Different loop arrangements of intramolecular human telomeric (3+1) G-quadruplexes in K⁺ solution

Anh Tuấn Phan*, Kim Ngoc Luu and Dinshaw J. Patel*

Structural Biology Program, Memorial Sloan-Kettering Cancer Center, New York, NY 10021, USA

Received August 1, 2006; Revised September 18, 2006; Accepted September 19, 2006

ABSTRACT

Intramolecular G-quadruplexes formed by the human telomeric G-rich strand are promising anticancer targets. Here we show that four-repeat human telomeric DNA sequences can adopt two different intramolecular G-quadruplex folds in K⁺ solution. The two structures contain the (3+1) G-tetrad core, in which three G-tracts are oriented in one direction and the fourth in the opposite direction, with one double-chain-reversal and two edgewise loops, but involve different loop arrangements. This result indicates the robustness of the (3+1) core G-quadruplex topology, thereby suggesting it as an important platform for structure-based drug design. Our data also support the view that multiple human telomeric G-quadruplex conformations coexist in K⁺ solution. Furthermore, even small changes to flanking sequences can perturb the equilibrium between different coexisting G-quadruplex forms.

INTRODUCTION

Telomeres, the ends of eukaryotic chromosomes, are important for chromosome stability (1–3). Human telomeric DNA, ~4–14 kb long (4), is composed of (TTAGGG/CCCTAA)ₙ repeats, with a 100–200 nt G-rich strand...
overhang at 3' ends (5–7). Telomerase, an enzyme that can elongate this strand, is not activated in normal human somatic cells, but highly expressed in 80–85% cancer cells and helps to maintain the length of telomeres in these cells (8). G-rich sequences can form G-quadruplex structures (9–11) stabilized by cations such as Na$^{+}$ or K$^{+}$. Intramolecular G-quadruplexes formed by the human telomeric G-rich strand are promising anticancer targets (1–3), as formation of such structures has been shown to inhibit the activity of telomerase (12–14).

The structure of intramolecular human telomeric G-quadruplexes is the subject of intense research (15–33) but still puzzling. Until recently only Na$^{+}$ solution (15) and K$^{+}$-containing crystal (16) structures were available and these structures are very different from each other. In the former structure (15), guanines around each tetrad are syn/syn/anti/anti, loops are successively edgewise–diagonal–edgewise and each G-tract has both a parallel and an antiparallel adjacent strands. In contrast, the crystal structure in the presence of K$^{+}$ revealed a completely different intramolecular G-quadruplex (16), in which all strands are parallel, guanines are anti and loops are double-chain-reversal. Because K$^{+}$ is much more abundant than Na$^{+}$ in cellular environments, the structure of telomeric G-quadruplexes in K$^{+}$ solution is most important. Data from many laboratories (17–33) indicated the presence of a mixture of several G-quadruplex forms in K$^{+}$ solution. Recently, we (34) and others (35,36) reported the folding topology of an intramolecular human telomeric G-quadruplex (Form 1) in K$^{+}$ solution, with our group also reporting on its solution structure (34). This so-called (3+1) G-quadruplex involves the following: one anti/syn/syn/syn and two syn/anti/anti/anti G-tetrads; one double-chain-reversal loop formed by the first TTA linker and two other edgewise loops; three G-tracts oriented in one direction and the fourth in the opposite direction. Here we report on a new intramolecular human telomeric G-quadruplex in K$^{+}$ solution. Intriguingly, this structure contains the same (3+1) core but differs from Form 1 by loop arrangements.

**MATERIALS AND METHODS**

**Sample preparation**

Unlabeled and site-specific low-enrichment (2% $^{15}$N-labeled) oligonucleotides were synthesized on an ABI 392 DNA synthesizer.
synthesizer and purified by HPLC (17). They were dialyzed successively against 50 mM KCl solution and against water. The strand concentration of the NMR samples was typically 0.5–5 mM; the solutions contained 70 mM KCl and 20 mM potassium phosphate (pH 7).

**NMR spectroscopy**

Experiments were performed on 600 MHz Varian and 800 MHz Bruker spectrometers at 25°C. Resonances were assigned unambiguously by using site-specific low-enrichment labeling and through-bond correlations at natural abundance (37–39). Experiments in H$_2$O used the jump-and-return (JR) water suppression (39,40) for detection.

**RESULTS AND DISCUSSION**

Favoring a new G-quadruplex conformation

We systematically examined NMR spectra of human telomeric sequences (21–27 nt) containing four G-tracts and found that in all cases there were mixtures of multiple G-quadruplexes in K$^+$ solution. However, certain spectral patterns could be recognized (Figure 1), which were apparently characteristic for different G-quadruplex forms, and the ratio between these forms varied among sequences (Figure 1). Possibly, some terminal residues could favor or disfavor a particular form by interacting with the loops (34). Indeed, NMR spectra (Figure 2) suggested that the 25-nt human telomeric d(TAGGG(TTAGGG)TT) sequence adopts predominantly (~70%) a G-quadruplex in K$^+$ solution,
which is different from Form 1, the major form of the 23-nt human telomeric d[TAGGG(TTAGGG)3] sequence reported previously (34).

### Determination of G-quadruplex folding topology

Sharp linewidths (~2–3 Hz for the sharpest peaks at 25°C) of the major form of this 25-nt sequence are consistent with a monomeric intramolecular G-quadruplex. Guanine imino protons were unambiguously assigned by the site-specific low-enrichment approach (37) using 2% 15N-labeled samples (Figure 3). Guanine H8 protons were assigned by natural-abundant through-bond correlations (38) to the already assigned imino protons (Figure 4). Some H8 proton assignments were independently confirmed in site-specific 2% 15N-labeled samples (37) (data not shown). Specific NOEs between imino and H8 protons (Figure 5a, c and d), established the fold (Form 2) of this G-quadruplex (Figure 5e), involving three G-tetrads: G3-G21-G15-G11, G4-G10-G16-G22 and G5-G9-G17-G23. The third linker T18-T19-A20 forms a double-chain-reversal loop, while the two other linkers T6-T7-A8 and T12-T13-A14 form edgewise loops. The G-tetrad core is of (3-1)-type (34,41–43), in which three G-tracts are oriented in one direction and the fourth in the opposite direction. The glycosidic conformations of the top G-tetrad are anti-syn-anti-syn, those of the two others are syn-anti-syn-anti, consistent with H1’-H8 NOE intensities observed for these residues (Figure 5b). This G-quadruplex fold is also consistent with other NOEs, such as the rectangular H8–H1’ patterns (44) (Figure 5b) of four 5’-syn-anti-3’ steps in the structure (Figure 5e). The G-quadruplex fold was supported by proton exchange data, which showed that imino protons of the central G-tetrad are the most protected from exchange with water (Figure 3b).

### Comparison with related G-quadruplex structures

Like Form 1, Form 2 human telomeric G-quadruplex observed in K+ solution here (Figure 5e) is very different from the structures observed in Na+ solution or in a K+-containing crystal, in terms of strand orientations, loop topologies and glycosidic conformations of guanines. Form 2 (Figure 6a) differs from Form 1 (Figure 6b) by loop arrangements: the double-chain-reversal loop is formed by the third linker instead of the first linker. The folding topology of Form 2 is similar to the (3-1) G-quadruplex formed by the four-repeat *Tetrahymena* telomeric sequence (42) and by a variant sequence of four G-tracts from the human *bcl-2* promoter (43). The three-repeat human telomeric sequence adopts a (3-1) G-quadruplex with yet another loop arrangement (41).

### Possible intramolecular (3-1) G-quadruplexes

Figure 6 shows four possible intramolecular human telomeric (3-1) G-quadruplexes, in which the 5’ end starts from different corners of the same G-tetrad core. It is worth mentioning that NMR spectra of the human telomeric d[TAGGG(TTAGGG)3TT] sequence show a downfield-shifted peak for A20(H8-H1’) of Form 2 together with two other downfield-shifted minor A(H8-H1’) peaks (red box, Figure 5b), which are characteristic of an adenine in the double-chain-reversal TTA loop (34). It is likely that the latter belong to minor species that contain double-chain-reversal loops, such as the parallel-stranded G-quadruplex observed in a K+-containing crystal (16) or possibly different (3-1) G-quadruplex conformations shown in Figure 6.

### CONCLUSION

We have determined the folding topology of a new intramolecular G-quadruplex conformation of the human telomere in K+ solution. Taken together with recent findings (34–36), this work has established that four-repeat human telomeric sequences can form two intramolecular (3-1) G-quadruplexes in K+ solution, which differ from each other only by the order of loop arrangements. This indicates the robustness of the (3-1) core G-quadruplex topology, first reported a decade ago (42), thereby suggesting it as an important platform for structure-based drug design. The presence of a mixture of multiple G-quadruplexes and the differential effects of terminal residues (or possible attached dyes) on the stability of different forms should be considered when one studies telomeric G-quadruplexes.

### ACKNOWLEDGEMENTS

This research was supported by National Institutes of Health Grant GM34504. D.J.P. is a member of the New York Structural Biology Center supported by National Institutes of Health Grant GM66354. Funding to pay the Open Access publication charges for this article was provided by National Institutes of Health Grant GM34504.

Conflict of interest statement. None declared.
REFERENCES

1. Neidle,S. and Parkinson,G. (2002) Telomere maintenance as a target for anticancer drug discovery. Nature Rev. Drug. Discov., 1, 383–393.
2. Hurley,L.H. (2002) DNA and its associated processes as targets for cancer therapy. Nature Rev. Cancer, 2, 188–200.
3. Mergny,J.L., Riou,J.F., Maillet,P., Teulade-Fichou,M.P. and Gilson,E. (2002) Natural and pharmacological regulation of telomerase. Nucleic Acids Res., 30, 839–865.
4. de Lange,T., Shiue,L., Myers,R.M., Cox,D.R., Naylor,S.L., Killely,A.M. and Varmus,H.E. (1990) Structure and variability of human chromosome ends. Mol. Cell. Biol., 10, 518–527.
5. 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.
6. McElligott,R. and Wellinger,J. (1997) The terminal DNA structure by STE-11 analysis of telomeric DNA. EMBO J., 16, 3705–3714.
7. Wright,W.E., Tesmer,V.M., Huffman,K.E., Levene,S.D. and Shay,J.W. (1997) Normal human chromosomes have long G-rich telomeric overhangs at one end. Genes Dev., 11, 2801–2809.
8. Kim,N.W., Ptatszeg,M.A., Prowse,K.R., Harley,C.B., West,M.D., Ho,P.L., Coviello,G.M., Wright,W.E., Weinrich,S.L. and Shay,J.W. (1994) Specific association of human telomerase activity with immortal cells and cancer. Science, 266, 2011–2015.
9. Simonsson,T. (2001) G-quadruplex DNA structures - variations on a theme. Biol. Chem., 382, 621–628.
10. Davis,J.T. (2004) G-quartets 40 years later: from 5'-GMP to molecular biology and supramolecular chemistry. Curr. Opin. Struct. Biol., 14, 668–699.
11. Phan,A.T., Kuryavyi,V. and Patel,D.J. (2006) DNA architecture: from sequences in Na+ solutions. Bioorg. Med. Chem. Lett., 16, 2723–2735.
12. Yale,N., Noguchi,Y. and Sugiyama,H. (2006) The new models of the telomeric repeat d(TTAGGG)4. Nucleic Acids Research, 2006, Vol. 34, No. 19, 5715–5719.
13. Zaug,A.J., Podell,E.R. and Cech,T.R. (2005) Human POT1 disrupts the hairpin structure of the human telomeric repeat d(TTAGGG)4. Nucleic Acids Res., 33, 3185–3192.
14. Phan,A.T., Kuryavyi,V. and Patel,D.J. (2006) DNA architecture: from sequences in Na+ solutions. Bioorg. Med. Chem. Lett., 16, 2723–2735.
15. Phan,A.T. (2000) Long-range imino proton-13C J-couplings and the G-quadruplex scaffold. Structure, 1, 263–282.
16. Zaug,A.J., Podel,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.
17. Oganessian,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.
18. Wang,Y. and Patel,D.J. (1993) Solution structure of the human telomeric repeat d[AG(T2AG3)3] G-tetraplex. Structure, 1, 263–282.
19. Parkinson,G.N., Lee,M.P.H. and Neidle,S. (2002) Crystal structure of parallel quadruplexes from human telomeric DNA. Nature, 417, 876–880.
20. Phan,A.T. and Patel,D.J. (2003) Two-repeat human telomeric d(TAGGGTTAGGG) 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.
21. Ying,L., Green,J.J., Li,H., Klenerman,D. and Balusubramanian,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.
22. Wang,Y. and Patel,D.J. (1993) Solution structure of the human telomeric repeat d[AG(T2AG3)3] G-tetraplex. Structure, 1, 263–282.
23. Xu,Y. and Sugiyama,H. (2004) Highly efficient photochemical 2'-deoxyribonucleotide formation at the diagonal loop of a 5'-iodouracil-containing antiparallel G-quartet. J. Am. Chem. Soc., 126, 6274–6279.
24. D’Isa,G., Galeone,A., Oliviero,G., Piccialli,G., Varra,M. and Mayo,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.
25. Phan,A.T. and Patel,D.J. (2002) A site-specific low-enrichment-15N,13C isotope-labeling approach to unambiguous NMR spectral assignments in nucleic acids. J. Am. Chem. Soc., 124, 1160–1161.
26. Phan,A.T. (2000) Long-range imino proton-13C J-couplings and the through-bond correlation of imino and non-exchangeable protons in unlabeled DNA. J. Biomol. NMR, 16, 175–178.
27. Phan,A.T., Guéron,M. and Leroy,J.L. (2001) Investigation of unusual DNA motifs. Methods Enzymol., 338, 341–371.
28. Plateau,P. and Guéron,M. (1982) Exchangeable proton NMR without base-line distortion using strong pulse sequences. J. Am. Chem. Soc., 104, 7310–7311.
29. Zhang,N. (2005) (3+1) Assembly of three human telomeric repeats into an asymmetric dimeric G-quadruplex. J. Am. Chem. Soc., 127, 17277–17285.
30. Wang,Y. and Patel,D.J. (1994) Solution structure of the Tetrahymena telomeric repeat d(T2AG3) G-tetraplex. Structure, 2, 1141–1156.
31. Dai,J., Dexeheimer,T.S., Chen,D., Carver,M., Ambrus,A., Jones,R.A. and Yang,D. (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.
32. Feigon,J., Koshlap,K., and Smith,F.W. (1995) 1H NMR spectroscopy of DNA triplexes and quadruplexes. Methods Enzymol., 261, 225–255.