SUPPORTING INFORMATION

Visualization of Long Human Telomere Mimics by Single-Molecule Fluorescence Imaging

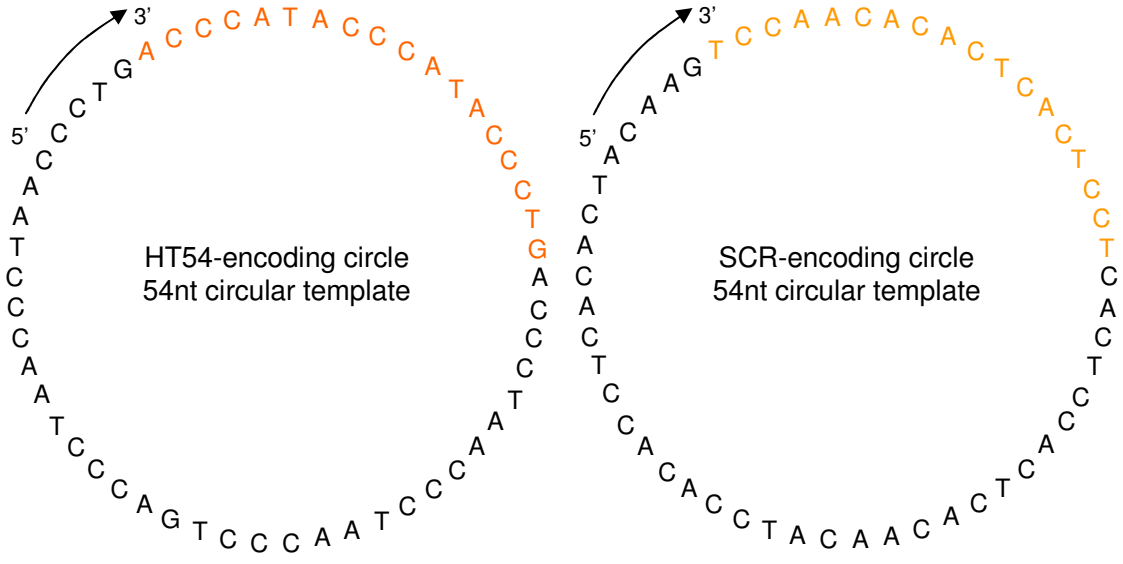
Andrea K. Pomerantz, W. E. Moerner*, and Eric T. Kool*

Department of Chemistry, Stanford University, Stanford, CA 94305

*To whom correspondence should be addressed: wmoerner@stanford.edu; kool@stanford.edu

Part I. Materials and methods for oligonucleotide preparation

The circular templates used in these experiments are derived from those used by Lindström *et al.*[1] for the preparation of human telomere mimics. Instead of being fully repetitive, however, our human telomere mimic strand contains variations in the TTA loop regions of the TTAGGG motif that were incorporated to simplify the synthesis of the circular template. These variations (TAT and TCA) appear to be non-disruptive towards the formation of G-quadruplexes in the RCR product, as the changes lie in the loop regions of the quadruplex motifs. This was verified with circular dichroism spectroscopy (Part II below). Circular 54nt ssDNA templates were prepared by enzymatic ligation of the 5' phosphorylated linear precursors with T4 DNA ligase (New England Biolabs) in the presence of an 18nt splint, following the method described by Diegelman and Kool.[2] For all DNA sequences used in this study, solid-phase oligonucleotide synthesis was carried out on an ABI 394 DNA synthesizer (Applied Biosystems) at the 1 μmole scale with standard β-cyanoethyl phosphoramidite chemistry. Preparative-scale denaturing polyacrylamide gel electrophoresis (PAGE) was used to purify and isolate the full-length oligos. Figure S1 shows the sequences of the circular templates, their corresponding primers (HT54-P and SCR-P), and all linear complements and short RCR product mimics used in this study.



Name	Length	Sequence (5' to 3')
HT54-P	18	biotinTEG-CAGGGTATGGGTATGGGT
SCR-P	18	biotinTEG-AGGAGTGAGTGTGTTGGA
HT1	18	ACCCATACCCATACCCTG
HT2	18	ACCCTAACCCTAACCCTG
SCR1	18	CACACCTCACACTACAAG
H54	54	TGGGTCAGGGTTAGGGTTAGGGTCAGGGTTAGGGTT AGGGTCAGGGTATGGGTA
S54	54	AGGAGTGAGTGTGTTGGACTTGTAGTGTGAGGTGTG GATGTTGTGAGTGGAGTG

Figure S1. Sequences of HT54 human telomere mimic circular template and SCR scrambled circular template, corresponding biotinylated primers (HT54-P and SCR-P, respectively), linear complements and 54nt RCR product mimics (H54 and S54, respectively) described in this study. Primer binding regions are shown in orange on the circular templates, and polarity of the template is indicated with arrows.

Purified oligonucleotide samples were resuspended in Nanopure water (NP H_2O) and studied in the following buffers: Tris (10 mM Tris-Cl, pH 7.4), Tris-Na^+ (10 mM Tris-Cl, 100 mM NaCl, pH 7.4) and/or Tris-K^+ (10 mM Tris-Cl, 100 mM KCl, pH 7.4).

Part II. CD spectroscopic characterization of telomeric and scrambled sequences

To determine whether or not our RCR products would serve as effective telomere structural mimics, we acquired circular dichroism (CD) spectra for the 54nt ssDNA product sequences H54 and S54 as well as for a 69nt fully repeating telomere mimic T69, $GGG(TTAGGG)_{11}$, in a variety of buffers known to stabilize different forms of G-quadruplexes.

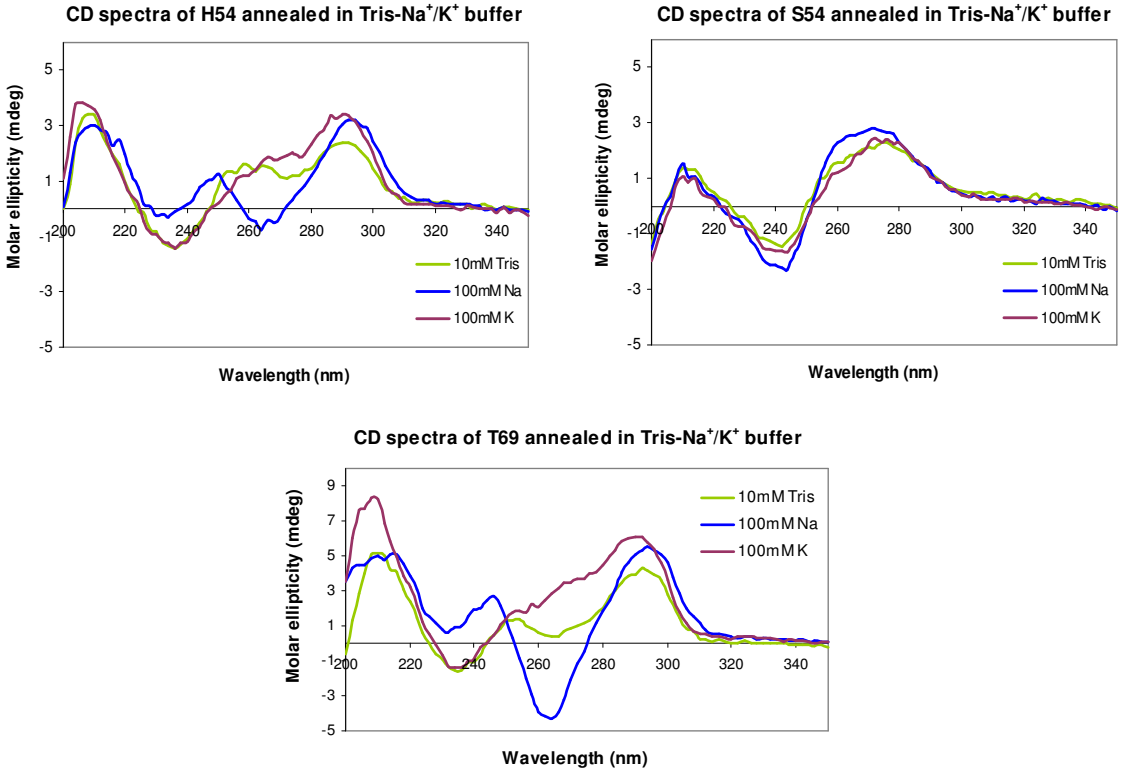


Figure S2. Circular dichroism spectra for 54nt/69nt telomere sequence mimics and 54nt scrambled sequence in various buffers.

The bottom plot in Figure S2 shows that T69 produces a strong negative peak at ~265 nm and a strong positive peak at ~290 nm in the presence of 100mM Na^+, characteristics of an antiparallel G-quadruplex spectrum.[3] Of note is the fact that similar peaks are observed for H54 but not for

S54 in Tris-Na^+ buffer. This indicates that H54 possesses quadruplex-like structure and is in fact structurally distinct from the scrambled sequence under these buffer conditions.

Part III. Protocols for single-molecule (SM) sample preparation, imaging, and data analysis

SM sample preparation

In order to prepare long (>1,000nt) repeating telomeric sequence mimics and scrambled sequence products, rolling circle replication was carried out in flow cells using biotinylated primers (sequences in Figure S1) attached to PEG-coated coverslips as described by Joo *et al.*[4] Rolling circle replication using bacteriophage phi29 DNA polymerase (Epicentre Biotechnologies) and initial primer-template concentrations of 1 nM - 5 nM was carried out at room temperature for varying lengths of time, typically between 1.5 h - 5 h. Complementary single-stranded oligos were added at a concentration of 50 μM and incubated for 10 minutes at room temperature before samples were imaged; SYBR Gold was then added to the sample chamber by syringe pump at a dilution of 1:10,000 to give a final concentration of approx. 2 μM. Flow buffers were degassed for approximately 30 minutes prior to use to minimize photobleaching of SYBR Gold and subsequent photodamage of the labeled DNA.

Sets of short single-stranded DNA complements were designed to bind different lengths and at different locations of the rolling circle replication products as diagrammed in Figure S3, where RCR products are shown in dark blue, and the original primers in orange. Of note is the fact that HT2 binds twice to each 54nt repeat of the HT54 sequence due to the repeating sequence of the HT54 rolling circle replication product. The other 18nt complements (HT1 and SCR1) can bind only once per 54nt repeat, assuming no mismatches.

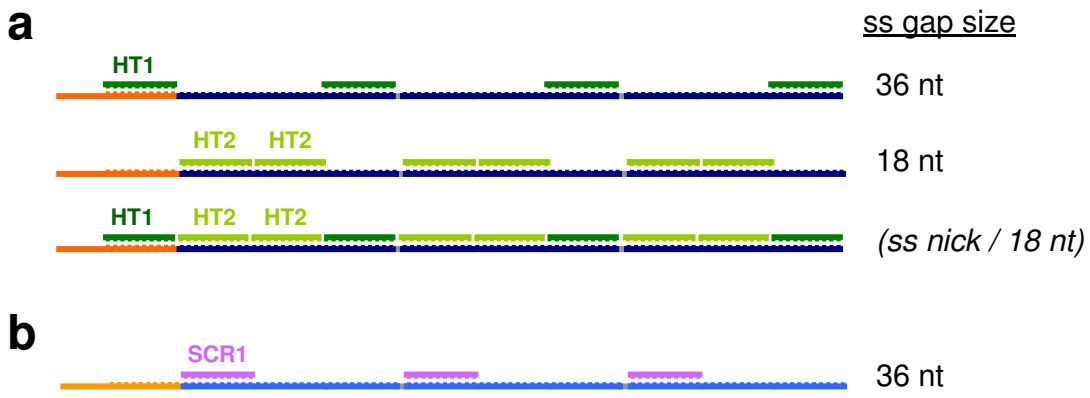


Figure S3. Schematic diagram of complement-binding regions in rolling circle replication products. (a) Complements binding to RCR products encoded by circular template derived from the human telomeric sequence. (b) Complements binding to RCR products encoded by circular template with scrambled sequence. The column on the right notes the minimum sizes of the single-stranded gap regions/nick intervals present after addition of the corresponding complements.

Melting temperatures (T_m) for the various complements in Figure S3 were calculated with DINAMelt[5] in order to ensure that the chosen sequences possessed comparable target binding affinities under buffer conditions similar to those used in our assay.

Name	Length	Sequence (5' to 3')	T_m (ºC)
HT1	18	ACCCATACCCATACCCTG	60.7
HT2	18	ACCCTAACCCTAACCCTG	60.6
SCR1	18	CACACCTCACACTACAAG	58.0

Table S1. Calculated melting temperatures for complements of HT54- and SCR-encoded rolling circle replication products. Simulation parameters: 1 μM complement, 0.1 μM target, 100 mM Na^+, temperature ramp from 0ºC to 100ºC by 1ºC increments.

As shown in Table S1, estimated T_m for the 18nt complements differed by 0.1ºC within the sets for a given template (HT1 vs. HT2), and by 2.6ºC across the sets for the two different templates (HT1 and HT2 vs. SCR1). In the absence of native structure, it would be expected that these sequences would be of comparable stability when bound to target based on their estimated thermodynamic parameters. In all cases, a large excess (10^5-fold vs. the initial primer concentration) of complement in 10 nM Tris buffer was added to the flow cell after DNA synthesis, although the exact stoichiometry of our system is indeterminate, given the surface-immobilized nature of the DNA and the variability in the length of individual strands.

SM imaging procedures

Imaging protocols were adapted from those described in a recent publication from R. Phillips *et al.*, who described SYBR Gold binding to individual DNA molecules ejected from viral capsids.[6] In our setup, SYBR Gold was supplied in solution so that dye concentration was kept relatively constant during buffer flow, and total internal reflection (TIR) illumination of the sample surface was used to reduce background from the unbound/out-of-focus fluorophores. The flow cell was constructed by sandwiching a piece of double-sided polyester cleanroom tape (UltraTape) between a polymer-coated coverslip and a glass microscope slide with 1 mm-diameter drilled inlet/outlet ports. A 2.4 mm by 13 mm linear channel was manually cut in the 8-mil-thick (~200 μm) double-sided tape, creating a chamber with a total volume of approximately 6 μL. A syringe pump (Harvard Apparatus) was used to deliver buffer to the flow cell at a rate of 4 mL/min. We estimate that this flow rate corresponds to an applied force in the range of 0.1 pN - 10 pN; this estimate is based on the similarity of our flow cell dimensions and flow rate to those used by Granéli *et al.* in their flow-stretching of end-tethered lambda DNA.[7] The broad

range of our force estimate is due to the fact that different forces are required to stretch ssDNA vs. dsDNA (~5 pN vs. ~0.1 pN)[8], and our SM stretching experiments involve hybrid ss/dsDNA strands, for which force-extension behavior is less well-defined than for duplex DNA. Both Tris-Na^+ and Tris-K^+ buffers were tested in the presence of telomere mimic and scrambled RCR products, but we noted no significant differences between strands stretched in Na^+ vs. K^+ buffer, and so all experiments described here were carried out in Tris-Na^+ buffer.

Data analysis

Unless otherwise noted, all data sets consisted of 50-frame tiff stacks comprising the video data acquired for each experimental condition tested. The data were acquired with 50-msec exposure times, for a total observation time of 2.5 seconds. Buffer flow (4 mL/min) was manually activated in the first second of data acquisition, and flow was applied for the remainder of the acquisition time so that the majority of the frames in each data set consisted of images of the elongated strands of DNA, as shown above in Figure S4.

Each tiff stack was processed in ImageJ to produce a single-frame average intensity z-projection file. The average z-projection files were then manually analyzed to obtain the lengths of each DNA strand or single-molecule fluorescent spot, and these lengths were plotted in Excel to produce mean-length histograms. The minimum length value present in each data file was set to 0.22 μm, or 4 pixels, to compensate for error in manual measurement of the smallest fluorescent spots observed; therefore, all single-spot measurements less than 4 pixels in length were adjusted up to 4 pixels. This particular value was chosen as it represents the diameter of a single diffraction-limited fluorescent spot under our experimental conditions.

Part IV. Evidence of RCR-based synthesis of kilobase-scale DNA strands

Evidence of enzymatic synthesis of kilobase-scale telomere mimic strands by phi29 polymerase under our experimental conditions was obtained by imaging individual RCR product lengths at the end of various reaction times. Products were visualized as described in Part III above; representative single-frame images are shown in Figure S4.

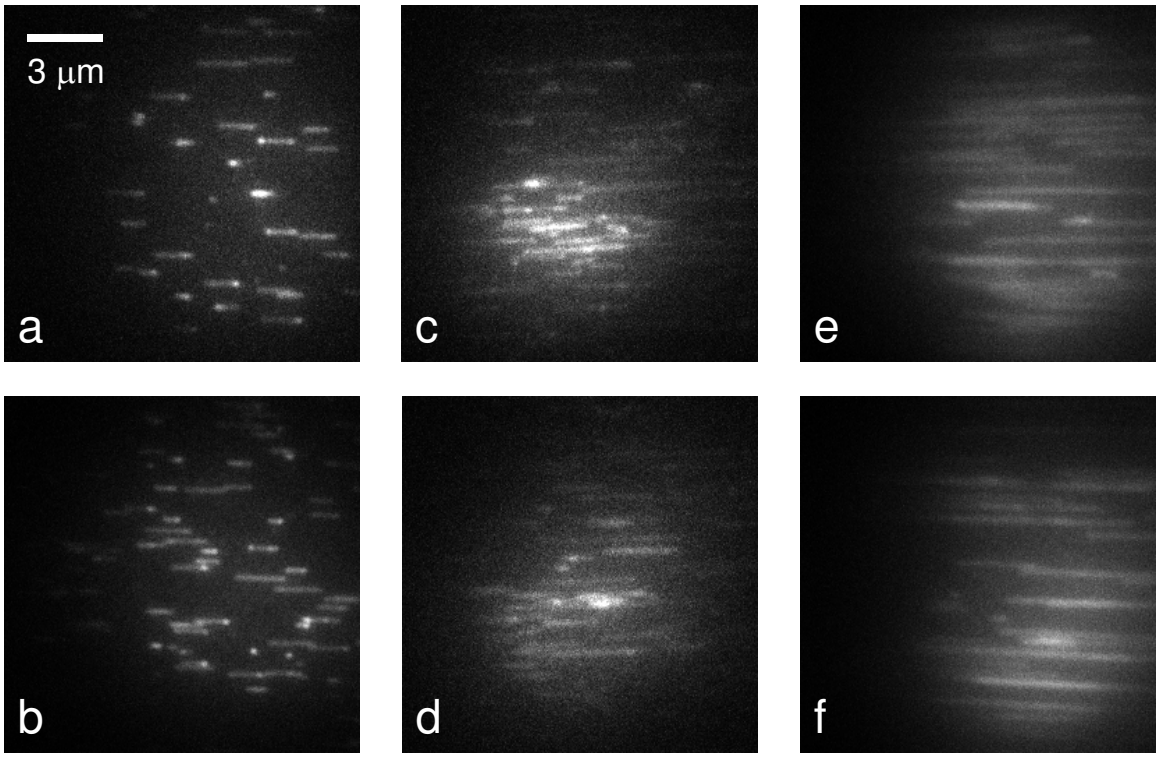


Figure S4. Images of individual telomeric repeat DNA strands produced by RCR and stretched by applied buffer flow in the presence of HT2 complement. (a,b) 2.25 h incubation; (c,d) 4.2 h incubation; (e,f) 5.25 h incubation.

The RCR product length was then converted to a total number of nucleotides by using a distance estimate of 0.34 nm between DNA bases/basepairs and multiplying by 0.67 to account for the fact that only two-thirds of the total strand length is in duplex form in the presence of HT2, assuming maximal binding of complement (see Figure 2B in manuscript). The mean strand

lengths for five different reactions and the corresponding estimated nucleotide insertion rates are shown in Table S2.

Mean strand length (μm)	# of strands measured	Mean # of nucleotides (nt)	Reaction time (s)
0.76	51	1500	9000
0.54	56	1100	13500
1.63	60	3200	15120
1.29	33	2500	17100
4.28	50	8400	18900

Table S2. Estimated strand lengths in nucleotides based on mean detected strand length in microns and an internucleotide spacing of 0.34 nm for two-thirds of the total length of five different SM RCR samples.

It is important to note that the values in the rightmost column of Table S2 represent the lower bounds of the DNA strand lengths for the associated reaction times. Possible sources of inaccuracy in determining the true DNA strand length include the following: the diffraction-limited resolution of the fluorescence imaging setup, such that 0.22 μm is the smallest fluorescent spot size that can be observed; finite extension of the ssDNA regions (currently counted as zero length in our estimate above, where two-thirds of the DNA is assumed to be duplex and to constitute the primary contribution to the observed strand length); incomplete extension of the DNA strand due to strong intramolecular interactions or sticking to the coverslip surface (the latter minimized with the use of polymer coatings); rapid photodamage and cleavage of strands in the presence of SYBR Gold (which we aimed to diminish by degassing buffers prior to use in order to remove oxygen); and extension of the strands beyond the bounds of the excitation beam or out of the TIR focal volume (this latter point is unlikely, as strands display essentially uniform brightness along their length, rather than a decreasing intensity near the free

end that would be indicative of extension out of the evanescent field of illumination). The last four factors would all result in underestimates of the actual strand length, and so it is reasonable to assume that the actual strand lengths could be larger than those listed in Table S2. For future refinement of our flow-stretching scheme, visualization of end-tethered, flow-stretched λ DNA strands of known length (16.3 μm) in our experimental setup should be a straightforward calibration measurement. This would allow us to determine whether or not we were collecting data in the presence of conditions that would necessitate further corrections to our SM strand length estimates; moreover, such a calibration would allow us to investigate the influence of SYBR Gold dye/TMPyP4 binding on the detected length of the DNA. In any case, it is clear that our assay allows us to synthesize easily-detected single strands of surface-immobilized DNA ranging in length from 1kb to >10kb in the span of a few hours at room temperature.

Part V. Protocol for application of G-quadruplex-binding agent TMPyP4

A variety of concentrations of TMPyP4 ranging from 10 nM to 10 μM were tested in our experiments for their impact on SM DNA extensibility, and 100 nM was selected as the optimal experimental concentration to promote an ideal binding stoichiometry of 4:1 TMPyP4:G-quadruplex, where 1 G-quadruplex = $G_3(TTAGGG)_3$, as referenced in Wei *et al.*[9] Furthermore, the influence of TMPyP4 was probed by addition during synthesis, with the assumption that the small molecule quadruplex-binding agent would then be able to access DNA strands as they were synthesized and potentially "lock" G-quadruplexes in place before larger aggregations within the long strands could form, thereby functioning as a sort of molecular chaperone.[10] As in previous SM flow-stretching experiments, we noted no significant differences between strands stretched in Na^+ vs. K^+ buffer.

The three-step procedure used to assess the impact of G-quadruplex-binding agents on long repeating sequences was as follows: first, RCR was carried out in the presence of 100 nM TMPyP4 or BMVC; next, the sample was stained with SYBR Gold, buffer flow was activated, and strand lengths were detected in the absence of complement; and finally, a large excess of complement was added to the sample with a 10-minute room-temperature incubation time, and the strands were stretched once more. Figure S5 highlights the contrasting effects of TMPyP4 on the morphology of telomere-mimicking and scrambled sequence strands, as the former displays a large change in strand appearance when TMPyP4 is present, whereas the latter remains relatively unchanged in the presence of TMPyP4.

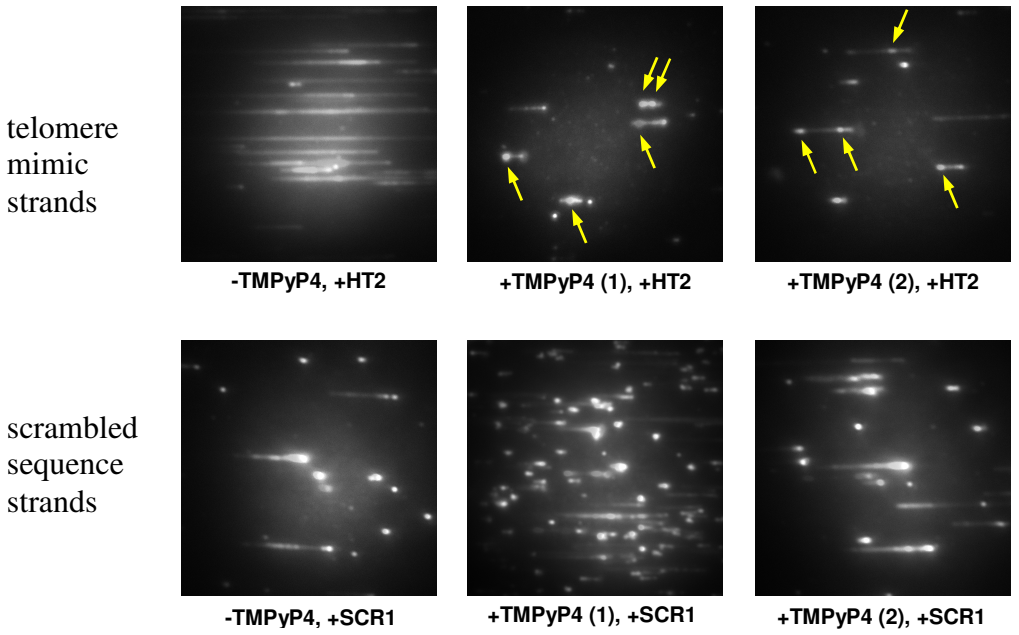


Figure S5. Average intensity projections for telomere-mimicking and scrambled sequence RCR products in the presence of quadruplex-binding agent TMPyP4 highlighting the change in mean strand length and strand morphology after complement addition. Yellow arrows indicate bright clusters in telomere mimic strands in the presence of TMPyP4 and complement HT2.

Part VI. SM video descriptions

SM1.avi: HT54 strands, no complement present (7 sec. total time; 256 × 256 pixels, 14 μm × 14 μm area, 50 ms integration, 4 mL/min applied buffer flow; 5 h 15 min RCR reaction time)

SM2.avi: HT54 strands + large excess of HT2 complement (7 sec. total time; 256 × 256 pixels, 14 μm × 14 μm area, 50 ms integration, 4 mL/min applied buffer flow; 5 h 15 min RCR reaction time)

SM3.avi: HT54 strands + large excess of HT2 complement (2.5 sec. total time; 256 × 256 pixels, 14 μm × 14 μm area, 50 ms integration, 4 mL/min applied buffer flow; 1 h 30 min RCR reaction time)

SM4.avi: HT54 strands + 100nM TMPyP4 + large excess of HT2 complement (2.5 sec. total time; 256 × 256 pixels, 14 μm × 14 μm area, 50 ms integration, 4 mL/min applied buffer flow; 1 h 30 min RCR reaction time)

References

(1) Lindström, U. M.; Chandrasekaran, R. A.; Orbai, L.; Helquist, S. A.; Miller, G. P.; Oroudjev, E.; Hansma, H. G.; Kool, E. T. Artificial human telomeres from DNA nanocircle templates. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 15953-15958.

(2) Diegelman A. M.; Kool, E. T. Chemical and enzymatic methods for preparing circular single-stranded DNAs. *Current Protocols in Nucleic Acid Chemistry* **2000**, 5.2.1-5.2.27.

(3) Porumb, H.; Monnot, M.; Fermandjian, S. Circular dichroism signatures of features simultaneously present in structured guanine-rich oligonucleotides: a combined spectroscopic and electrophoretic approach. *Electrophoresis* **2002**, *23*, 1013-20.

(4) Joo, C.; McKinney, S. A.; Nakamura, M.; Rasnik, I.; Myong, S.; Ha, T. Real-time observation of RecA filament dynamics with single monomer resolution. *Cell* **2006**, *126*, 515-527.

(5) Markham, N. R.; Zuker, M. DINAMelt web server for nucleic acid melting prediction. *Nucleic Acids Res.*, 2005, **33**, W577-W581.

(6) Grayson, P.; Han, L.; Winther, T.; Phillips, R. Real time observations of single bacteriophage λ DNA ejections *in vitro*. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 14652-14657.

(7) Granéli, A.; Yeykal, C. C.; Prasad, T. K.; Greene, E. C. Organized arrays of individual DNA molecules tethered to supported lipid bilayers. *Langmuir* **2006**, *22*, 292-299.

(8) van Oijen, A. Honey, I shrunk the DNA: DNA length as a probe for nucleic-acid enzyme activity. *Biopolymers* **2007**, *85*, 144-153.

(9) Wei, C.; Jia, G.; Yuan, J.; Feng, Z.; Li, C. A spectroscopic study on the interactions of porphyrin with G-quadruplex DNAs. *Biochemistry* **2006**, *45*, 6681-6691.

(10) De Cian, A.; Mergny, J.-L. Quadruplex ligands may act as molecular chaperones for tetramolecular quadruplex formation. *Nucleic Acids Res.* **2007**, *35*, 2483-2493.