Supplementary Figures and Supplementary Tables

FOLDING AND PERSISTENCE TIMES OF INTRAMOLECULAR G-QUADRUPLEXES TRANSIENTLY EMBEDDED IN A DNA DUPLEX

Phong Lan Thao Tran[1], Martin Rieu[2,3], Samar Hodeib[2,3], Alexandra Joubert[1], Jimmy Ouellet[4], Patrizia Alberti[1], Anthony Bugaut[1], Jean-François Allemand[2,3], Jean-Baptiste Boulé[1]*, Vincent Croquette[2,3,5]*

[1] Structure et Instabilité des Génomes, Museum National d'Histoire Naturelle, Alliance Sorbonne Université, INSERM, CNRS, 43 rue Cuvier, 75005 Paris, France.

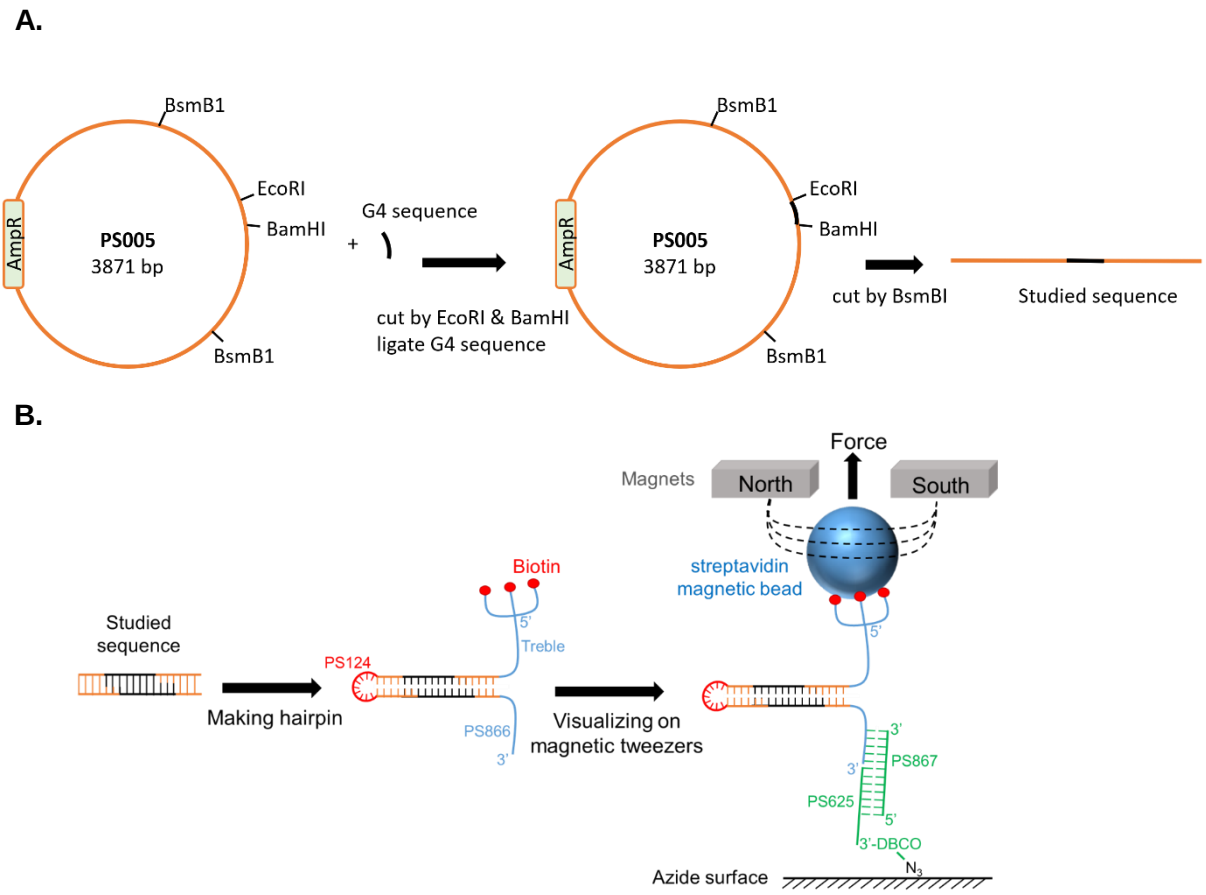
[2] Laboratoire de physique de L'École normale supérieure de Paris, CNRS, ENS, Université PSL, Sorbonne Université, Université de Paris, 75005 Paris, France

[3] Institut de Biologie de l'Ecole Normale Supérieure (IBENS), Ecole normale supérieure, CNRS, INSERM, Université PSL, 75005, Paris, France

[4] Depixus SAS, 3-5 Impasse Reille, 75014 Paris, France

[5] ESPCI Paris, PSL University, 10 rue Vauquelin, 75005 Paris, France

*Corresponding authors: vincent.croquette@lps.ens.fr and jean-baptiste.boule@mnhn.fr

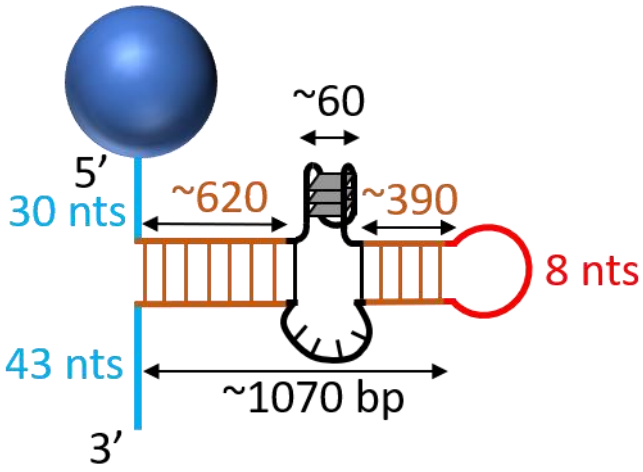


Supplementary Figure S1: Synthesis scheme of the single molecule substrate. A. Cloning of a G4 sequence in the PS005 plasmid. Oligonucleotides containing respectively the G-rich sequences and their complementary strands were incubated at equimolar ratios at 95°C in 1X Tris-EDTA buffer (pH 7) and digested by EcoRI and BamHI. The PS005 vector was digested by the same restriction sites and ligated with the double-stranded insert containing the potential G4 forming sequence (**Supplementary Table 2**). The recombinant plasmid containing the G4 sequence was then amplified in *E. Coli*. The purified plasmid was then digested with BsmB1, which generates a ~1070 kilobase pair DNA fragment containing the G4 motif and non-palindromic overhangs on each end. **B. Diagram of the hairpin synthesis scheme and sequences involved in tethering the DNA hairpin on azide coated glass coverslip.** The double-stranded DNA fragment was ligated on one end with a loop sequence (PS124) and in the other end with a Y-shape molecule resulting from the annealing of DNA oligonucleotide PS 866 to the treble oligonucleotide (**Supplementary Table 1**). The DNA hairpin contains three biotins at the 5'-end which allow a strong binding on a streptavidin-coated magnetic bead. The DNA molecule is tethered to a glass cover slip through annealing to oligonucleotide PS867 at 25°C, which itself is annealed to oligonucleotide PS625. PS625 is covalently linked at its 3' end to the coverslip through a dibenzocyclooctyne group (DBCO)-azide reaction.

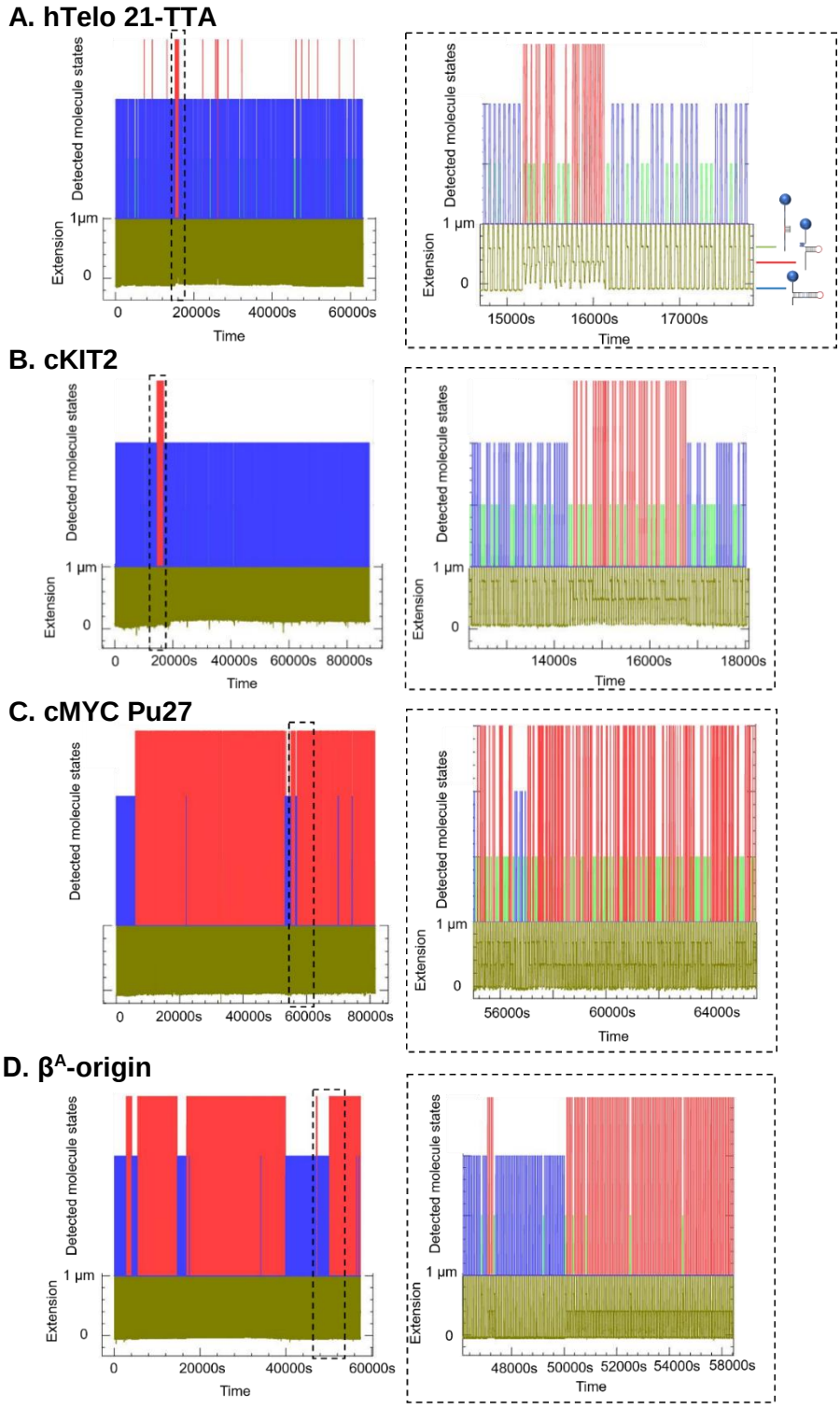
A.

5'BiotinATGGTCAAGCATGCCGCTTTTCGGTTCCCGTGTGTCTTTTGGTCTTTCTGGTGCTCTTCgaatggagAcGAGCTCAGGCCTTAGAGTCAAGACTACAGTAGATACTAATGAC
TGACAGTACCATCAGCATCTCTATATGTAACAGAGTCTTAGAACTGTAGTACTCATGTCATGATCAGATCCATCTTGATGTACTGTCAGACTATCCTCATAAGTCTAGTGAGTACTG
AAGATACATAGCTCAGATCTGATCATACACGACAGGAGTAGAAGTATACACTGAAGTAGATAGATCGATTGACTATGACCATGCTATGACTGACTGCTACTATTCTAGAGTGACA
GTATTACTAAGTACGAGCATCTAACAAGATGAACCTGACAGACACTGACATAGCGATTATCATAGCTAGTCACTCTGACGTATCTCAGTCTAACATCAGCACATGAATGTATGACT
ACAGTTACTATCAAGTACACTCTGAGTCAGCATCATCAGTAGATCGATCAGTACAGTATCTAGGACATGACTACTTGAGTGAGAGAATCTTCTACTATCTGTAGTCAGCTTCGTAA
CACTCTGATGTCATCTTAATGCTATGATTGTCTACTTGATCGACTACTAGATAACTATAGTTCAGCAGTACg----------G-strand (upper strand)----------
gatccgatatcACTAGTGTACACTGATCACATGTACTAGAATACCAGGTTACTATGAATGAGTACGATAGCTATCAGATGTACTACTACTAGATACACGTTGTCATTAGTCACTAGGA
TCAGTACTACTTCGATCACTAGATGTTACAGTCAGAGTCCTCTTAGATTGATACAGTATACTACATGACATATTAGTCCAGTACATACAGACATGATTCACGTGACATTATGAGTAC
CATCATTCAGATCAGTCACCATTACTGTGAGATACTAGCTAGAGTAGACTACTGATCGACTACATGTCAGTCTCAGTACGTTGTATCATAGTCACATTACATCATGTATGCTATTCA
GATTGAGTTAACGGTACCGTCTcagcttGCACTGAGAGCGCGGCCTCTCAGTGCtgAGACGGTACCGTTAACTCAATCTGAATAGCATACATGATGTAATGTGACTATGATACAACG
TACTGAGACTGACATGTAGTCGATCAGTAGTCTACTCTAGCTAGTATCTCACAGTAATGGTGACTGATCTGAATGATGGTACTCATAATGTCACGTGAATCATGTCTGTATGTACT
GGACTAATATGTCATGTAGTATACTGTATCAATCTAAGAGGACTCTGACTGTAACATCTAGTGATCGAAGTAGTACTGATCCTAGTGACTAATGACAACGTGTATCTAGTAGTAGT
ACATCTGATAGCTATCGTACTCATTCATAGTAACCTGGTATTCTAGTACATGTGATCAGTGTACACTAGTgatatcg----------C-strand (lower strand)----------
aattcGTACTGCTGAACTATAGTTATCTAGTAGTCGATCAAGTAGACAATCATAGCATTAAGATGACATCAGAGTGTTACGAAGCTGACTACAGATAGTAGAAGATTCTCTCACTCA
AGTAGTCATGTCCTAGATACTGTACTGATCGATCTACTGATGATGCTGACTCAGAGTGTACTTGATAGTAACTGTAGTCATACATTCATGTGCTGATGTTAGACTGAGATACGTCA
GAGTGACTAGCTATGATAATCGCTATGTCAGTGTCTGTCAGGTTCATCTTGTTAGATGCTCGTACTTAGTAATACTGTCACTCTAGAATAGTAGCAGTCAGTCATAGCATGGTCAT
AGTCAATCGATCTATCTACTTCAGTGTATACTTCTACTCCTGTCGTGTATGATCAGATCTGAGCTATGTATCTTCAGTACTCACTAGACTTATGAGGATAGTCTGACAGTACATCAA
GATGGATCTGATCATGACATGAGTACTACAGTTCTAAGACTCTGTTACATATAGAGATGCTGATGGTACTGTCAGTCATTAGTATCTACTGTAGTCTTGACTCTAAGGCCTGAGCT
CgTctccattcGAAGAGCACCAGAAAGACCAAAAGACACAAAGGGTCAGTGCTGCAACCCACTTCCTAATCTGTCATCTTCTG-3'

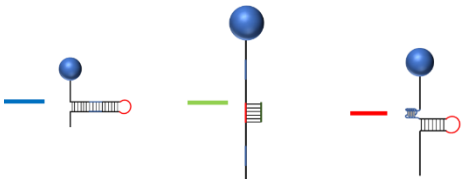
B.



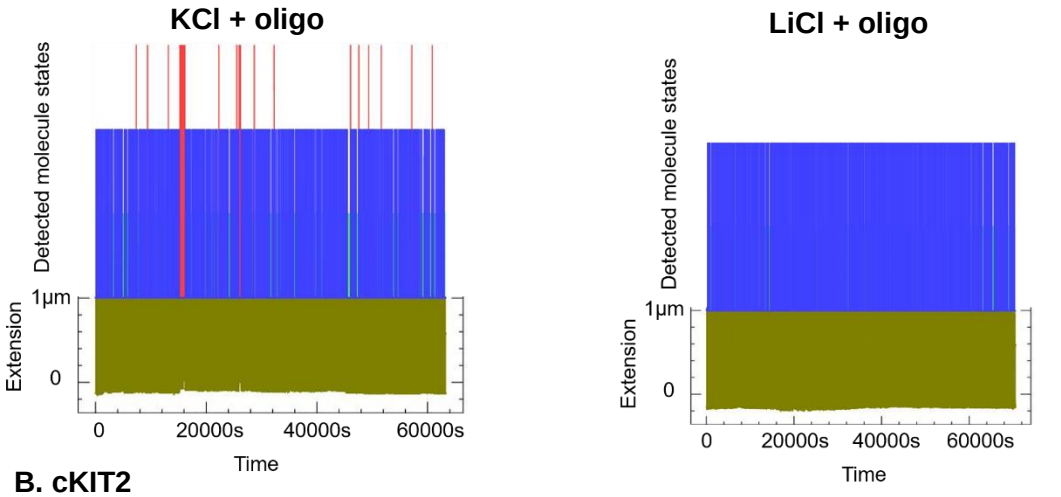
Supplementary Figure S2: A. Sequence of the DNA hairpin scaffold. ssDNA arms are shown in blue. The hairpin ssDNA loop sequence is shown in red. Constant dsDNA regions are shown in brown. Variable DNA regions containing inserted G4 forming sequences are shown in black. **B. Schematics of the dsDNA hairpin molecule.**



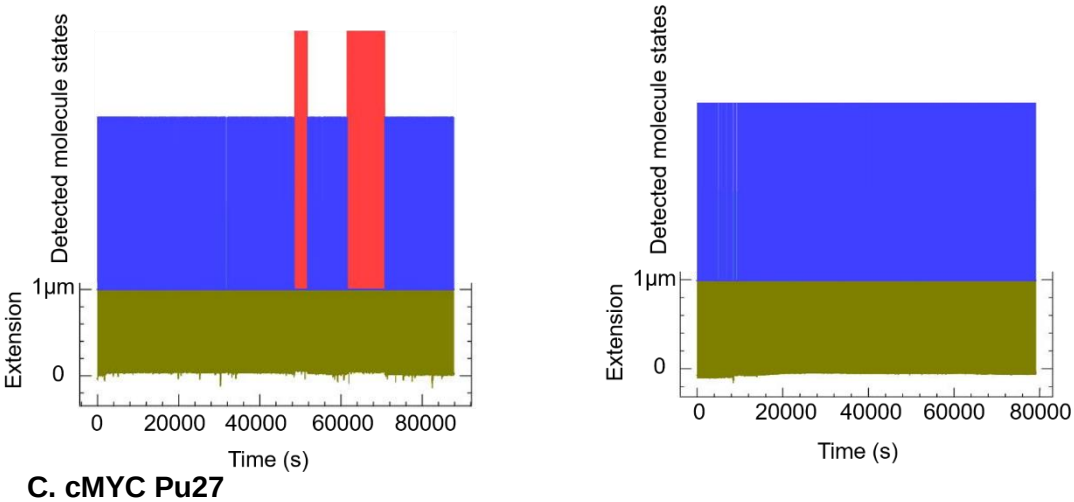
Supplementary Figure S3: Automatic detection of molecular states from full-time courses of recorded DNA extension traces obtained with hTelo 21-TTA (A), cKIT-2 (B), cMYC Pu27 (C) and β^A origin (D) in 100 mM KCl plus oligonucleotide. The left panel shows the full course of an experiment from one bead. Zoom on the dashed timeframe in the left panel is shown on the right. Traces of DNA extension are shown in khaki green. Blockages of the hairpin by a G4 structure or by the binding of oligonucleotide to its loop were automatically detected for each cycle. During hairpin refolding at 7 pN (F_{hold}), three detected molecule states are: complete refolding of the hairpin (0 µm extension state is indicated in blue), binding of blocking oligonucleotide (0.75 µm extension state is in green) and G4 folding (0.45 µm extension state is indicated in red).



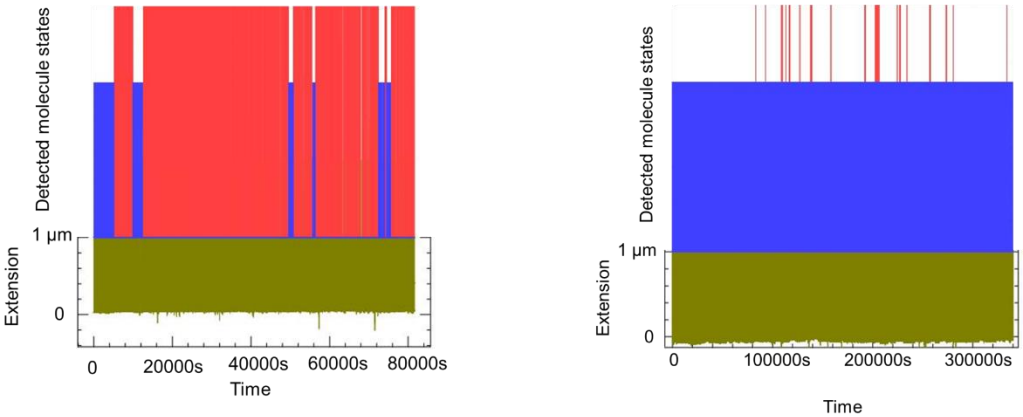
A. hTelo 21-TTA



B. cKIT2

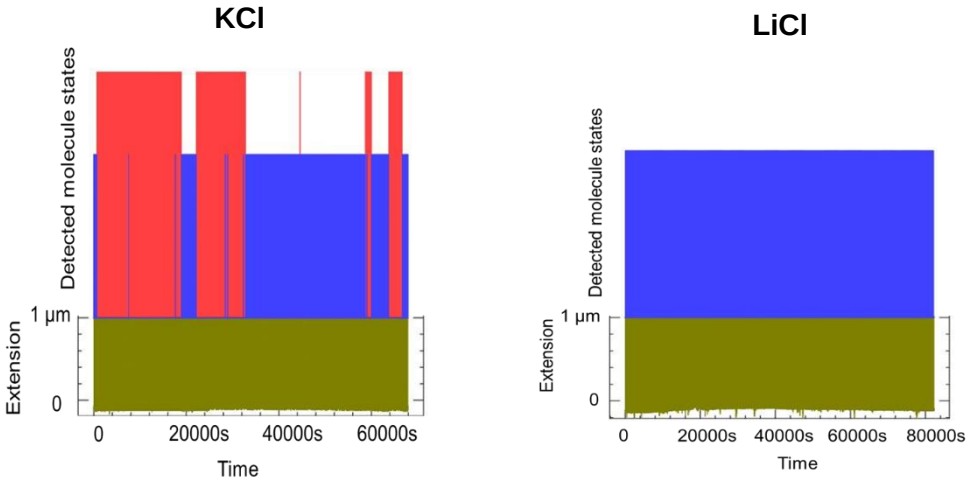


C. cMYC Pu27

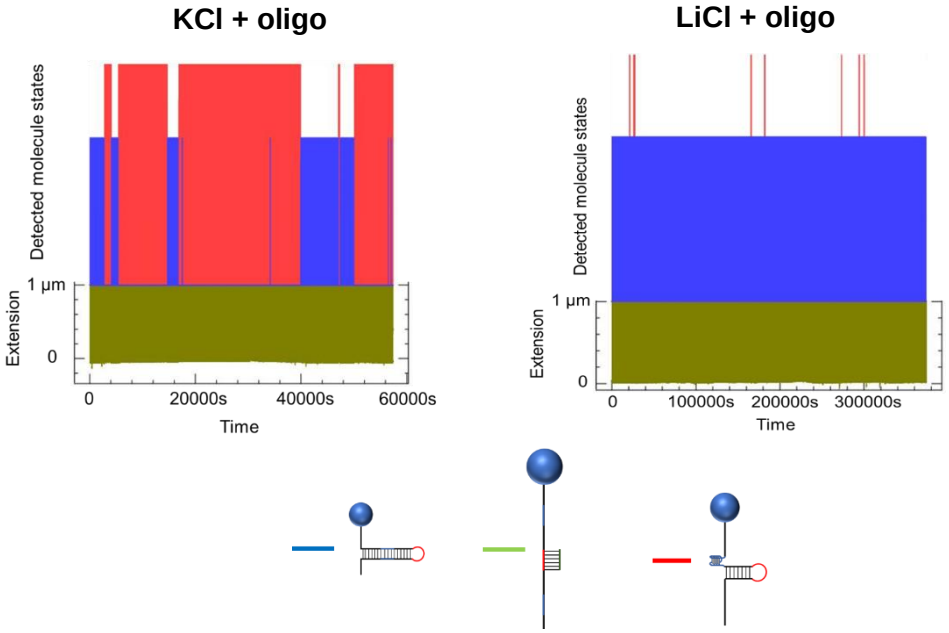


Supplementary Figure S4: Automatic analyses of molecular states from full-time courses of recorded DNA extension traces obtained with hTelo 21-TTA (**A**), cKIT-2 (**B**), cMYC Pu27 (**C**) in different ion conditions: KCl + oligo (**left panel**) and LiCl + oligo (**right panel**). Three detected molecule states are: complete refolding of the hairpin (in blue), binding of blocking oligonucleotide (in green) and G4 folding (in red).

A. In the absence of blocking oligonucleotide



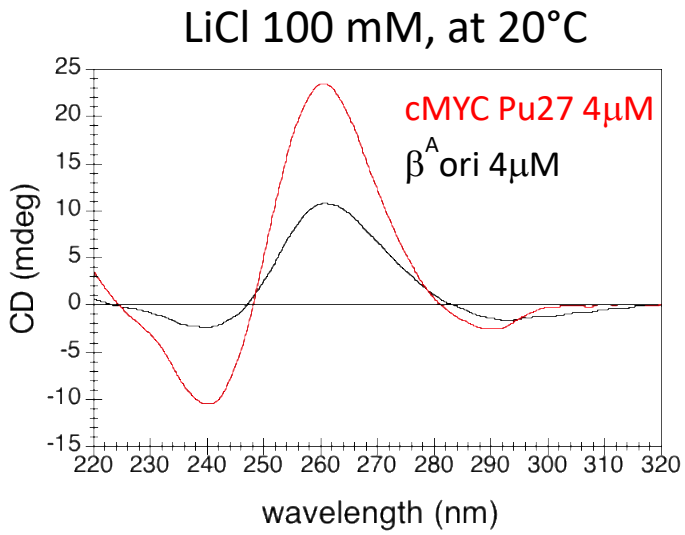
B. In the presence of blocking oligonucleotide



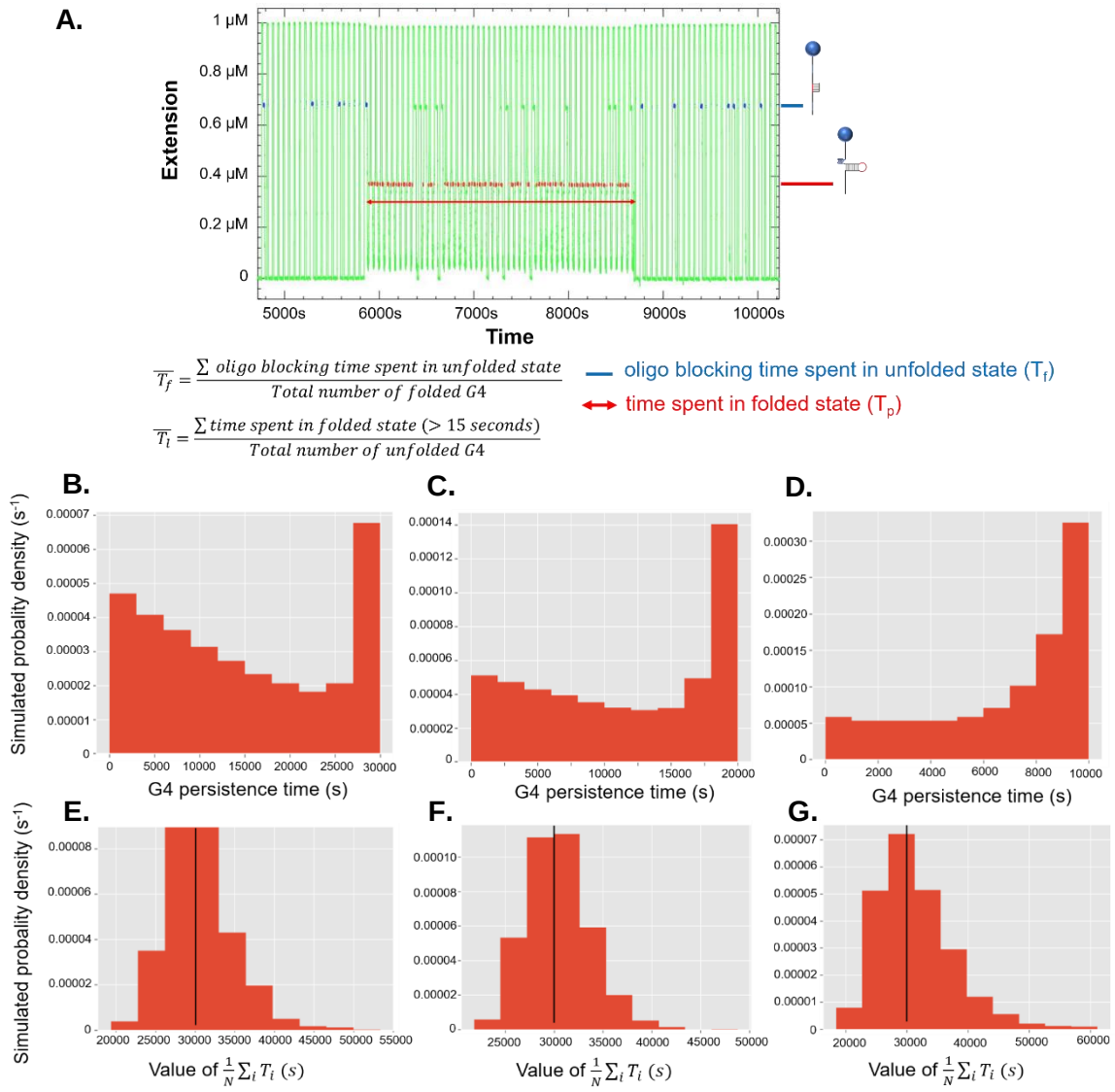
Supplementary Figure S5: Automatic analyses of molecular states from full-time courses of recorded DNA extension traces obtained with ori β^A sequence in different ion conditions: KCl (**left panel**) and LiCl (**right panel**) without (**A**) or with (**B**) blocking oligonucleotide. Three detected molecule states are: complete refolding of the hairpin (in blue), binding of blocking oligonucleotide (in green) and G4 folding (in red).

cMYC Pu27 : 5'-FAM-TGGGGAGGGTGGGGAGGGTGGGGAAGG-TAMRA-3'

β^A origin : 5'-FAM-GGGGGGGGGGGGGCGGG-TAMRA-3'



Supplementary Figure S6: CD (Circular Dichroïsm) spectra at 20°C of cMYC Pu27 (in red) and β^A origin sequences (in black) in 100 mM LiCl at 4 µM oligonucleotide strand concentration. The CD spectra indicate the formation of a parallel G4 conformation for both G4 sequences.



Supplementary Figure S7: Calculation of $\overline{T_f}$ and $\overline{T_p}$.
A. Diagram explaining the calculation method of the mean of G4 folding ($\overline{T_f}$) and G4 persistence times ($\overline{T_p}$). B-G. Simulations illustrating the impact of the experiment time of $\overline{T_f}$ and $\overline{T_p}$. Experiments are simulated on 100 beads by taking single exponential distributions for the probabilities of folding ($\overline{T_f}$ = 100 s) /unfolding ($\overline{T_p}$ = 30000s) of the G-quadruplex as well as for the probabilities of binding ($\overline{T_{bound}}$ = 30 s)/unbinding ($\overline{T_{unbound}}$ = 10s) of the oligonucleotide.

The distribution of G4 lifetimes averaged over 1000 experiments are drawn for different measurement times in panel **B** (T_{exp} = 10000s), **C** (T_{exp} = 20000s) and **D** (T_{exp} = 30000s). These figures show the impact of a short experiment time on the observed T_p distributions, when the experiment time is smaller or comparable to the persistence time of a given G-quadruplex.

The panel **E, F** and **G** show the corresponding distribution over 1000 experiments of the resulting $\overline{T_p}$ as inferred on one experiment by using the method in the paper, that is by dividing the whole time where a G-quadruplex is observed by the number of unfolding events (described above in **A**). They validate the accuracy of the inference process even in cases when the T_p is larger than the experiment time. Black lines represent the simulated $\overline{T_p}$ of the G4 (30000s), i.e. the correct value to be inferred. Standard deviation of the inferred value is respectively 3250s, 4310s and 6500s. The error of the inference increases as the measurement time decreases because of the corresponding decrease of the number of observed events.

Supplementary Figure S8: Error estimations of folding and persistence times

The method used in this paper involves a periodic testing of the presence of a G-quadruplex in the substrate every cycle, each cycle lasting a time T_c. During a fraction of this cycle ($\alpha_i T_c$), the presence of the G4 cannot be detected, either because the hairpin is open or because the G4 is embedded in double-stranded DNA. Thus, there is a small probability that a G4 structure unfolds and then refolds during this time $\alpha_i T_c$. Then, we would observe two successive blockages at the G4 position that we would consider as a unique structure while in reality, two successive folding and unfolding events happened. In this section, we assess the error made on the estimation of the kinetic parameters of G4 folding (k_f) and unfolding (k_u) due to the possibility of these hidden events.

Given the knowledge that a G-quadruplex is present in the molecule at t_0, the probability P_1 that there is still a structure at a time $t_0 + \alpha_i T_c$, independently of whether it unfolds and refolds one or several times is:

$$P_1 = \frac{k_f}{k_u + k_f} + \frac{k_u}{k_u + k_f} e^{-(k_f + k_u)\alpha_i T_c}$$

Thus, the probability to detect a G4 during N successive cycles is:

$$\begin{aligned} P(N) &= \left[\underbrace{\left(e^{-k_u(1-\alpha_i T_c)}\right)}_{P_0} \left(\underbrace{\frac{k_f}{k_u + k_f} + \frac{k_u}{k_u + k_f} e^{-(k_f + k_u)\alpha_i T_c}}_{P_1} \right) \right]^N \\ &= \left[\underbrace{\left(e^{-k_u(1-\alpha_i T_c)}\right)}_{P_0} \left(\underbrace{1 - (1 - e^{-(k_f + k_u)\alpha_i T_c}) \frac{k_u}{k_u + k_f}}_{P_1} \right) \right]^N, \end{aligned}$$

where P_0 is the probability that the G4 does not unfold during the part of the cycle where it is detectable and P_1 is defined just above. In particular, the probability to observe the G4 during N cycles if it were detectable at all times ($\alpha_i = 0$) would be $e^{-k_u N T_c}$.

Using the inequality $(1+x)^N \leq e^{Nx}$, one gets:

$$P(N) < \left(e^{-k_u(1-\alpha_i T_c)}\right)^N e^{-N(1-e^{-(k_f+k_u)\alpha_i T_c})\frac{k_u}{k_u+k_f}}$$

Now considering that $e^{-x} < 1 - x + x^2/2$, we get:

$$\begin{aligned} P(N) &< \left(e^{-k_u(1-\alpha_i T_c)}\right)^N e^{-N\frac{k_u}{k_u+k_f}((k_f+k_u)\alpha_i T_c + (k_f+k_u)^2 \alpha_i{}^2 T_c{}^2)} \\ &< e^{-k_u T_c N} e^{-N(k_f+k_u)k_u \alpha_i^2 T_c^2/2} \\ &< e^{-T_c N \underbrace{k_u(1-(k_f+k_u)\alpha_i^2 T_c/2)}_{k_{u,obs}}} \end{aligned}$$

This means that the observed folding kinetics $k_{u,obs}$ is slightly different from the real rate $k_{u,real}$, with:

$$k_{u,real} > k_{u,obs} > k_{u,real}(1 - (k_f + k_u)\alpha_i^2 T_c/2)$$

In particular, this means that we slightly overestimate the persistence time $T_p = \frac{1}{k_u}$ and that the amplitude of the overestimation verifies:

$$T_{p,real} < T_{p,obs} < \frac{T_{p,real}}{1 - (k_f + k_u)\alpha_i^2 T_c/2}$$

Conversely, we make a similar error on $T_f = \frac{1}{k_f}$, as it is possible that between two successive detections steps, one was folded and then unfolded:

We thus have,

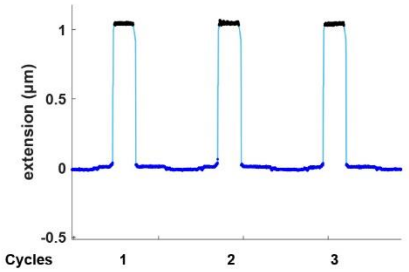
$$T_{f,real} < T_{f,obs} < \frac{T_{f,real}}{1-(k_f+k_u)\alpha_i^2 T_c/2}$$

For our data, these corrections are negligible (< 10%), except for the β^A-origin sequence that folds very rapidly. This overestimation of the measured times is taken into account in all our error bars. We add to the low error bars of our data the maximum relative error that can be made due to these hidden events: $(k_f+k_u)\alpha_i^2 T_c^2/2$.

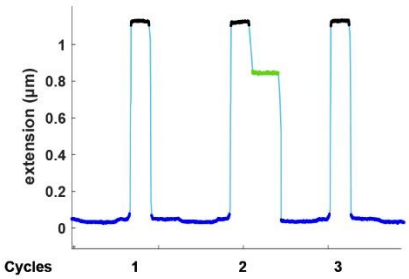
In all the data shown in this paper, we take as a lower error the sum of the probabilistic error due to the above effect and of the sampling errors (obtained through bootstrap-resampling of the data). However, we take as upper error the bootstrapped error only as the hidden events cannot induce any underestimation of the times:

$$\sigma_{up} = \sigma_{bootstrap}$$

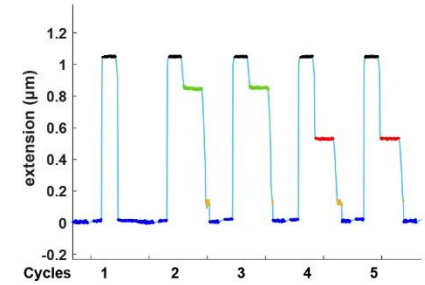
$$\sigma_{down}^2 = \sigma_{bootstrap}^2 + max((k_{real} - k_{obs}))^2$$



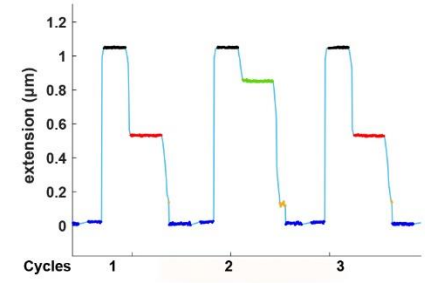
No blockage: no oligonucleotide and no G4



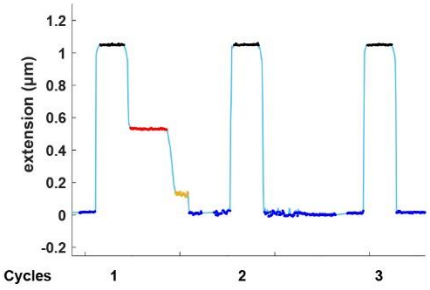
Succesive traces with only oligonucleotide blockage at ~0.75 µm (green): no G4



Oligo blockage followed by successive blocage at ~0.45 µm (red) : appearance of a persistent G4 structure. Here, cycle 4 is the first cycle included in the calculation of Tp. Cycles 2 and 3 are included in the calculation of Tf, which can introduce a small error. This error is taken into account in the estimation of the error of Tf and Tp. Blockage at 0.12 µm (orange) is not used to calculate Tf or Tp, as it is not systematically present when a G4 is present.

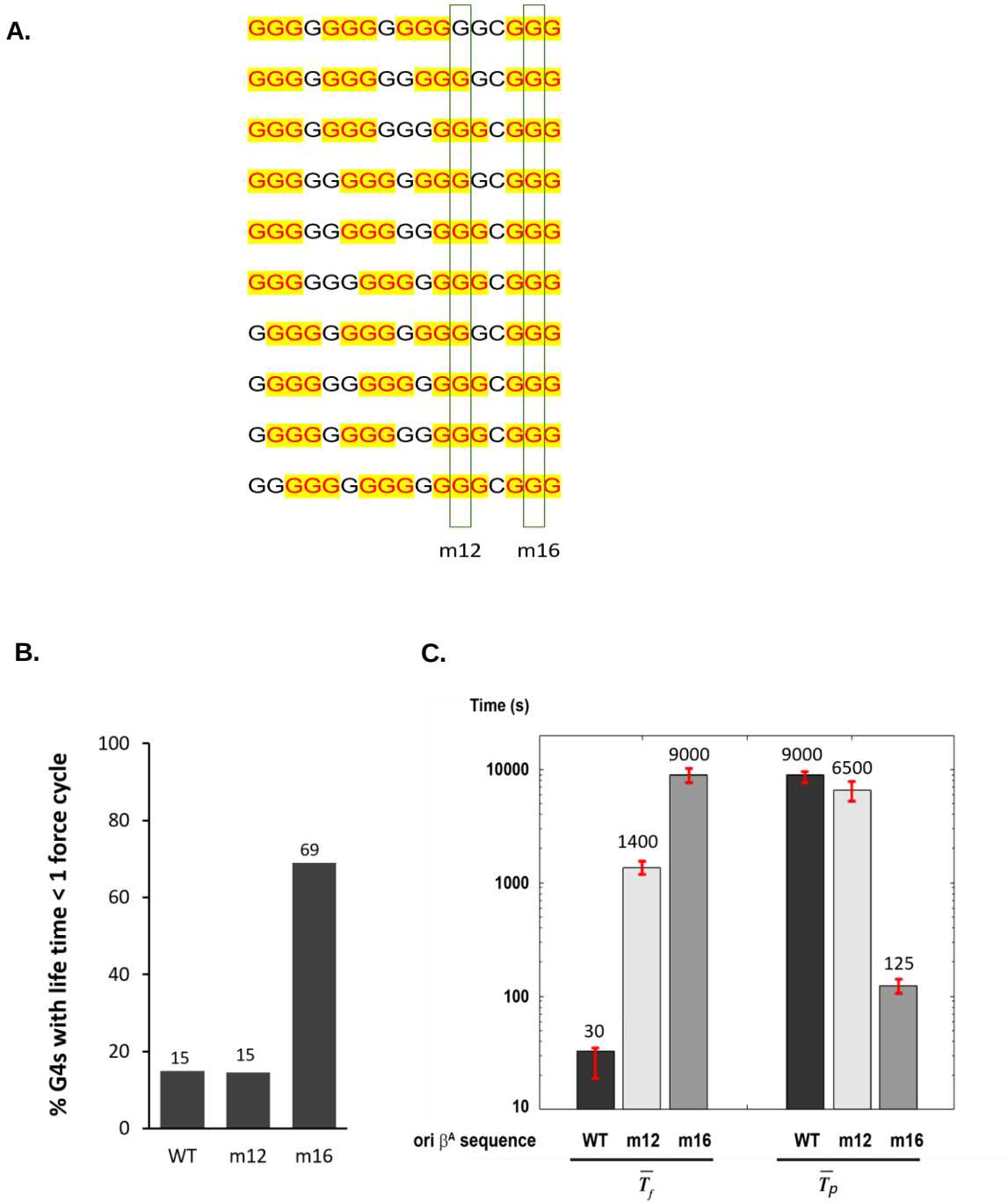


The oligo blockage at cycle 2 hides the detection of the G4 structure, which can be nevertheless confirmed by the presence of the blockage at 0.12 µm (orange). Probability of the G4 structure detected at cycle 1 and 3 being the same structure is explicited in Supplementary Data 9, independently of the presence of a blockage at 0.12 µm.

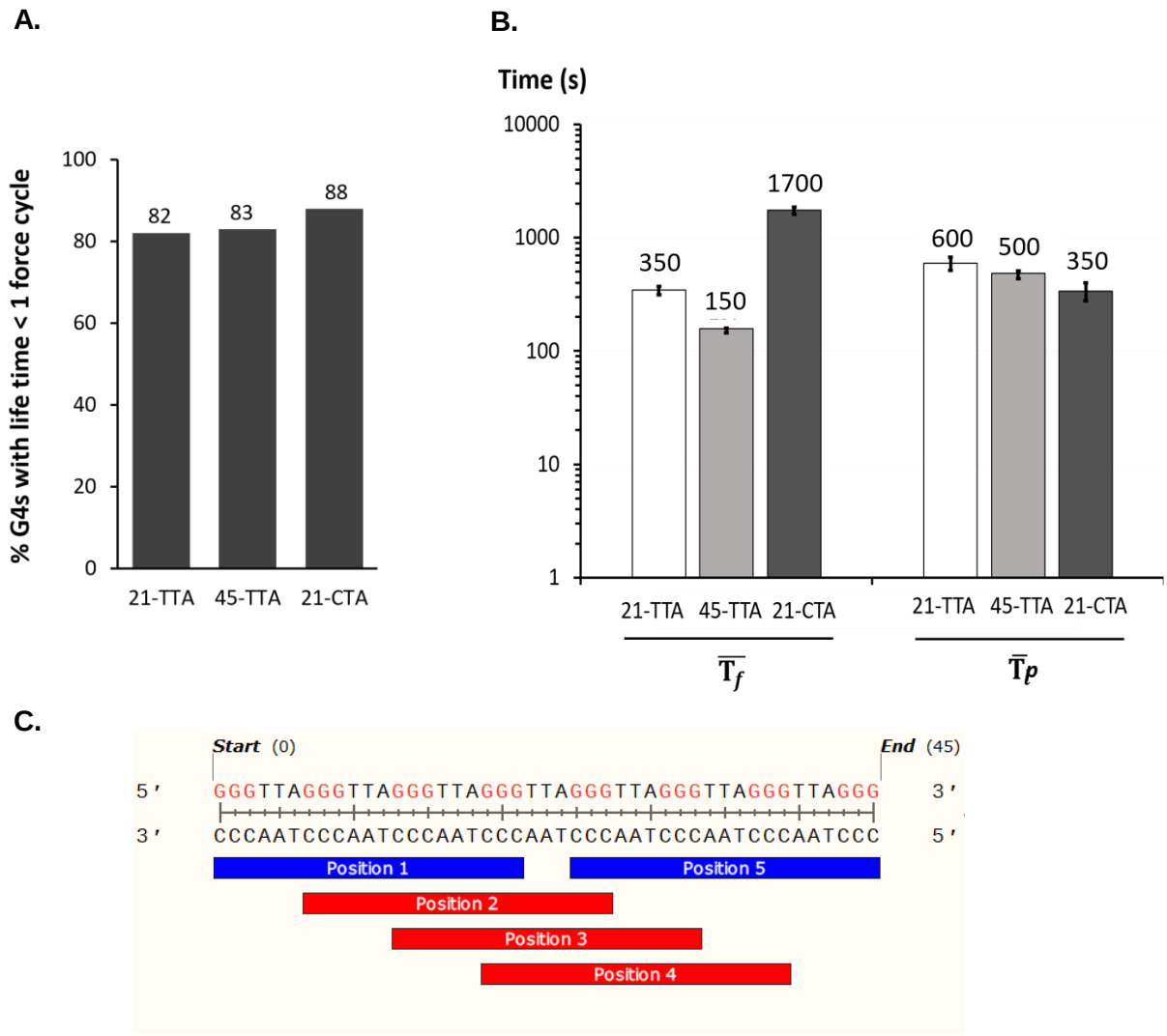


Blockage at 0.45 µm at cycle 1 indicates a G4 structure. Cycles 2 and 3 contain no blockage, indicating that the G4 structure is no longer present. In this case, cycle 1 is the last cycle used for calculating Tp.

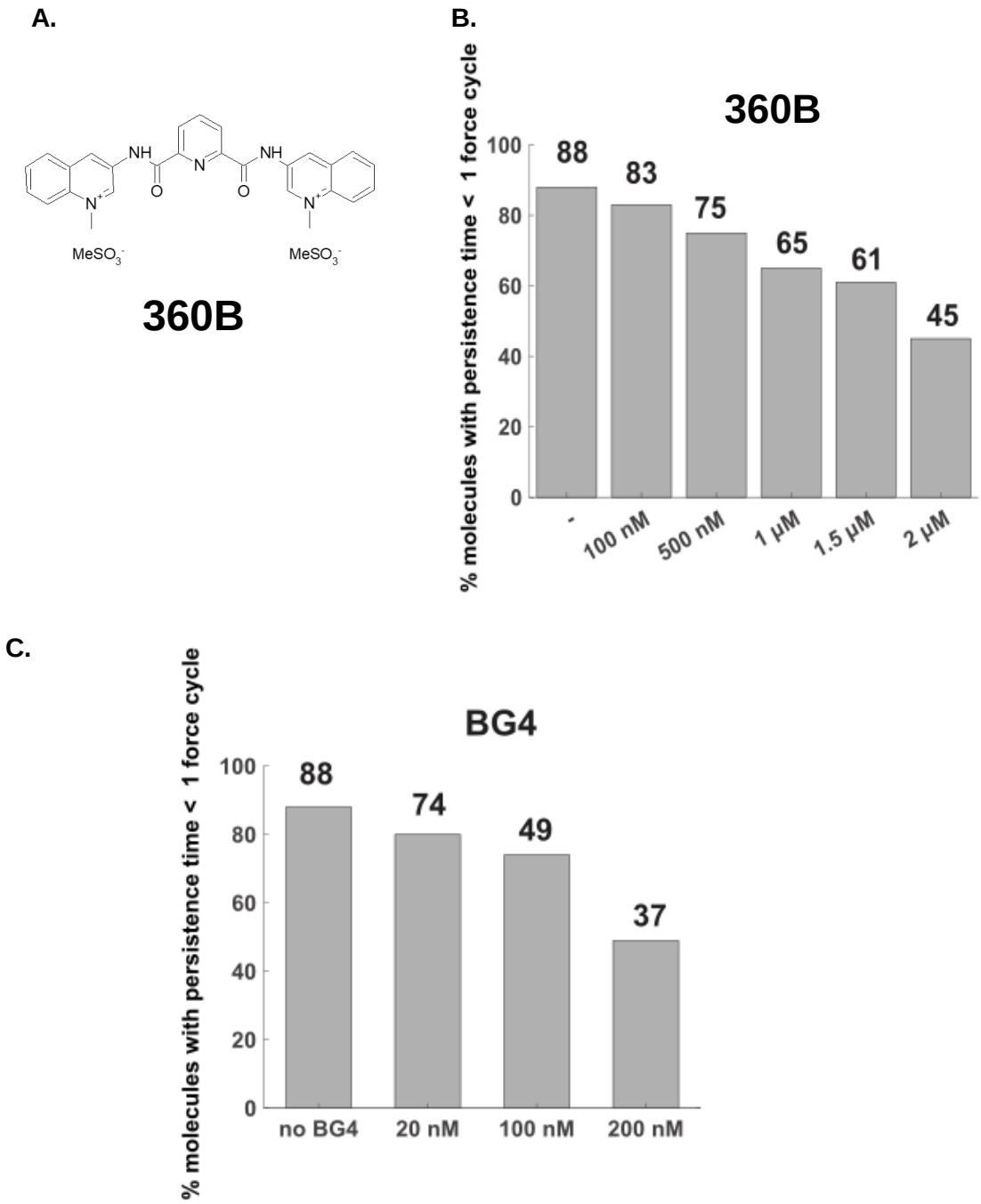
Supplementary Figure S9: Different scenarios for the detection of presence/absence of a G4 structure within the hairpin molecule based on extension patterns during successive force cycles. Fully opened hairpin at 20 pN (black), fully closed hairpin (blue), oligo blockage at 7 pN (green), G4 blockage at 7 pN (red), G4 blockage at 2 pN (orange).



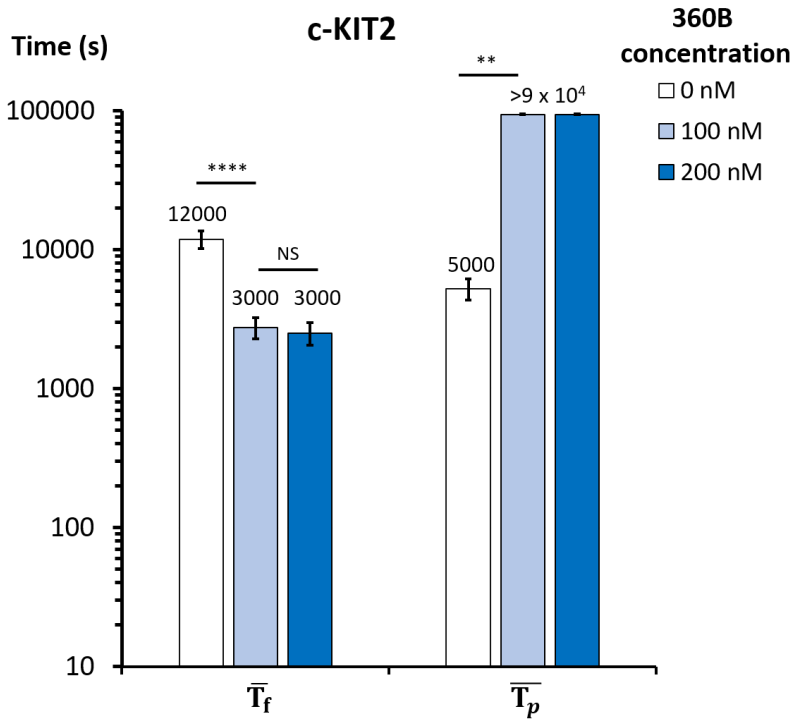
Supplementary Figure S10: Folding and persistence times of G4 structures formed by ori β^A wild type sequence and point mutants m12 and m16. A. Position of the m12 and m16 mutants in the ori β^A sequence. All 10 possibilities for forming an intramolecular G4 implicating 4 runs of 3 consecutive Gs are shown. In the case of the m12 mutant, only one possibility remains, none in the case of m16. **B.** Percentage of G4 with persistence time less than 1 force cycle. **C**. Mean folding ($\overline{T_f}$) and persistence ($\overline{T_p}$) times of wild type (WT) and mutant ori β^A G4 forming sequences.



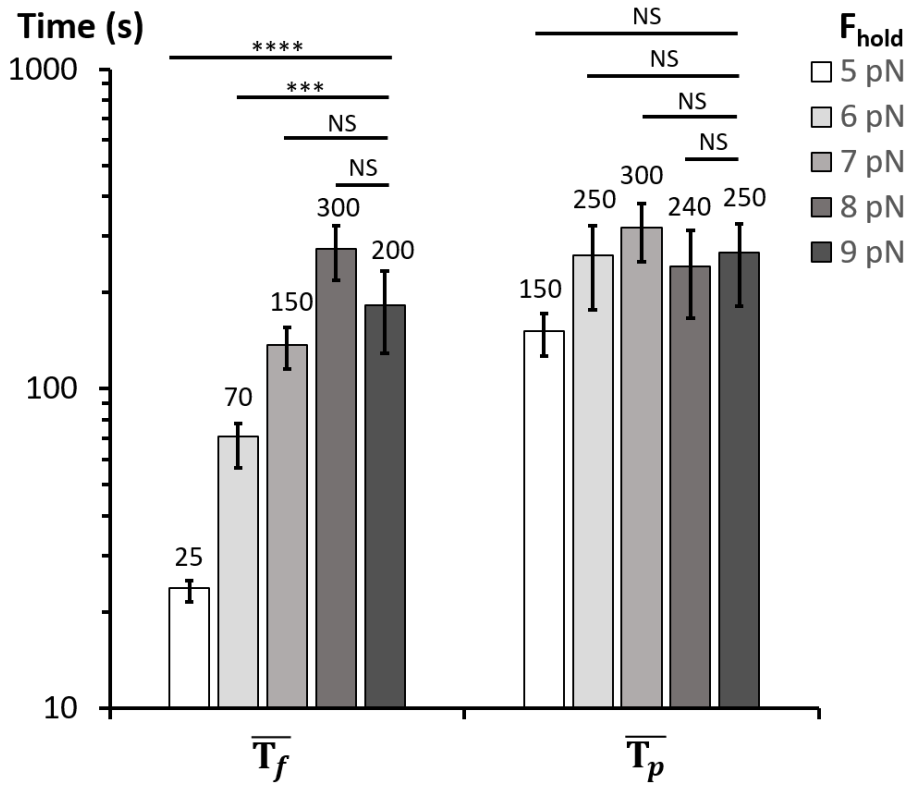
Supplementary Figure S11: Folding and persistence of G4 structures formed by human telomeric sequences. A. Percentage of G4 with persistence time less than 1 force cycle. **B**. Mean folding ($\overline{T_f}$) and persistence ($\overline{T_p}$) times of human telomeric sequences. **C.** Different possible G4 nucleation sites folded by 45-TTA sequence.



Supplementary Figure S12: Effect of G4 ligand 360B and BG4 antibody on the persistence of 21-CTA G4. (**A**) Chemical formula of 360B ligand. Percentage of 21-CTA G4 with persistence time less than 1 force cycle in presence of increasing concentration of 360B (**B**) and BG4 antibody (**C**).



Supplementary Figure S13: The mean of folding and persistence times of c-KIT2 at different concentrations of 360B. P-values are from ~ 100 events and are calculated using a two-tailed t-test. NS stands for non-significative ($p > 0.1$). *** $p \leq 0.005$ and **** $p < 0.00001$



Supplementary Figure S14: Mean folding time and persistence time of human telomeric G4 forming sequences (45-TTA) at different holding forces. We performed different cycle experiments on human telomeric 45-TTA sequence by varying F_{hold} from 5 to 9 pN (force that we can observe blocking oligonucleotide and G4 formation). The results show that the lifetime of 45-TTA G4 did not depend on the applied F_{hold}. However, its folding time increased with the force until reaching a plateau at ~ 7-9 pN. We have then chosen the holding force (F_{hold}) at 7 pN for all G4 experiments where we have reasonable oligo-blocking time (about 10-15s) and where most of DNA hairpins were closed in the absence of oligonucleotide and G4 formation. P-values are from ~ 100 events and are calculated using two-tailed t-test. NS stands for non-significative ($p > 0.1$). ****** $p \leq 0.0001$ and ******** $p \leq 0.00001$

Name	Description	Sequence (5' ➔ 3')
PS124	Hairpin loop	GCTTGCACTGAGAGCGCGGCCTCTCAGTGC
Treble	Tri-biotinylated 5' end oligonucleotide allowing binding of the hairpin to a streptavidin covered magnetic bead	*Biotin(x3)*_ATGGTCAAGCATGCCGCTTTTCGGTTCCCGTGTGTCTTTTGGTCTTTCTGGTGCTCTTC
PS866	3'-end oligonucleotide allowing annealing of the hairpin to oligonucleotide PS867	ATTCGAAGAGCACCAGAAAGACCAAAAGACACAAAGGGTCAGTGCTGCAACCCACTTCCTAATCTGTCATCTTCTG
PS867	Bridge oligonucleotide to link PS866 with PS625	GTGTCTTTTGGTCTTTCTGGTGCTCTTCGAATCAGAAGATGACAGATTAGGAAGTGG
PS625	sequence containing a 3' DBCO group which covalently binds to the azide coated cover slip	ATTCGAAGAGCACCAGAAAGACCAAAAGACACAGACAGATATCGCGCTTCCTCCTACTTTGAATGCTAT_*DBCO*
Blocking oligonucleotide	7 bp blocking oligonucleotide	GCCGCGC

Supplementary Table 1: Sequences of DNA oligonucleotides used for hairpin synthesis and for tethering hairpins to the azide coated coverslip. In black are single-stranded DNA regions. The regions with identical colors (*i.e.* red, purple, orange and green) are complementary sequences.

Name	Description	Sequence (5' ➔ 3')
hTelo 21-CTA	variant human telomeric sequence (single G4 unit)	AATTCTCATAGCATGATACCAATGGGCTAGGGCTAGGGCTAGGGATTGGTACGTAGACCATG
hTelo 21-CTA mut	mutated hTelo 21-CTA	AATTCTCATAGCATGATACCAATGcGCTAGcGCTAGcGCTAGcGATTGGTACGTAGACCATG
cMYC Pu27	promoter of oncogene c-MYC	ATTCTCATAGCATGATAAGGGGAGGGTGGGGAGGGTGGGGAAGGTTGGTACGTAGACCATG
cmyc Pu27 mut	mutated c-MYC Pu27	AATTCTCATAGCATGATAAGGGGAGcGTGccGAGcGTGccGAAGGTTGGTACGTAGACCATG
cKIT2	promoter of oncogene c-KIT	ATTCTCATAGCATGATACCAATGGGCGGGCGCGAGGGAGGGGAATTGGTACGTAGACCATG
ckit2 mut	mutated c-KIT2	AATTCTCATAGCATGATACCAATGcGCGcGCGCGAGcGAGccGAATTGGTACGTAGACCATG
β^A ori	Chicken origin of replication β^A	AATTCTCATAGCATGATACCAATGGGGGGGGGGGGGCGGGATGCATTGGTACGTAGACCATG
β^A ori - M12	mutated ori β^A (G_{12}□A)	ATTCTCATAGCATGATACCAATGGGGGGGGGGGaGCGGGATGCATTGGTACGTAGACCATG
β^A ori - M16	mutated ori β^A (G_{16}□A)	AATTCTCATAGCATGATACCAATGGGGGGGGGGGGGCGaGATGCATTGGTACGTAGACCATG
hTelo 21-TTA	human telomeric sequence (single G4 unit)	AATTCTCATAGCATGATACCAATGGGTTAGGGTTAGGGTTAGGGATTGGTACGTAGACCATG
hTelo 21-TTA mut	mutated hTelo 21-TTA	AATTCTCATAGCATGATACCAATGcGTTAGcGTTAGcGTTAGcGATTGGTACGTAGACCATG
hTelo 45-TTA	human telomeric sequence (two G4 units)	AATTCTCATGGGTTAGGGTTAGGGTTAGGGttaGGGTTAGGGTTAGGGAGACCATG
hTelo 45-TTA mut	mutated hTelo 45-TTA	AATTCTCATGcGTTAGcGTTAGcGTTAGcGttaGcGTTAGcGTTAGcGTTAGcGAGACCATG

Supplementary Table 2: G4 forming sequences and control sequences used in this study: in red, guanines involved in the G4 structure, and in lower-cases, base mutated to destabilize G4 structures.

Conditions	cMYC Pu27	cKIT-2	ori β^A	hTelo TTAGGG
LiCl	**nd**	**nd**	**no G4** extensions: 1µm (1a) 0µm (5b)	**nd**
LiCl + oligo	**G4** extensions: 1µm (1a/b/c) 0.75µm (2a/b) 0.45µm (3) 0.12µm (4) 0µm (5a/b)	**no G4** extensions: 1µm (1a/b) 0.75µm (2a) 0µm (5b)	**G4** extensions: 1µm (1a/b/c) 0.75µm (2a/b) 0.45µm (3) 0.12µm (4) 0µm (5a/b)	**no G4** extensions: 1µm (1a/b) 0.75µm (2a) 0µm (5b)
KCl	**no G4** extensions: 1µm (1a) 0µm (5b)	**no G4** extensions: 1µm (1a) 0µm (5b)	**G4** extensions: 1µm (1a/c) 0.45µm (3) 0.12µm (4) 0µm (5a/b)	**no G4** extensions: 1µm (1a) 0µm (5b)
KCl + oligo	**G4** extensions: 1µm (1a/b/c) 0.75µm (2a/b) 0.45µm (3) 0.12µm (4) 0µm (5a/b)	**G4** extensions: 1µm (1a/b/c) 0.75µm (2a/b) 0.45µm (3) 0.12µm (4) 0µm (5a/b)	**G4** extensions: 1µm (1a/b/c) 0.75µm (2a/b) 0.45µm (3) 0.12µm (4) 0µm (5a/b)	**G4** extensions: 1µm (1a/b/c) 0.75µm (2a/b) 0.45µm (3) 0.12µm (4) 0µm (5a/b)

Supplementary Table 3: Molecule extension states observed for each G4 sequence under different buffer conditions. nd means “not determined”. Structures 1 to 5 mentioned in parentheses are defined in Figure 1a in the main text.