Supporting Material

G-quadruplex and i-motif are mutually exclusive in double-stranded ILPR DNA

Soma Dhakal, Zhongbo Yu, Ryan Konik, Yunxi Cui, Deepak Koirala, and Hanbin Mao*

Department of Chemistry & Biochemistry, Kent State University, Kent, Ohio 44242, USA

*To whom correspondence should be addressed. Email: hmao@kent.edu

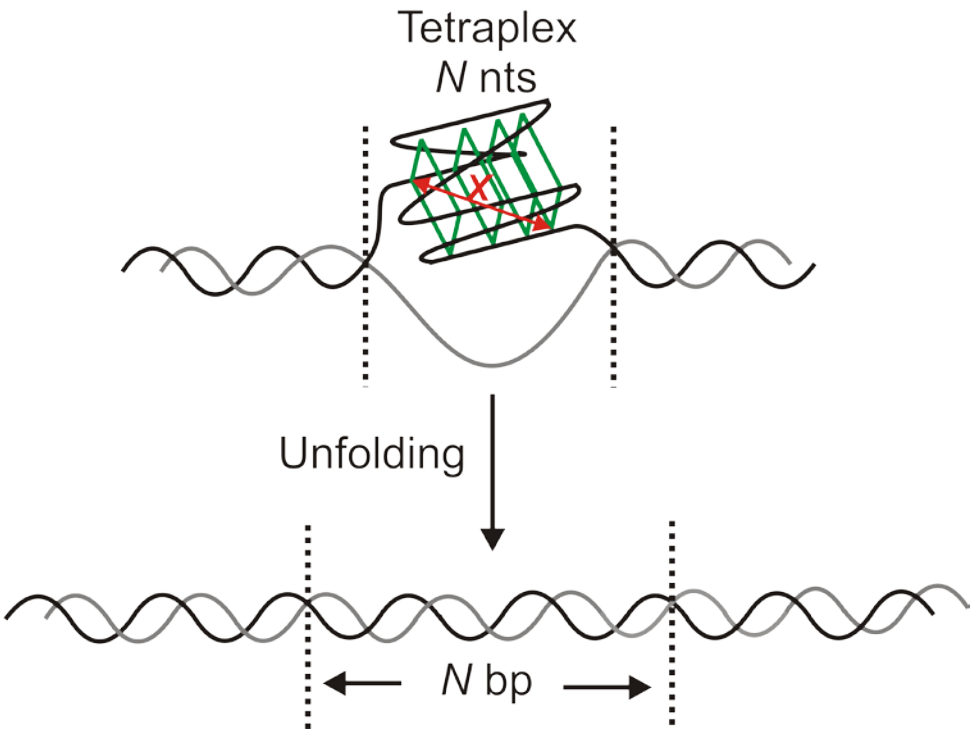

FIGURE S1. Calculation of the number of nucleotides (*N*) involved in the tetraplex structures. The diagram shows that unfolding of a tetraplex structure, a G-quadruplex for example (*top*, *green frames* represent the G-quartets), leads to a dsDNA (*bottom*). The red double-head arrow represents the end-to-end distance (*X*) for the structure. The number of nucleotides (*N*) involved in the structure is calculated using the following equation (1-4),

$$N * L_{bp} - X = \Delta L \quad \ldots\ldots\ldots\ldots\ldots\ldots (1)$$

where *N* is the total number of nucleotides involved in the tetraplex structure, L_{bp} is the contour length of each base pair (bp) in the B form of dsDNA (0.34 nm) (5), *X* is the end-to-end distance, and ΔL is the change in contour length due to the unfolding of the structure.

With ΔL = 6.7 (± 0.2) nm (Table 1) and an estimation of *X* from various G-quadruplex structures after counting for different number of G-quartet stacks (*X* = 1.5 nm for the parallel G-quadruplex [PDB code, 1KF1]; *X* = 1.6 nm for the hybrid-1 and hybrid-2 quadruplexes [PDB codes, 2HY9 and 2GKU]; and *X* = 2.1 nm for the basket type quadruplex [PDB code, 143D]) (6-9), we obtained *N* as 24 ± 1, 25 ± 1, and 26 ± 1 nts for respective structures. Similarly, with ΔL = 7.2 (± 0.1) nm (Table 1) and an estimation of *X* from various i-motif structures (PDB codes, 1ELN, 1A83 and 1YBR, *X* = 0.8 nm for average narrow groove distance and *X* =1.3 nm for average wide groove distance) (10-12), we obtained *N* as 24 ±1 and 25 ±1 nts for respective structures. These calculations confirmed the formation of fully folded tetraplex structures in the ILPR duplex.

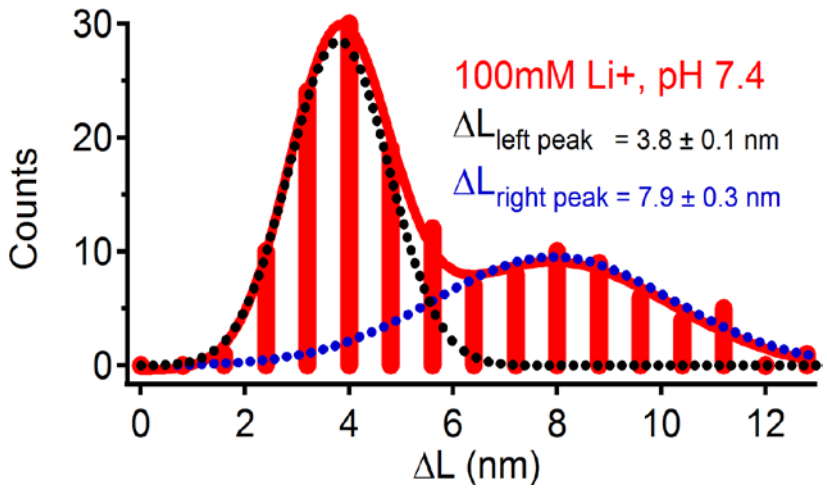

FIGURE S2. Histogram of change in contour length (ΔL) in a pH 7.4 Tris buffer with 100 mM Li^+ at 23 °C. The histogram was fitted with a two-peak Gaussian function (*red solid curve*) and further deconvoluted randomly into left (*black-dotted curve*) and right (*blue-dotted curve*) populations (see Materials and Methods). These two populations represent a partially folded (*left*) and a fully folded (*right*) species.

TABLE S1. The percentage of mechanical unfolding events for the control experiments in different buffers at 23 °C.

Buffer/pH	Ion	DNA Construct	Unfolding Events (%)
10 mM Tris/pH 7.4	100 mM K^+	dsDNA handles only	1.9 (Fully Folded)
10 mM MES/pH 5.5	100 mM Li^+	dsDNA handles only	1.0 (Part. Folded) 1.6 (Fully Folded)
10 mM Tris/pH 7.4	100 mM Li^+	dsDNA with G/C rich sequences	3.1 (Part. Folded) 3.2 (Fully Folded)

As shown in Table S1, we observed rare unfolding events (< 2%) when a DNA construct without G-quadruplex/i-motif forming sequences (dsDNA handles only) was mechanically stretched at pH 5.5 with 100 mM Li^+ or at pH 7.4 with 100 mM K^+. When a DNA construct with ILPR G-quadruplex/i-motif forming sequence was used, ~ 3% unfolding events were observed for partially or fully folded populations (Figure S3) at pH 7.4 with 100 mM Li^+. These results were consistent with the DMS footprinting of a 87-bp dsDNA in the same buffer where no protection of the G4 tracts (Figure 1*A*, *lane 7*) or the C4 tracts (Fig.1*B*, *lane 1*) was observed. As a comparison, 18% unfolding events were observed at pH 7.4 with 100 mM K^+, 44% were observed at pH 5.5 with 100 mM Li^+ (10% partially folded and 34% fully folded, see Figure 2*C*, *middle panel*, and Table 1); and 33% were observed at pH 5.5 with 100 mM K^+ (14% *left peak population* and 19% *right peak population*, see Figure 2*D*, *bottom panel*, and Table 1). These results confirmed the formation of G-quadruplex or i-motif in the DNA construct containing the ILPR G-quadruplex/i-motif forming sequence.

SUPPORTING REFERENCES

1. Yu, Z., Schonhoft, J.D., Dhakal, S., Bajracharya, R., Hegde, R., Basu, S. and H. Mao. 2009. ILPR G-Quadruplexes Formed in Seconds Demonstrate High Mechanical Stabilities. *J. Am. Chem. Soc.* 131: 1876-1882.
2. Dhakal, S., Schonhoft, J.D., Koirala, D., Yu, Z., Basu, S. and H. Mao. 2010. Coexistence of an ILPR i-Motif and a Partially Folded Structure with Comparable Mechanical Stability Revealed at the Single-Molecule Level. *J. Am. Chem. Soc.* 132: 8991–8997.
3. Dietz, H. and M. Rief. 2004. Exploring the energy landscape of GFP by single-molecule mechanical experiments. *Proc. Natl. Acad. Sci. USA*. 101: 16192-16197.
4. Greenleaf, W.J., Frieda, K.L., Foster, D.A.N., Woodside, M.T. and S.M. Block. 2008. Direct Observation of Hierarchical Folding in Single Riboswitch Aptamers. *Science*. 319: 630-633.
5. Sinden, R.R. 1995. DNA Structure and Function. Academic Press, San Diego, CA.
6. Parkinson, G.N., Lee, M.P.H. and S. Neidle. 2002. Crystal structure of parallel quadruplexes from human telomeric DNA. *Nature*. 417: 876-880.
7. Dai, J., Carver, M., Punchihewa, C., Jones, R.A. and D. Yang. 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.* 35: 4927-4940.
8. Luu, K.N., Phan, A.T., Kuryavyi, V., Lacroix, L. and D.J. Patel. 2006. Structure of the human telomere in K+ solution: an intramolecular (3 + 1) G-quadruplex scaffold. *J. Am. Chem. Soc.* 128: 9963-9970.
9. Wang, Y. and D.J. Patel. 1993. Solution structure of the human telomeric repeat d[AG3(T2AG3)3] G-tetraplex. *Structure*. 1: 263-282.
10. Phan, A.T., Gueron, M. and J.L. Leroy. 2000. The Solution Structure and Internal Motions of a Fragment of the Cytidine-rich Strand of the Human Telomere. *J. Mol. Biol.* 299: 123-144.
11. Han, X., Leroy, J.L. and M. Gueron. 1998. An intramolecular i-Motif: the Solution Structure and Base-pair Opening Kinetics of d(5mCCT3CCT3ACCT3CC). *J. Mol. Biol.* 278: 949-965.
12. Esmaili, N. and J.L. Leroy. 2005. i-motif solution structure and dynamics of the d(AACCCC) and d(CCCCAA) tetrahymena telomeric repeats. *Nucleic Acids Res.* 33: 213-224.