Supporting information for "Insight into G-DNA Structural Polymorphism and Folding from Sequence and Loop Connectivity through Free Energy Analysis" *J. Amer. Chem. Soc. (2011),* http://dx.doi.org/10.1021/ja107805r *by Xiaohui Cang, Jiří Šponer, and Thomas E. Cheatham, III.*

Table S1. Free energies (kcal/mol) calculated using the MM-PBSA method for the single-loop models.

model	time interval	G_{total}	G_{stem}	G_{loop}	ΔG_{total}	ΔG_{loop}	–TS (Nmode)	–TS (quasi-harmonic)
T-L_n	40-50ns	-2848.5 ± 2.5	-2647.0 ± 2.3	-201.5 ± 0.6	**7**	0	-354.0	-121.2
T-P_3a	40-50ns	-2855.2 ± 2.0	-2653.7 ± 1.9	-201.5 ± 0.6	**0**	0	-355.1	-124.5
T-P_3p	40-50ns	-2829.4 ± 2.0	-2626.3 ± 1.9	-203.1 ± 0.6	26	**-2**		
T-P_2a	11-21ns	-1790.9 ± 1.5	-1600.4 ± 1.5	-190.5 ± 0.7		11		
T-P_2a	35-45ns	-1790.4 ± 1.7	-1599.4 ± 1.4	-191.0 ± 0.7		**11**		
T-P_4a†	40-50ns	-3904.1 ± 2.6	-3694.2 ± 2.5	-209.9 ± 0.7		-8		
T_2-L_n	40-50ns	-3003.0 ± 2.0	-2657.6 ± 1.9	-345.4 ± 0.8	**2**	3	-377.7	-127.3
T_2-L_w*	40-50ns	-3007.2 ± 1.8	-2652.8 ± 1.6	-354.4 ± 0.9	-3	-6		
T_2-L_w-2	40-50ns	-3000.5 ± 2.3	-2648.2 ± 2.2	-352.3 ± 0.9	**4**	-4	-377.1	-126.5
T_2-P_3a	40-50ns	-3004.6 ± 2.1	-2655.9 ± 2.2	-348.7 ± 0.9	**0**	0	-380.4	-124.1
T_2-P_3p	15-25ns	-2978.6 ± 2.1	-2634.0 ± 1.7	-344.6 ± 2.4	26	**4**		
T_2-P_3p	38-48ns	-2972.1 ± 2.5	-2630.2 ± 2.2	-341.9 ± 0.9	33	7		
T_2-P_2a	5-15ns	-1938.3 ± 2.0	-1604.2 ± 1.7	-334.1 ± 0.9		15		
T_2-P_2a	18-26ns	-1936.8 ± 1.8	-1601.6 ± 1.7	-335.2 ± 0.8		**13**		
T_2-P_2a	40-50ns	-1934.4 ± 2.2	-1599.4 ± 2.0	-335.0 ± 1.0		14		
T_2-P_4a	40-50ns	-4054.2 ± 2.2	-3704.0 ± 2.0	-350.2 ± 0.9		-2		
T_2-P_4p	15-25ns	-4043.4 ± 2.9	-3696.5 ± 2.6	-346.9 ± 0.9		**2**		
T_3-L_n	40-50ns	-3150.7 ± 2.5	-2656.7 ± 2.2	-494.0 ± 1.1	**-5**	-3	-401.8	-129.7
T_3-L_w	17-27ns	-3149.7 ± 2.5	-2652.3 ± 2.1	-497.4 ± 1.2	-4	-6		
T_3-L_w	40-50ns	-3155.2 ± 2.1	-2657.9 ± 2.0	-497.3 ± 1.0	**-9**	-6	-402.0	-128.6
T_3-L_w-2	40-50ns	-3151.0 ± 2.8	-2658.0 ± 2.4	-493.0 ± 1.3	-5	-2	-401.3	
T_3-D	40-50ns	-3145.4 ± 2.5	-2651.9 ± 2.1	-493.5 ± 1.2	**1**	-3	-403.3	-132.1
T_3-P_3a	40-50ns	-3145.9 ± 2.0	-2655.0 ± 1.9	-490.9 ± 1.0	**0**	0	-403.8	-126.8
T_3-P_3p	40-50ns	-3124.4 ± 2.4	-2634.0 ± 2.3	-490.4 ± 1.0	22	**1**		
T_3-P_2a	30-40ns	-2080.8 ± 1.9	-1602.4 ± 1.7	-478.4 ± 1.0		**13**		
T_3-P_4a	40-50ns	-4191.9 ± 2.5	-3702.4 ± 2.2	-489.5 ± 1.2		**1**		

T_3-P_4p	40-50ns	-4182.9	±	2.6	-3691.0	±	2.4	-491.9	±	1.2		**-1**
TTA-L_n	40-50ns	-3195.7	±	2.4	-2653.9	±	2.3	-541.8	±	1.1	**-4**	-8
TTA-L_w	20-30ns	-3196.0	±	2.0	-2652.7	±	1.8	-543.3	±	1.0	-4	-10
TTA-L_w	40-50ns	-3197.4	±	2.3	-2654.9	±	2.0	-542.5	±	1.1	**-5**	-9
TTA-D	30-40ns	-3190.3	±	1.8	-2650.0	±	1.6	-540.3	±	1.3	**2**	-7
TTA-P_3a	40-50ns	-3192.2	±	2.4	-2658.7	±	2.2	-533.5	±	1.1	**0**	0
TTA-P_3p	30-40ns	-3159.4	±	2.4	-2629.0	±	2.2	-530.4	±	1.1	33	**3**
TTA-P_2a	40-50ns	-2128.9	±	2.2	-1604.4	±	1.9	-524.5	±	1.2		**9**
TTA-P_4a	40-50ns	-4238.5	±	2.8	-3701.9	±	2.7	-536.6	±	1.0		**-3**
TTA-P_4p	40-50ns	-4215.3	±	2.6	-3687.4	±	2.0	-527.9	±	1.4		**6**
T_4-L_n	40-50ns	-3291.1	±	2.4	-2656.0	±	2.2	-635.1	±	1.2	**-7**	-6
T_4-L_n-2	40-50ns	-3286.5	±	2.2	-2653.0	±	1.8	-633.5	±	1.1	-3	-4
T_4-L_w	40-50ns	-3303.2	±	2.3	-2660.4	±	1.9	-642.8	±	1.2	**-19**	-14
T_4-L_w-2	40-50ns	-3298.7	±	2.3	-2657.2	±	2.0	-641.5	±	1.1	-15	-12
T_4-L_w-3*	40-50ns	-3307.8	±	2.0	-2660.4	±	1.9	-647.4	±	1.1	-24	-18
T_4-D	40-50ns	-3293.7	±	2.8	-2659.3	±	2.5	-634.4	±	1.1	**-10**	-5
T_4-P_3a	40-50ns	-3283.9	±	2.1	-2654.7	±	1.9	-629.2	±	1.1	**0**	0
T_4-P_3p	30-40ns	-3264.1	±	2.1	-2634.7	±	2.0	-629.4	±	1.3	20	0
T_4-P_3p	40-50ns	-3262.1	±	2.6	-2630.0	±	2.3	-632.1	±	1.3	22	**-3**
T_4-P_2a	40-50ns	-2230.0	±	1.9	-1602.7	±	1.5	-627.3	±	1.2		**2**
T_4-P_4a	40-50ns	-4340.1	±	2.7	-3703.8	±	2.3	-636.3	±	1.3		**-7**
T_4-P_4p	40-50ns	-4316.0	±	2.5	-3682.7	±	2.2	-633.3	±	1.7		**-4**

The data shown in bold and with underscores are used for relative rankings and are shown in the main manuscript in Table 1.
The T-**L_w**, T-**D**, T-**P_4p** and T_2-**D** models were unstable in the MD simulations, so these models are not included in the table.
All of the non-residue loop models (i.e. where the two strands are directly bonded without any intervening residues) were unstable except 0-**P_3p**, therefore free energies were not calculated for those models.
*: The extreme stability for the over-stacked loop geometry may be an artifact of the current force field (ff99+parmbsc0).
†: The abnormal geometry of the sugar ring of the loop-connecting stem residue makes the ΔG_{loop} value misleading, therefore the data of T-**P_4a** is not shown in the main manuscript Table 1.
Entropies were calculated with both normal mode and quasi-harmonic methods for a subset of the models to demonstrate that the solution entropies do not significantly affect the relative ranking of different loop types.

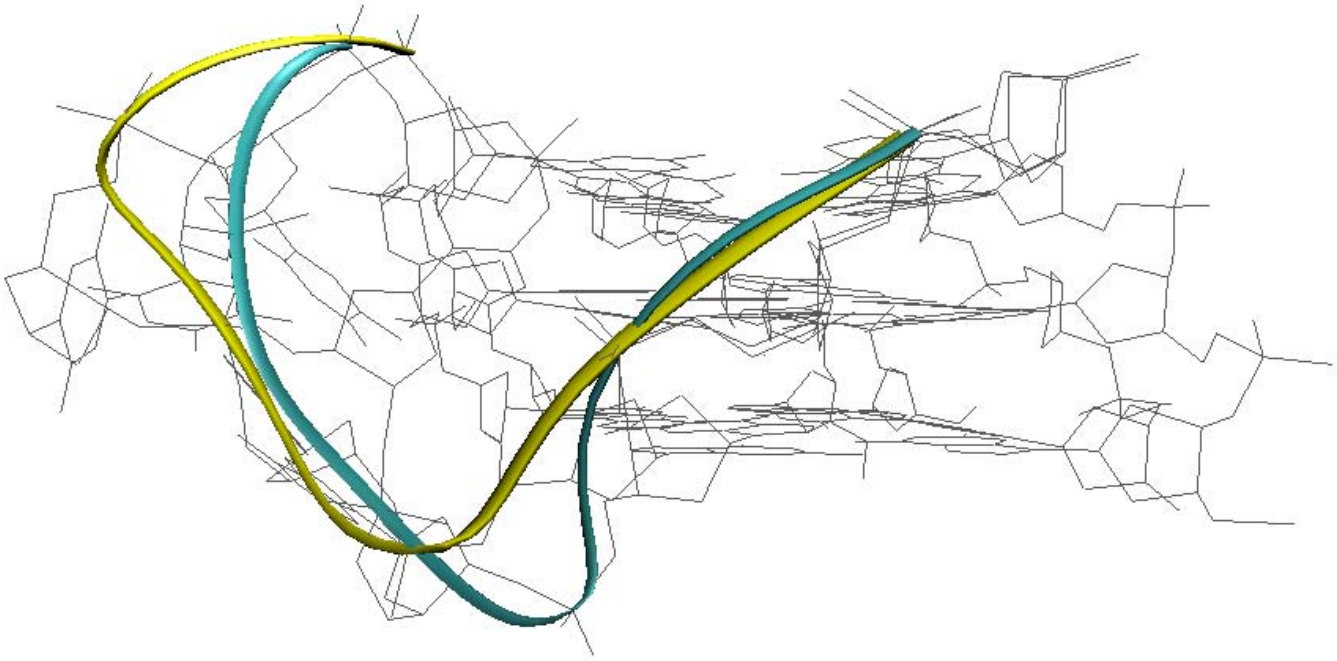

Figure S1: Backbone differences for the parallel and anti-parallel propeller dTTA loops. The loop and strand direction follows that of Figure 2 in the main text. Yellow: the parallel propeller loop (from the crystal structure 1KF1); cyan: the anti-parallel propeller loop (from the NMR structure 2GKU). Compared to the anti-parallel reversal loop, the parallel reversal loop is more flat and the final loop residue is further way from the connecting stem residue.

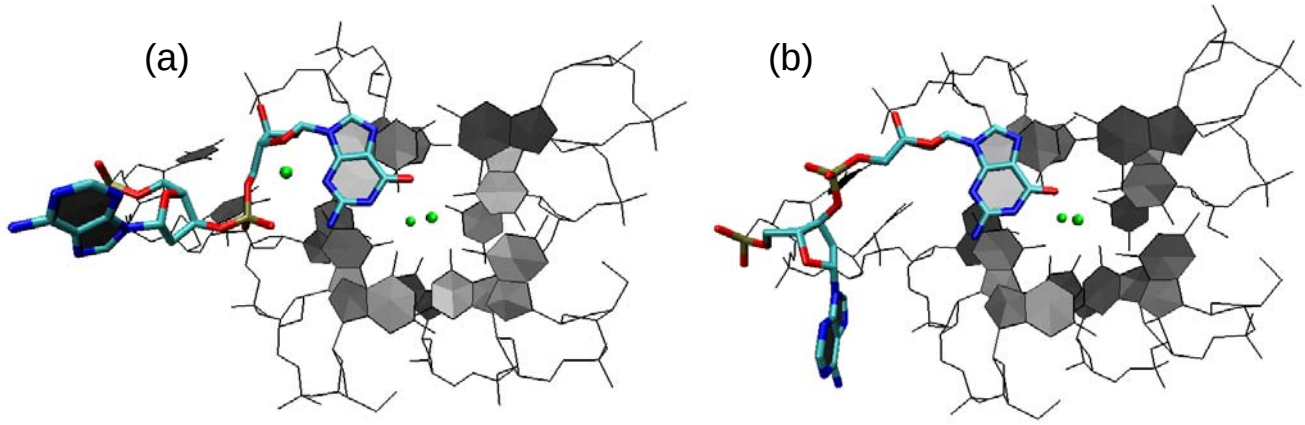

Figure S2. The stable cation binding site specific for the anti-parallel propeller loop models. The cation binding site is formed from the phosphate group, O4' atom and N3 atom of the *syn* dG residue connected with the 3' side of the propeller loop. (a) When the phosphate group is close to the other atoms, the binding site forms; (b) When the loop changed to a geometry that moved the phosphate group away the ion leaves quickly.

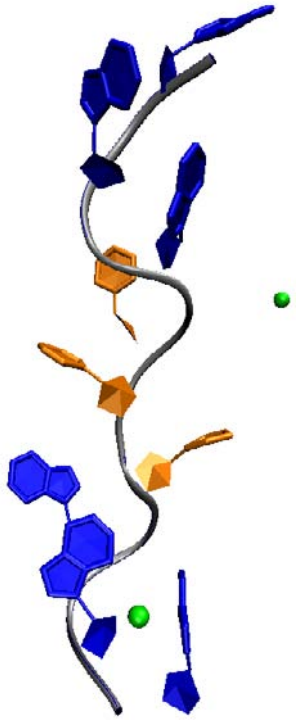

Figure S3: The extended structure of the linear sequence d(GGGTTTGGG) at the end of 500ps MD simulation. The guanine residues are shown in blue; the thymine residues are shown in orange; and the two closest K^+ cations are shown as green spheres.

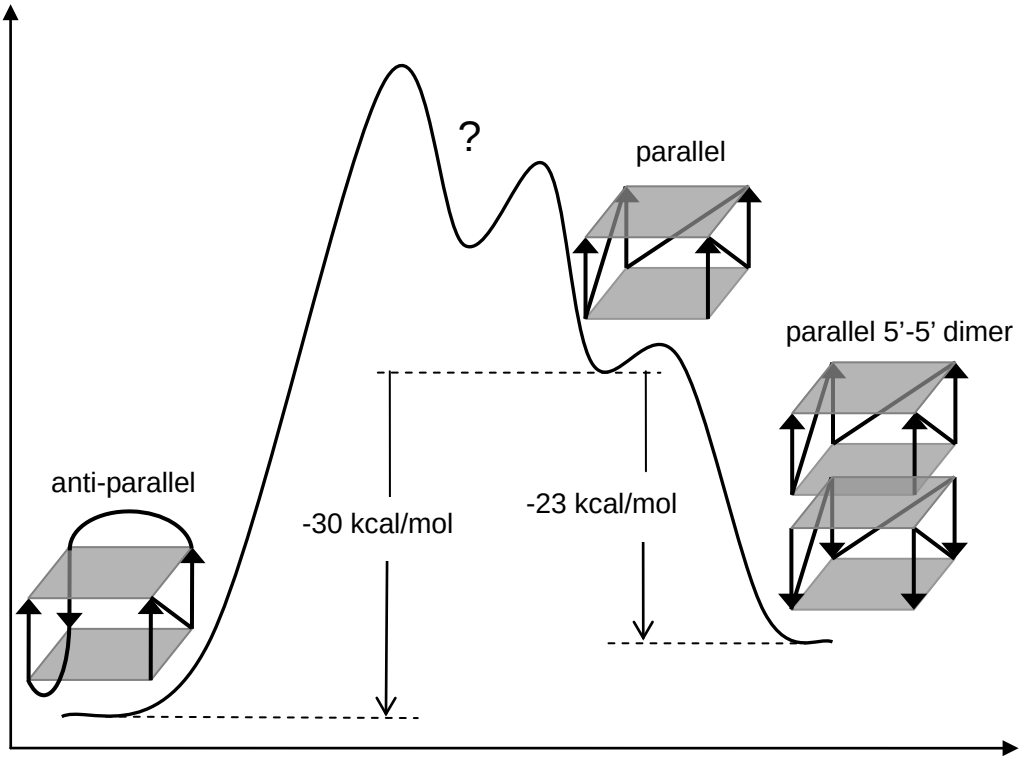

Figure S4. Hypothetical energy curves between anti-parallel monomer, parallel monomer and parallel 5’-5’ dimer of human telomeric G-quadruplexes. When the loops are not very short anti-parallel structures are more stable than parallel structures. However, parallel structures can be favored through multimer formation. The balance between inter-conversions of anti-parallel and parallel structures is still unclear.