2-2014

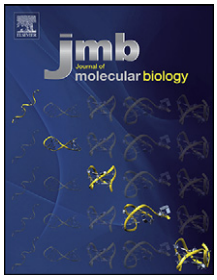
Kinetics and Structures on the Molecular Path to the Quadruplex Form of the Human Telomere

W. David Wilson
Georgia State University, wdw@gsu.edu

Ananya Paul
Georgia State University, apaul@gsu.edu

Recommended Citation

Kinetics and structures on the molecular path to the quadruplex form of the human telomere. W. David Wilson and A. Paul. doi: http://dx.doi.org/10.1016/j.jmb.2014.02.001



Kinetics and Structures on the Molecular Path to the Quadruplex Form of the Human Telomere

W. David Wilson and Ananya Paul

Department of Chemistry, *Georgia State University, Atlanta, GA 30303, USA*

Correspondence to W. David Wilson: *Department of Chemistry, Natural Science Center, Georgia State University, 50 Decatur Street, Atlanta, GA 30303, USA.*
http://dx.doi.org/10.1016/j.jmb.2014.02.001
Edited by S. A. Woodson

Telomeres, which form the ends of chromosomes in eukaryotic organisms, have a number of important functions from protection of the chromosome ends from degradation to processes in the expression of some genes, as well as roles in protection of cells from aging due to loss of DNA in replication and in cancer development [1–4]. The human telomere starts at the ends of chromosomal DNA with the repeat duplex sequence 5′-(TTAGGG)-3′ 5′-(CCCTAA)-3′ and moves to a single-stranded, terminal 3′ overhang of the G-rich strand. It has been known for quite some time that G nucleotides and G-rich sequences can associate into planar tetrads that are stabilized by tetrad stacking, multiple H-bonds and cation interactions (Fig. 1a) [5,6]. The tetrad structure creates a central hole that is lined by four oxygen atoms to form an excellent macrocyclic receptor for K^+ ions. Sodium ions and other cations can also fill the hole, but K^+ has the optimum properties and generally provides the highest stability [7–9]. There are a number of different conformational possibilities for telomere sequences, and two important examples, found experimentally, are shown in Fig. 1. Some of the more important features of the folded states include strand directionality, parallel or antiparallel, whether the bases are *syn* or *anti*, and the connecting TTA loop arrangements [8–12].

Wang and Patel determined a structure of a single G-rich strand of the human telomere, 5′-d[AGGG(TTAGGG)$_3$], (Htelo) in Na^+ solution by using NMR and molecular dynamics methods (Fig. 1b) [11]. The structure has three stacked G-tetrads, as expected from the GGG repeats, which are formed through folding of the single strand. The strands in the Na^+ structure are connected into an antiparallel folding pattern by folding with two lateral loops and one central TTA loop. This structure creates four grooves, one wide, two medium and one narrow. The G nucleotides have both *syn* and *anti* orientations. Neidle and coworkers.[12] solved an X-ray crystal structure of the Htelo sequence by crystallization from K^+ media in polyethylene glycol and found a very different structure (Fig. 1c) [12]. In the K^+ crystal structure, all four DNA strands are parallel with the three linking TTA-connecting sequences oriented in chain-reversal loops along the exterior of the tetrad stack. This gives the structure a propeller-type appearance when viewed from the top. As in the Na^+ structure, the G-tetrads are stacked and stabilized by π–π interactions, H-bonds and cation complexes. All G nucleotides in this structure are in the *anti* conformation, and the stacking pattern results in four almost equivalent grooves. These two very different structures for Htelo suggested environment-dependent polymorphism of the structures formed from this sequence [11–15].

Detailed analysis of the solution structure of Htelo in K^+ solution revealed a mixture of structures and confirmed the polymorphic possibilities of the Htelo sequence [9,14–18]. Through modification of the terminal nucleotides, it has been possible to isolate different single species from the polymorphic ensemble in solution and to determine their structural details by NMR [16–18]. Interestingly, hybrid structures that have both parallel and antiparallel strands are a common observation [19,20]. Two related structures of particular interest have been referred to as Hybrid-1 and Hybrid-2, and their folding patterns are shown in Fig. 2. As can be seen, both of these have three parallel strands and one antiparallel strand that are connected by two lateral loops and one chain-reversal loop. The two structures are stabilized by different terminal stacking nucleotides and have the strand orientations and loops in different positions.

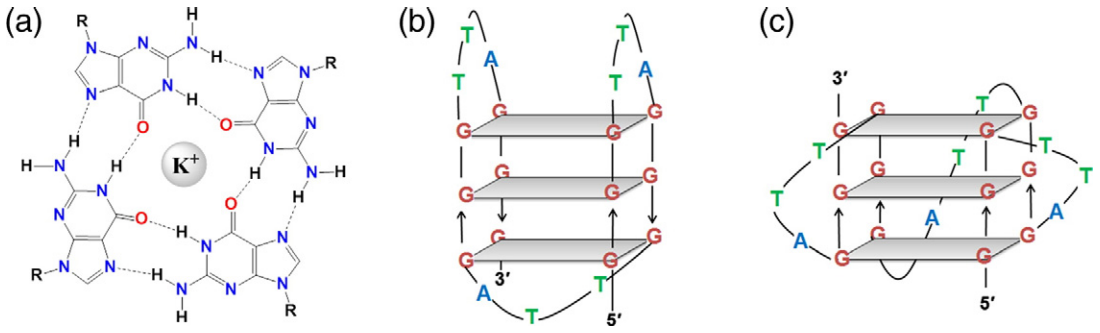


Fig. 1. Structures for (a) a G-quadruplex tetrad interaction, (b) a basket-type intramolecular G-quadruplex found in Na^+ solution [11] and (c) a propeller-type parallel-stranded intramolecular G-quadruplex found in the presence of K^+ in the crystalline state [12].

In addition to the structures of Htelo, a functionally very important and interesting question is how does the G-rich Htelo sequence fold from the open stranded into the quadruplex conformation [19]? Protein folding has been a topic of intense interest and research for over 50 years, and more recently, the folding mechanisms of complex RNA 3′ structures have attracted interest. DNA duplex formation has been less extensively studied, but with the increasing evidence that complex DNA structures, such as the quadruplex, exist in cells and have important functions in both normal and diseased cells [21], the question of how these structures fold, as well as the kinetics of folding, has taken on greater significance. The Sugiyama laboratory [20] performed extensive molecular dynamic simulations to model possible quadruplex folding pathways and were the first to propose a triple-helix intermediate, but those models were not experimentally tested until now.

In the *Journal of Molecular Biology* [19], the authors have used three different, powerful spectroscopic probes, circular dichroism, FRET (*f*luorescence *r*esonance *e*nergy *t*ransfer) and 2AP (*2-a*rnino*p*urine) fluorescence changes, to monitor Htelo folding in the most detailed study to date [19]. The advantage of using these three is that each monitors a structurally different aspect of Htelo folding. CD primarily monitors global quadruplex folding since it is most sensitive to stacking of the G-tetrads. Different quadruplexes, as well as DNA duplexes and triplexes, usually have different CD spectra that provide patterns for their recognition in solution [22]. The FRET signals arise from interaction of probes at the 5′ and 3′ ends of the G-rich sequence and monitor the distance between the two ends [23]. Fluorescence from 2AP arises from replacement of an A base in the quadruplex TTA loops, and fluorescence from this group can provide a monitor of folding of each loop [24]. With these probes, the authors were able to monitor the kinetics of folding at both the global level and at specific sites. They used three DNA sequences, the d[AGGG(TTAGGG)$_3$] natural sequence (Tel22), and the two hybrid

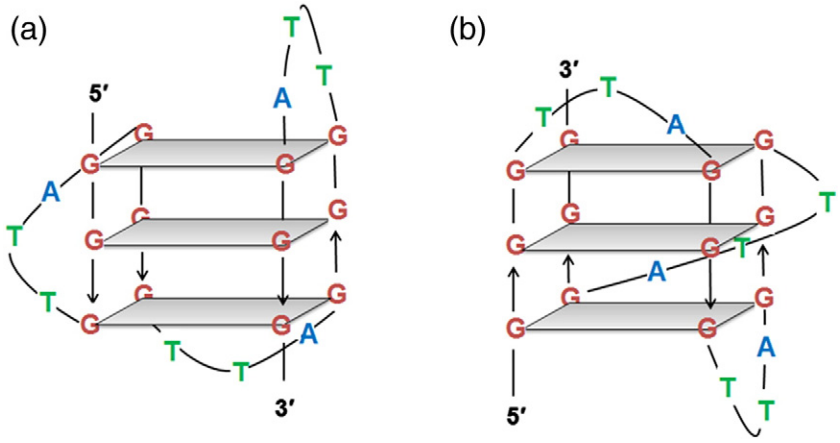


Fig. 2. Schematic outline of folding topologies of the Hybrid-1 (a) and Hybrid-2 (b) intramolecular human telomere G-quadruplexes in K^+ solution.

$$U \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} I_1 \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} I_2 \underset{k_{-3}}{\overset{k_3}{\rightleftharpoons}} I_3 \underset{k_{-4}}{\overset{k_4}{\rightleftharpoons}} F$$

Scheme 1. Folding pathway to the final structure (F) of the human telomere quadruplex in solution.

sequences found by NMR analysis to give primarily single structures (Fig. 2), to follow Htelo folding.

Interestingly, in spite of their simpler, ensemble of structures, Hybrid-1 and Hybrid-2 follow a similar complex folding pathway to the natural sequence. Folding of all these sequences, which can be induced by adding K^+ to U, appears to follow a multistep mechanism that requires at least three intermediates (Scheme 1): where *U* represents unfolded oligonucleotide conformers; I_1, I_2 and I_3 are folding intermediates and *F* is the final folded state or ensemble of states. By using the spectroscopic methods described above along with SVD (*s*ingle *v*alue *d*ecomposition) analysis, the authors have been able not only to define the number of species but also to propose structures for each intermediate that fit well into what we know about DNA duplex and quadruplex folding in general.

The first intermediate, I_1, forms very rapidly and is proposed to be dominated by hairpin conformations formed by folding of the G-rich strand to allow G–G interactions that are relatively weak and dynamic (Fig. 3). The I_2 folded state forms next, in under 10 s, but the authors were able to obtain a CD spectrum for I_2 by conducting short-time (2 s) CD kinetics studies as a function of wavelength. This intermediate spectrum is very informative and is characteristic of an antiparallel structure, such as the Na^+ structure in Fig. 1, but the precise conformation cannot be defined by CD. It seems likely that the state is an ensemble of antiparallel structures.

The steps in conversion of $I_2 \rightarrow I_3 \rightarrow F$ are significantly slower and, as a result, can be evaluated by manual mixing of *U* with K^+. CD spectra were collected after the initial 10-s folding up to 10^4 s. SVD and fitting analyses in this time region gave a time constant of ~3700 s for $I_2 \rightarrow I_3$ and a faster final folding of $I_3 \rightarrow F$ with a time constant of ~160 s. From SVD analysis, the significant spectral species could be associated with the most likely structure for each intermediate. As noted above, I_2 is likely an antiparallel G4 structure while the SVD results suggest that I_3 is most likely a triple-helix conformation that folds into the final hybrid conformation in K^+ (Figs. 3 and 12 in Ref. [19]).

Kinetic results monitored by FRET and by replacing the 5′ A and the A in each of the three loops with 2AP support the complex kinetic Scheme 1 proposed from CD experiments as well as the proposed structures. All of these methods share relaxation times of ~0.2, 21 and 1300 s. The unfolded state gives an ensemble of dynamic hairpin duplexes that can fold in a direct manner into antiparallel quadruplex structures (such as the "chair" conformation in Fig. 3). Although the antiparallel conformation cannot be exactly specified, the chair conformation seems most likely since such structures can be formed by simple folding of initial hairpins. Opening of the chair from either the 5′ or the 3′ end can give the proposed triple-helical structures. The triple helices allow refolding into the alternative orientations of either Hybrid-1 or Hybrid-2 [19]. Opening of the relatively stable antiparallel structure (I_2) to give the triple helix has high activation energy and represents a significant kinetic barrier to formation of the final folded hybrid quadruplex. The *anti–syn* conversions of some nucleotides required for the final structures are also analogous to the *cis–trans* proline conversions required in folding of some proteins. The unfolding of the hybrid quadruplex structures was followed by trapping the unfolded state with the complementary C-rich strand since formation of the trapped duplex is faster than unfolding of the quadruplex *F* state. Complex, multiphasic kinetics was also observed in unfolding and generally supports the proposed mechanism in Scheme 1 and Fig. 3.

To test their proposed conformational scheme, the authors conducted constrained simulation studies of the folding process and they present movies of the folding process in Supplementary Materials [19].

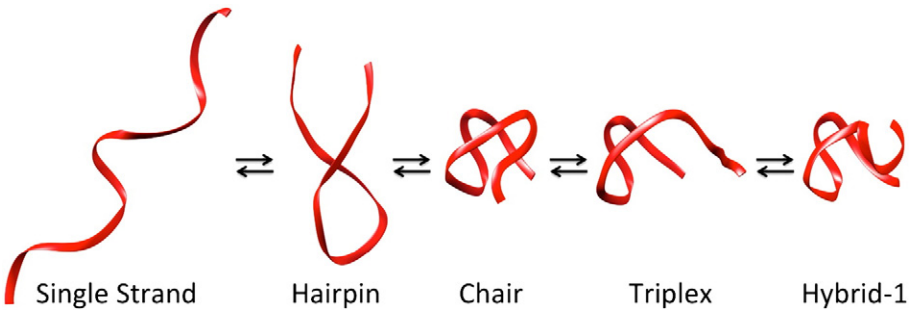


Fig. 3. Proposed folding mechanisms for the formation of the human telomere G-quadruplex.

These simulations provide strong support for the proposed conformational conversions and give very useful visual assistance in understanding the possible folding steps in the kinetic pathway to the final quadruplex. It will be of great interest to conduct similar detailed kinetics studies on DNA quadruplexes of different structure from Htelo.

References

[1] Blackburn EH. Structure and function of telomers. Nature 1991;350:569–73.

[2] Rhodes D, Giraldo R. Telomere structure and function. Curr Opin Struct Biol 1995;5:311–22.

[3] O'Sullivan RJ, Karlseder J. Telomeres: protecting chromosomes against genome instability. Nat Rev Mol Cell Biol 2010;11:171–81.

[4] Cech TR. Life at the end of chromosome: telomers and telomerase. Angew Chem Int Ed 2000;39:34–43.

[5] Blackburn EH, Szostak JW. The molecular structure of centromers and telomers. Annu Rev Biochem 1984;53:163–94.

[6] Meyne J, Ratliff RL, Moyzis RK. Conservation of the human telomere sequence (TTAGGG)n among vertebrates. Proc Natl Acad Sci U S A 1989;86:7049–53.

[7] Hardin CC, Henderson E, Watson T, Prosser JK. Monovalent cation induced structural transitions in telomeric DNAs: G-DNA folding intermediates. Biochemistry 1991;30:4460–72.

[8] Gray RD, Chaires JB. Kinetics and mechanism of K^+- and Na^+-induced folding of models of human telomeric DNA into G-quadruplex structures. Nucleic Acids Res 2008;36:4191–203.

[9] Jixun D, Megan C, Chandanamali P, Roger AJ, Danzhou Y. 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 2007;35:4927–40.

[10] Yang D, Okamoto K. Structural insights into G-quadruplexes: towards new anticancer drugs. Future Med Chem 2010;2:619–46.

[11] Wang Y, Patel DJ. Solution structure of the human telomeric repeat D[Ag3(T2Ag3)3] G-tetraplex. Structure 1993;1:263–82.

[12] Parkinson GN, Lee MP, Neidle S. Crystal structure of parallel quadruplexes from human telomeric DNA. Nature 2002;417:876–80.

[13] Li J, Correia JJ, Wang L, Trent JO, Chaires JB. Not so crystal clear: the structure of the human telomere G-quadruplex in solution differs from that present in a crystal. Nucleic Acids Res 2005;33:4649–59.

[14] Dai JX, Carver M, Yang DZ. Polymorphism of human telomeric quadruplex structures. Biochimie 2008;90:1172–83.

[15] Buscaglia R, Miller MC, Dean WL, Gray RD, Lane AN, Trent JO, et al. Polyethylene glycol binding alters human telomere G-quadruplex structure by conformational selection. Nucleic Acids Res 2013;41:7934–46.

[16] Phan AT, Luu KN, Patel DJ. Different loop arrangements of intramolecular human telomeric (3 + 1) G-quadruplexes in K^+ solution. Nucleic Acids Res 2006;34:5715–9.

[17] Zhang Z, Dai J, Veliath E, Jones RA, Yang D. Structure of a two-G-tetrad intramolecular G-quadruplex formed by a variant human telomeric sequence in K^+ solution: insights into the interconversion of human telomeric G-quadruplex structures. Nucleic Acids Res 2010;38:1009–21.

[18] Ambrus A, Chen D, Dai J, Bialis T, Jones RA, Yang D. Human telomeric sequence forms a hybrid-type intramolecular G-quadruplex structure with mixed parallel/antiparallel strands in potassium solution. Nucleic Acids Res 2006;34:2723–35.

[19] Gray RD, Trent JO, Chaires JB. Folding and unfolding pathways of the human telomeric G-quadruplex. J Mol Biol 2014;426. http://dx.doi.org/10.1016/j.jmb.2014.01.009.

[20] Mashimo T, Yagi H, Sannohe Y, Rajendran A, Sugiyama H. Folding pathways of human telomeric type-1 and type-2 G-quadruplex structures. J Am Chem Soc 2010;132:14910–8.

[21] Biffi G, Tannahill D, McCafferty J, Balasubramanian S. Quantitative visualization of DNA G-quadruplex structures in human cells. Nat Chem 2013;5:182–6.

[22] Karsisiotis AI, Hessari NM, Novellino E, Spada GP, Randazzo A, Webba da Silva M. Topological characterization of nucleic acid G-quadruplexes by UV absorption and circular dichroism. Angew Chem Int Ed 2011;50:10645–8.

[23] Torimura M, Kurata S, Yamada K, Yokomaku T, Kamagata Y, Kanagawa T, et al. Fluorescence-quenching phenomenon by photoinduced electron transfer between a fluorescent dye and a nucleotide base. Anal Sci 2001;17:155–60.

[24] Gray RD, Petraccone L, Buscaglia R, Chaires JB. 2-Aminopurine as a probe for quadruplex loop structures. Methods Mol Biol 2010;608:121–36.