,A., Piccialli,G., Petraccone,L., Giancola,C. and Mayol,L. (2004) Effects of an 8-bromodeoxyguanosine incorporation on the parallel quadruplex structure $[d(TGGGT)]_4$. *Org. Biomol. Chem.*, **2**, 313–318.
49. Borman,S. (2006) Quadruplex in its elements: structures of human telomeric quadruplex in cell-like solution have implications for anticancer therapeutics. *Chem Eng News*, **84**, 46.
50. Dai,J., Punchihewa,C., Ambrus,A., Chen,D., Jones,R.A. and Yang,D. (2007) Structure of the intramolecular human telomeric G-quadruplex in potassium solution: a novel adenine triple. *Nucleic Acids Res.*, **35**, 2440–2450.
51. Matsugami,A., Xu,Y., Noguchi,Y., Sugiyama,H. and Katahira,M. (2007) Structure of a human telomeric DNA sequence stabilized by 8-bromoguanosine substitutions, as determined by NMR in a K^+ solution. *FEBS J.*, **274**, 3545–3556.
52. Phan,A.T. and Patel,D.J. (2002) A site-specific low-enrichment ^{15}N,^{13}C isotope-labeling approach to unambiguous NMR spectral assignments in nucleic acids. *J. Am. Chem. Soc.*, **124**, 1160–1161.
53. Phan,A.T., Modi,Y.S. and Patel,D.J. (2004) Two-repeat *Tetrahymena* telomeric d(TGGGGTTGGGGT) sequence interconverts between asymmetric dimeric G-quadruplexes in solution. *J. Mol. Biol.*, **338**, 93–102.
54. Phan,A.T., Guéron,M. and Leroy,J.L. (2001) Investigation of unusual DNA motifs. *Methods Enzymol.*, **338**, 341–371.
55. Phan,A.T. (2000) Long-range imino proton-^{13}C J-couplings and the through-bond correlation of imino and non-exchangeable protons in unlabeled DNA. *J. Biomol. NMR*, **16**, 175–178.
56. Brünger,A.T. (1992) *X-PLOR: A System for X-ray Crystallography and NMR*. Yale University Press, New Haven, CT.
57. Bouaziz,S., Kettani,A. and Patel,D.J. (1998) A K-cation induced conformational switch within a loop spanning segment of a DNA quadruplex containing G-G-G-C repeats. *J. Mol. Biol.*, **282**, 637–652.
58. Parkinson,G.N., Ghosh,R. and Neidle,S. (2007) Structural basis for binding of porphyrin to human telomeres. *Biochemistry*, **46**, 2390–2397.