UNIVERSITY OF CALIFORNIA
SANTA CRUZ

ARCHITECTURE AND DYNAMICS OF TELOMERASE AND TELOMERES

A dissertation submitted in partial satisfaction
of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

CHEMISTRY

by

Joseph W. Parks

June 2016

The Dissertation of Joseph W. Parks
is approved:

_____
Professor Seth M. Rubin, Chair

_____
Professor Michael D. Stone, Advisor

_____
Professor Carrie L. Partch

_____
Tyrus Miller
Vice Provost and Dean of Graduate Studies

TABLE OF CONTENTS

LIST OF FIGURES AND TABLES

Figure 6. Exonuclease VII digestion reveals extrusion of 5' DNA during repeat addition processivity

CHAPTER V: The telomerase essential N-terminal domain promotes DNA synthesis by stabilizing short RNA-DNA hybrids.

ABSTRACT

Joseph Parks

Architecture and Dynamics of Telomerase and Telomeres

Telomeres and telomerase form a dynamic interplay in order to protect the ends of eukaryotic chromosomes. Telomeres are DNA capping structures that protect the chromosome ends from aberrant recognition as sites of double stranded DNA breaks. Telomerase is a conserved ribonucleoprotein that reverse transcribes G-rich repeats onto the 3' end of these structures in order to facilitate homeostatic telomere length. The reverse transcription reaction of the telomerase protein component is facilitated by an integral RNA template that codes for telomeric sequence. Catalytically, telomerase differs from canonical polymerases by the ability to add multiple template repeats during a single substrate binding event, an action known as translocation. Herein we attempt to dissect this translocation sub-step through single molecule FRET, a technique used for measuring distances and dynamics within macromolecular complexes. We find that the DNA substrate forms a dynamic interplay between multiple conformations during the catalytic cycle and is stabilized by a telomerase specific N-terminal domain. Similarly, we attempt to dissect the intricacies of the telomerase catalytic cycle by observing the integral RNA component during telomerase activity. We find that the telomerase RNA pseudoknot fold undergoes nanometer scale dynamics during telomerase processivity and likely plays an integral role in priming the telomerase complex for successful translocation of the DNA substrate.

Telomeres are DNA structures with highly repetitive sequence elements (TTAGGG in humans) with a 3' G-rich tail, which is the product of telomerase synthesis. Many studies have shown that these repetitive G-rich sequences are capable of forming complex, polymorphic tertiary structures known as quadraplexes. In this thesis we investigate the dynamics and structure of human quadruplexes in order to facilitate understanding of how protein binding partners and enzymes act on these unique DNA structures. Our study involved a recently developed methodology that integrates single molecule FRET with Magnetic Tweezers in order to actively manipulate and monitor conformational changes within the DNA. We find that these DNA structures are relatively brittle and can disrupted by destabilizing only a few base pairs of the structure.

ACKNOWLEDGEMENTS

First I must thank my advisor and mentor, Michael Stone. He has been an amazing resource throughout the entirety of my PhD studies, both emotionally and pedagogically. Not only is he a model scientist, but he is (perhaps more importantly) a truly good person that I admire entirely. Michael, you are one of the best people I know and it is an honor to have worked with you. Also I would like to thank Carrie Partch and Seth Rubin for their insight and guidance during my graduate career. I would like to thank everyone in the Stone lab for their support during my PhD studies as I am not always easiest person to work with. Specifically, I should recognize Martin Hengesbach and Salina Long for their invaluable training and mentorship. I would like to thank my identical twin brother Joshua Parks for encouraging me to strive for excellence. Without our internal competition and collaboration, I believe that my intellectual capacity and curiosity would never have flourished. Similarly, my spiritual and mental health throughout graduate school was largely grounded by you; follow your path my true friend. Lastly, I would like to thank my parents, Virgil and Kathy Parks, who have literally dedicated their entire lives toward providing me opportunities that they were never granted. I truly would not be here without your continued support and I am immeasurably grateful for your unconditional love and dedication, I love you both in a way that is indescribable through words.

CHAPTER I

A brief introduction to telomeres and telomerase

Eukaryotic chromosomes are linear, thereby requiring specialized capping and maintenance machinery to service chromosomal ends. These terminal DNA structures, known as telomeres, contain a repetitive DNA sequence which are composed of a long double stranded region as well as a short single stranded 3' tail [1,2]. Two major barriers must be overcome in order to maintain homeostatic telomeres, the first is known as the end replication problem [3]. Canonical DNA replication machinery fail to copy the telomere 3' terminus, resulting in a progressive shortening of telomere ends over multiple cell cycles. To counteract this problem, eukaryotes evolved the enzyme telomerase to extend the 3' telomere tail [4,5]. The second barrier arises from the structural similarity between the telomere end and a double stranded break [6]. In order to avoid the DNA damage response mechanism, a host of telomere interacting proteins known as the shelterin complex evolved to specifically recognize telomere DNA [7,8]. Proteins within this complex bind both the single stranded and double stranded DNA, forming a telomere capping structure known as the T-loop [9]. Although the entire mechanism of T-loop regulation remains unclear, the 3' tail must be selectively exposed for telomerase binding and telomere extension.

Telomerase achieves telomere extension through a specialized reverse transcription reaction. Like other reverse transcriptases, telomerase can bind an RNA:DNA hybrid with a reverse transcriptase protein domain that catalyzes

extension of the bound DNA substrate [10]. Telomerase differs by its ability to catalyze multiple rounds of template synthesis on a single DNA substrate, a dynamic process known as repeat addition processivity (RAP) [11]. A telomerase specific N-terminal domain (TEN) has evolved to assist with this processive, dynamic behavior [12]. Similarly, a second telomerase specific protein domain called the RNA binding domain associates with the telomerase template containing RNA, preventing template dissociation upon product release [13]. Telomerase RNA has gained other specialized functions including contribution to catalytic activity and telomerase recruitment to the telomeres [14,15].

The product of telomerase synthesis is single stranded DNA composed of G-rich repeats, a DNA sequence that can fold into G-quadruplex structures [16]. Quadruplexes themselves are composed of tetrameric base paired structures known as G-quartets [17]. Depending on the orientation these quartets form and stack, a single G-rich sequence can result in several different structures [18]. Many of these structures are near iso-energetic and readily interconvert at physiological conditions, resulting in highly polymorphic DNA structures. Telomere DNA sequence is highly repetitive and structurally unique, which provides a unique sequence dependent platform for telomere protein binding [19].

A dynamic complex of telomere associated proteins, known as shelterin in humans, have specifically evolved to recognize and interact with telomeric DNA. Some of these proteins specifically recognize either single stranded or double stranded telomere DNA [20,21]. Through interaction with bridging proteins, these

telomere binding proteins maintain a diverse set of functions ranging from telomerase processivity enhancement to T-loop formation [7,22]. More generally, shelterin function revolves around telomere homeostasis through interaction with the repetitive telomere sequence. Shelterin proteins resolve polymorphic telomere DNA in order to create a dynamic structure which protects the telomere from DNA damage response but also presents the telomere end for extension by telomerase.

Telomere DNA, telomere proteins, and telomerase maintain dynamic behaviors that are amenable to study at the single molecule level [23,24]. To successfully approach telomerase with single molecule techniques, one might look to a homologous system that has benefited largely from single molecule studies. One such case is the HIV reverse transcriptase, where crucial mechanistic insights were gained through the use of single molecule biophysics. Although initial crystallographic models HIV RT suggested a single binding mode with the RNA:DNA substrate, a single molecule FRET study provided the first direct evidence for alternative binding modes [25]. Closer analysis of the complex revealed the binding modes themselves regulate the catalytic activity of HIV RT [26]. Unlike the HIV RT, structural information about physiological telomerase, telomeres, and quadruplexes is quite limited. A fragmented structural perspective of telomeres and telomerase creates a great challenge for approaching questions with single molecule methods, but it also creates a demand for biophysical approaches that probe both dynamics and structure. Many questions in the telomerase field have already been answered through the use of single molecule biophysics, yet many remain unresolved [23,24].

In Chapter II, I describe the first single molecule dissection of the human telomerase mechanism. In this work, we use single molecule FRET to dissect the internal DNA dynamics associated with telomerase catalysis. We find that the DNA structure is highly dynamic during telomerase catalysis and the rate limiting step of telomerase translocation is not template hybrid dissociation, but rather re-entry of the newly formed hybrid back into the active site. During this rate limiting step, we also discovered that downstream DNA contacts are modified, allowing extrusion of the DNA product during translocation.

In Chapter III, I describe a collaborative study between the Stone and Das research group where we use a network of FRET measurements within the telomerase RNA component to map the telomerase RNP structure. By labeling the telomerase RNA component in multiple locations, we generate a network of label sites, and therefore a network of distance measurements through single molecule FRET. These distance constraints were integrated with Rosetta modeling to reveal the RNA architecture in the telomerase RNP. We also find that the essential telomerase RNA pseudoknot fold undergoes nanometer scale dynamics during catalysis. This motion was found to be directly coupled to motion of the template hybrid and reveals the importance of RNA connectivity between the template hybrid and the pseudoknot domain for telomerase processivity.

In Chapter IV, I describe a hybrid Magnetic Tweezers single molecule FRET approach that was used to dissect the folding landscape of a human telomere G-quadruplex. The simultaneous application of force to a single molecule FRET labeled

substrate allowed dissection of the quadruplex folding pathway at relatively low perturbative forces. This application was particularly useful for the G-quadruplex as the energy required for perturbing the DNA structure is relatively low. The main finding of the paper revealed that the G-quadruplex is a relatively brittle structure that readily unfolds after a few crucial contacts are disrupted.

In Chapter V, I describe a project that illuminated the functional significance of the *Tetrahymena* telomerase N-terminal (TEN) domain towards stabilizing the template hybrid. A single point mutant of the TEN domain can separate telomerase processivity from the reverse transcriptase activity. In the context of a single molecule FRET assay, this mutant shows increased DNA dynamics to an alternative conformation. Mutations in the TEN domain force the substrate DNA into an alternative conformation more readily when template hybrid is short. These results show that during the critical phases of telomerase processivity, when the template hybrid is short, the TEN domain acts to stabilize the DNA:RNA hybrid in the active site.

In Chapter VI, I discuss a review of the Magnetic Tweezers - single molecule FRET method where we delineate the hybrid method in detail for general use.

In Appendix I, I describe the use and design of custom written software for control, acquisition, and real-time analysis of an MT-smFRET setup. The software system is built predominantly in the LabVIEW runtime environment. In general, the software is used to track the bead center (x, y) and vertical (z) positions. At the end of

the chapter is a description of the DAQmx based code that allows automated control of laser shutters during single molecule FRET experiments.

REFERENCES

1 Blackburn, E. H. Structure and function of telomeres. *Nature* **350**, 569-573 (1991).

2 Wright, W. E. & Shay, J. W. Telomere dynamics in cancer progression and prevention: fundamental differences in human and mouse telomere biology. *Nature medicine* **6**, 849-851 (2000).

3 Cech, T. R. & Lingner, J. Telomerase and the chromosome end replication problem. *Ciba Foundation symposium* **211**, 20-28; discussion 28-34 (1997).

4 Greider, C. W. & Blackburn, E. H. The telomere terminal transferase of Tetrahymena is a ribonucleoprotein enzyme with two kinds of primer specificity. *Cell* **51**, 887-898 (1987).

5 Morin, G. B. The human telomere terminal transferase enzyme is a ribonucleoprotein that synthesizes TTAGGG repeats. *Cell* **59**, 521-529 (1989).

6 de Lange, T. Protection of mammalian telomeres. *Oncogene* **21**, 532-540 (2002).

7 de Lange, T. Shelterin: the protein complex that shapes and safeguards human telomeres. *Genes & development* **19**, 2100-2110 (2005).

8 Palm, W. & de Lange, T. How shelterin protects mammalian telomeres. *Annual review of genetics* **42**, 301-334 (2008).

9 Griffith, J. D. *et al.* Mammalian telomeres end in a large duplex loop. *Cell* **97**, 503-514 (1999).

10 Nugent, C. I. & Lundblad, V. The telomerase reverse transcriptase: components and regulation. *Genes & development* **12**, 1073-1085 (1998).

11 Greider, C. W. Telomerase is processive. *Molecular and cellular biology* **11**, 4572-4580 (1991).

12 Jacobs, S. A., Podell, E. R. & Cech, T. R. Crystal structure of the essential N-terminal domain of telomerase reverse transcriptase. *Nature structural & molecular biology* **13**, 218-225 (2006).

13 Lai, C. K., Mitchell, J. R. & Collins, K. RNA binding domain of telomerase reverse transcriptase. *Molecular and cellular biology* **21**, 990-1000 (2001).

14 Bachand, F. & Autexier, C. Functional regions of human telomerase reverse transcriptase and human telomerase RNA required for telomerase activity and RNA-protein interactions. *Molecular and cellular biology* **21**, 1888-1897 (2001).

15 Jady, B. E., Richard, P., Bertrand, E. & Kiss, T. Cell cycle-dependent recruitment of telomerase RNA and Cajal bodies to human telomeres. *Molecular biology of the cell* **17**, 944-954 (2006).

16 Parkinson, G. N., Lee, M. P. & Neidle, S. Crystal structure of parallel quadruplexes from human telomeric DNA. *Nature* **417**, 876-880 (2002).

17 Williamson, J. R., Raghuraman, M. K. & Cech, T. R. Monovalent cation-induced structure of telomeric DNA: the G-quartet model. *Cell* **59**, 871-880 (1989).

18 Dai, J., Carver, M. & Yang, D. Polymorphism of human telomeric quadruplex structures. *Biochimie* **90**, 1172-1183 (2008).

19 Court, R., Chapman, L., Fairall, L. & Rhodes, D. How the human telomeric proteins TRF1 and TRF2 recognize telomeric DNA: a view from high-resolution crystal structures. *EMBO reports* **6**, 39-45 (2005).

20 Hanaoka, S., Nagadoi, A. & Nishimura, Y. Comparison between TRF2 and TRF1 of their telomeric DNA-bound structures and DNA-binding activities. *Protein science : a publication of the Protein Society* **14**, 119-130 (2005).

21 Lei, M., Podell, E. R. & Cech, T. R. Structure of human POT1 bound to telomeric single-stranded DNA provides a model for chromosome end-protection. *Nature structural & molecular biology* **11**, 1223-1229 (2004).

22 Latrick, C. M. & Cech, T. R. POT1-TPP1 enhances telomerase processivity by slowing primer dissociation and aiding translocation. *The EMBO journal* **29**, 924-933 (2010).

23 Lin, J., Kaur, P., Countryman, P., Opresko, P. L. & Wang, H. Unraveling secrets of telomeres: one molecule at a time. *DNA repair* **20**, 142-153 (2014).

24 Hengesbach, M., Akiyama, B. M. & Stone, M. D. Single-molecule analysis of telomerase structure and function. *Current opinion in chemical biology* **15**, 845-852 (2011).

25 Rothwell, P. J. *et al.* Multiparameter single-molecule fluorescence spectroscopy reveals heterogeneity of HIV-1 reverse transcriptase:primer/template complexes. *Proceedings of the National Academy of Sciences of the United States of America* **100**, 1655-1660 (2003).

26 Abbondanzieri, E. A. *et al.* Dynamic binding orientations direct activity of HIV reverse transcriptase. *Nature* **453**, 184-189 (2008).

CHAPTER II

Coordinated DNA dynamics during the human telomerase catalytic cycle.

*Originally published in Nature Communications 2014

ABSTRACT

The human telomerase reverse transcriptase (hTERT) utilizes a template within the integral RNA subunit (hTR) to direct extension of telomeres. Telomerase exhibits repeat addition processivity (RAP) and must therefore translocate the nascent DNA product into a new RNA:DNA hybrid register to prime each round of telomere repeat synthesis. Here we use single-molecule FRET and nuclease protection assays to monitor telomere DNA structure and dynamics during the telomerase catalytic cycle. DNA translocation during RAP proceeds through a previously uncharacterized kinetic sub-step during which the 3'-end of the DNA substrate base pairs downstream within the hTR template. The rate constant for DNA primer re-alignment reveals this step is not rate-limiting for RAP, suggesting a second slow conformational change repositions the RNA:DNA hybrid into the telomerase active site and drives the extrusion of the 5'-end of the DNA primer out of the enzyme complex.

INTRODUCTION

Specialized nucleoprotein structures called telomeres protect the ends of eukaryotic chromosomes from aberrant recognition and processing by DNA damage repair proteins[1,2]. The inability of conventional replication machinery to generate a complete copy of telomere DNA results in gradual telomere shortening with each round of cell division and limits the proliferative capacity of cells[3]. However, in

rapidly dividing cell types, including the majority of human cancers, the enzymatic activity of telomerase maintains telomere length by adding short telomere DNA repeats (GGTTAG in humans) to chromosome ends[4–6]. The function of the telomerase ribonucleoprotein complex relies upon contributions from multiple evolutionarily conserved domains within the telomerase reverse transcriptase (hTERT) protein and RNA (hTR) subunits[7,8]. Telomerase recognizes the 3'-end of single-stranded telomere DNA via Watson-Crick base pair interactions with the hTR template, forming an RNA:DNA hybrid which is then extended by hTERT to generate a complete telomere DNA repeat[8,9]. Synthesis of a single telomere DNA repeat by telomerase proceeds through a mechanism similar to other DNA polymerases and reverse transcriptases[10]. In contrast, telomerase displays a unique repeat addition processivity (RAP) activity, during which the nascent telomere RNA:DNA hybrid dissociates and re-aligns downstream in the hTR template, permitting the addition of multiple telomere repeats to the growing DNA chain prior to dissociation (Fig. 1a)[11–13].

Telomere DNA primer length and 5'-sequence composition are critical determinants of RAP, indicating that binding interactions apart from the RNA:DNA hybrid contribute to telomerase function[6,14,15]. DNA crosslinking experiments have identified specific points of contact between the telomerase essential N-terminal (TEN) domain of hTERT and telomere DNA primers[16,17]. In addition, hTERT harbors a conserved RT primer grip motif which mediates interactions with the RNA:DNA hybrid and is crucial for proper RAP[18–20]. Thus, discrete points of telomerase-DNA

contact prevent enzyme dissociation during the complex DNA handling required for RAP[21]; however, the precise details of how the DNA binding properties of telomerase are modulated during multiple rounds of telomere repeat synthesis remain unknown.

Single-molecule approaches have been used to interrogate conformational rearrangements of numerous polymerase-nucleic acid complexes, providing novel insight into how template and product strands are dynamically bound and restructured during both DNA and RNA synthesis[22,23]. For example, single molecule Förster resonance energy transfer (smFRET) revealed a DNA-scrunching mechanism for RNA polymerase during early stages of transcriptional initiation[24], mapped the exit path of nascent messenger RNA out of eukaryotic RNA polymerase II[25], and revealed the nucleic acid binding orientation dynamics of the HIV reverse transcriptase[26,27]. Here, we have used smFRET paired together with nuclease protection mapping experiments to characterize the structural dynamics and kinetic sub-states of telomere DNA substrates during telomerase catalysis. Our results reveal that single-stranded telomere DNA primers exhibit a compact conformation during early stages of repeat synthesis. Upon completion of a telomere repeat, the 3'-end of the nascent DNA dynamically samples a new base pairing register in the downstream region of the hTR template. Interestingly, kinetic characterization of this primer re-alignment sub-step indicates that a subsequent rearrangement of the telomerase complex must be rate-limiting for complete translocation during RAP. We further characterized DNA rearrangements during the rate-limiting step of RAP using biochemical foot printing experiments which suggested that repositioning of the re-

aligned RNA:DNA hybrid in the active site is accompanied by the extrusion of the 5'-end of the DNA substrate out of the enzyme complex. Taken together, our results support a model for telomerase repeat addition processivity, during which coordinated DNA movements serve to both prime the synthesis of subsequent telomere repeats as well as render the 5'-end of the DNA substrate accessible to *trans*-acting proteins such as the POT1-TPP1 processivity factor[28].

MATERIALS AND METHODS

Preparation of RNAs

***In vitro* transcription**: *In vitro* reconstitution of telomerase enzyme was performed using the established two RNA fragment system in which the essential RNA pseudoknot domain and the CR4/CR5 domain are added *in trans*[52,53]. Unlabeled CR4/CR5 RNA (hTR 239-328) was generated by *in vitro* transcription using T7 RNAP and PAGE purified. Site-specific RNA labeling within the RNA pseudoknot domain was achieved using DNA-splinted RNA ligation techniques as described previously[54]. Unlabeled fragments of the RNA pseudoknot domain (hTR 63-195 and hTR 63-201) were prepared by *in vitro* transcription using T7 RNAP and purified using a 6% denaturing PAGE gel. Subsequently the RNA was phosphatase (CIP, NEB) treated to remove the 5' triphosphate, phenol/chloroform extracted to remove CIP, T4 PNK (NEB) treated to add a 5' monophosphate, phenol/chloroform extracted to remove PNK, and PAGE purified to remove excess nucleotide. Both enzymatic reactions were carried out using NEB specifications.

Synthetic, dye-labeled RNA fragments: Synthetic RNA fragments were purchased from Dharmacon containing internal aminoallyl uridines (5-N-U) used for dye labeling reactions. Specific RNA sequences are listed in Supplementary Table 1. RNAs were generally processed as previously described[54] with the following exceptions. All ethanol precipitation steps were replaced by purification over a GE PD-midiTRAP G25 column. RNA in each fraction was identified by NanoDrop or by eye once a dye label was present. The fractions were pooled and evaporated to dryness using a vacuum centrifuge set at 45°C. If a 5' phosphate was required for ligation, the dye-labeled oligos were phosphorylated by PNK treatment, followed by phenol/chloroform extraction.

Ligation of RNA Fragments: Modifications of the previously reported procedure[55] included addition of 1 U/μL RNasin and carrying out the ligation overnight at 30°C. The DNA splints for each RNA ligation are listed in Supplementary Table 1. Note for the doubly-RNA labeled construct the U29 and U42 fragments were ligated and purified prior to ligation with the *in vitro* transcribed fragment.

Preparation of telomerase enzyme

Telomerase reconstitution: Telomerase reconstitution was performed using the TnT Quick Coupled Transcription/Translation system (Promega)[55]. For every 100 μL of TnT Quick mix, 2 μg of pNFLAG-hTERT was added. RNAs were then added to the reaction mixture where dye-labeled RNAs were added to a final concentration of 0.1 μM whereas all unlabeled RNAs were added to 1 μM. The reconstitution mixture was held at 30°C for 3 hours.

Telomerase Purification: Directly after reconstitution, telomerase enzymes were purified using an N-terminal FLAG tag on hTERT. To pull down the enzyme, Sigma Anti-FLAG M2-agarose bead slurry was used at 50 µL bead slurry per 200 µL lysate. The beads were first washed four times with wash buffer (50 mM Tris HCl pH 8.3, 3 mM $MgCl_2$, 2 mM DTT, 1 mM spermidine, 100 mM KCl), spinning at 2350 rcf, 4°C, for 1 min between washes. The beads were blocked while being agitated in 750 µL blocking buffer (50 mM Tris HCl pH 8.3, 3 mM $MgCl_2$, 2 mM DTT, 1 mM spermidine, 0.5 mg/mL BSA, 0.05 mg/mL glycogen, 0.1 mg/mL yeast tRNA) for 15 minutes at 4°C. After spinning the beads at 2350 rcf, 4°C, for 1 min the supernatant was discarded. The blocking step was repeated and then the beads were added to the lysate. The beads were agitated at 4°C for two hours allowing the prepared enzyme to associate with the anti-FLAG beads. The mixture was spun at 2350 rcf, 4°C, for 1 min and the supernatant was discarded. The beads were then washed with 750 µL of wash buffer and spun at 2350 rcf, 4°C, for 1 min. This process was repeated four times to remove residual lysate. To elute the enzyme, 60 µL of elution buffer (50 mM Tris HCl pH 8.3, 3 mM $MgCl_2$, 1 mM spermidine, 0.5 mg/mL BSA, 10% glycerol, 0.75 mg/mL 3 x FLAG peptide) was added for every 200 µL of lysate and the slurry was incubated at 4°C for one hour. Beads were removed by centrifugation filtration using a Nanosep MF 0.45 µm filter. From this solution 10 µL aliquots were prepared, flash frozen in liquid nitrogen, and stored at -70°C until use.

DNA primer extension assays: Activity assays were performed using 5 µL of purified enzyme diluted to a final volume of 10 µL in 1x activity buffer. Each

reaction maintained a final concentration of 1 μM primer substrate, 250 μM dATP, 250 μM dTTP, 1 μM α^{32}P-dGTP (PerkinElmer BLU514Z500UC), and 1x activity buffer (50 mM Tris HCl pH 8.3, 3 mM $MgCl_2$, 2 mM DTT, 1 mM spermidine). The reactions were initiated by addition of dNTPs and incubated at 30°C for 90 minutes unless noted otherwise. Reactions monitoring stepwise addition of dNTPs to a DNA primer were performed for 30 minutes in the presence of a combination of 250 μM dNTPs plus the appropriate chain terminating ddNTPs at 250 μM as well (Supplementary Fig. 2, lanes 8-11). Reactions were terminated by addition of 90 μL stopping buffer (10 mM Tris HCl pH 7.5, 1 mM EDTA, 0.1% w/v SDS). The reactions were mixed vigorously with 100 μL of phenol to remove protein components followed by a 10 minute spin at 15000 rcf. The aqueous layer was further extracted with 100 μL of chloroform and spun at 15000 rcf for two minutes. After isolation of the aqueous phase, the radiolabeled DNA was precipitated and then separated on a 12% denaturing PAGE gel (19:1 cross-linking ratio). The gel was dried, exposed on a phosphorimager screen overnight, and scanned using a Typhoon scanner. Quantification of the bands was accomplished by using SAFA v11b[56].

Preparation of Dot Blot Oligo Probe: The 24 nt DNA probe (Supp. Table 1) hybridizes to pseudoknot nt 130-153. 10 pmol probe oligo was 5' ^{32}P labeled using 10 μL γ^{32}P-ATP (PerkinElmer BLU002Z250UC). The reaction was carried out in 1 x PNK Buffer (NEB) at a total volume of 20 μL with 10 U of PNK (NEB). The probe was purified using two sequential Centri-Spin 10 columns following the manufacturer's protocol.

Dot Blot Quantification of Enzyme Concentration: To quantify the concentration of telomerase we determined the concentration of telomerase pseudoknot RNA in the Immuno-precipitated enzyme solution. First a serial dilution of in-vitro transcribed pseudoknot RNA was made as an internal standard (10 fmol/μL, 5 fmol/μL, 1 fmol/μL, and 0.5 fmol/μL). For each of the standard RNA solutions and the reconstituted enzymes, 1 μL of sample was made up with 9 μL of formamide buffer (90% formamide, 0.1% bromophenol blue, 0.1% xylene cyanol, 1x TBE). The samples were heated at 70°C for 5 minutes and the put on ice. After cooling, the samples were applied to a Hybond N+ membrane (GE Life Sciences RPN119B). The samples were air dried and then UV-crosslinked to the membrane using a Stratagene Stratalinker 1800 set to the "Auto-crosslink" program. The membrane was prehybridized at 55°C in 10 mL of Church buffer (1% BSA Fraction V, 1 mM EDTA pH 7.5, 500 mM Na_2HPO_4, pH 7.2, 7% SDS) for 30 min. A total of 1 x 10^6 cpm of ^{32}P labeled dot-blot oligo probe was then added to the hybridization buffer and incubated overnight at 55°C. The membrane was washed with 150 mL of 0.1xSSC/0.1xSDS three times to remove excess probe. After air-drying, the membrane was exposed on phosphorimager screen and imaged using a Typhoon scanner. Quantification of the blot intensities was accomplished using ImageJ 1.45S.

Single Molecule Experiments

Slide preparation: Quartz slides (Finkenbeiner Inc.) were cleaned and pegylated using the following protocol[57]. The slides were boiled in water to remove parafilm and coverslips from previous experiments. The slides were then scrubbed with

alconox (Alconox Inc.), rinsed, and placed into a solution containing 10% alconox. After sonication for 20 minutes, the slides were rinsed and placed in water and sonicated for five minutes. Slides were then sonicated in acetone for 15 minutes. The slides were directly transferred into 1M KOH and sonicated for 20 minutes. After thorough rinsing, the slides surfaces were flame dried and cleaned using a butane torch (BernzOmatic). While the slides were cooling, a silanizing solution containing 100 mL of methanol, 5 mL of glacial acetic acid, and 1 mL of N-(2-aminoethyl)-3-aminopropyltrimethoxysilane (UCT). The cooled slides were placed in the solution, sonicated for one minute, and allowed to stand in solution at room temperature for at least 20 minutes. While incubating slides 200 mg of mPEG-Succinimidyl Valerate MW 5,000 (Laysan Bio, Inc.), was resuspended in 400 μL of 0.1 M $NaHCO_3$. Also 2 mg of Biotin-PEG-Succinimidyl Valerate MW 5,000 (Laysan Bio, Inc.) was resuspended in 200 μL of 0.1 M $NaHCO_3$. After rinsing the slides with dH_2O and drying with nitrogen, the solutions were mixed and applied to each slide surface (150 μL per slide). A coverslip was used to cover the solution and the slides incubated overnight at room temperature in a humidor box. The following day, channels were assembled using Paramfilm as a spacer on the pegylated quartz slides and plasma cleaned coverslips were used as the second chamber face[55].

Incubation and molecule deposition: To prepare the slides for molecule deposition, the channels were incubated with 60 μL of 10 mg/mL BSA (NEB) for 20 minutes. The slide was then washed with 100 μL of T50 (10 mM Tris-HCl pH8, 50 mM NaCl) and then incubated with 60 μL of 0.2 mg/mL streptavidin in T50 for 5 minutes. After

coating the slide with streptavidin, the excess protein was washed away with 500 μL of T50. While preparing the slide, 10 μL of purified enzyme was incubated with 1 μL of 10 nM biotinylated primer for one hour at room temperature. After incubation, the molecules were diluted with 20 μL of imaging buffer (50 mM Tris HCl pH 8.3, 3 mM $MgCl_2$, 1 mM spermidine, 0.5 mg/mL BSA, 4% glucose), and flowed over the slide surface. The number of molecules on the surface was observed in real-time to reach the desired density for smFRET measurements. Once telomerase molecules were deposited on the surface, the stable complexes were retained and imaged after multiple buffer exchanges. Imaging of the FRET-labeled enzymes required an oxygen scavenging system and a triplet state quencher to increase dye lifetime. Imaging buffer did not substantially impact telomerase activity (Supplementary Fig. 2, lanes 14 and 15). The buffer was saturated with Trolox (triplet state quencher), passed through a 0.22 μm filtration unit, and brought to pH 8.3. Once ready for imaging, the solution was brought to 2 μg/mL catalase and 1 mg/mL glucose oxidase. The pH was found to be stable for more than 30 minutes with these conditions (data not shown).

DNA primer extension by surface-immobilized telomerase-DNA complexes:

After telomerase-DNA complexes were immobilized on the microscope slide, the indicated dNTPs or combination of dNTPs + ddNTPs were added as described in the text at a concentration of 200 μM. The reaction was allowed to proceed for 30 minutes prior to washing the channel with imaging buffer and collecting single molecule FRET data.

Microscope Setup:

Data was collected using a custom built prism-type microscope, allowing hundreds of molecules to be studied in parallel[54]. Cy3 donor dyes were directly excited with a 532 nm laser (Laserglow). Depending on the proximity Cy5 acceptor dye to the donor dye, a variable degree of energy transfer will occur from the excited donor dye to the acceptor dye. Due to this FRET process, both dyes are capable of fluorescing and emit at distinct wavelengths. Emitted fluorescence was split into separate donor and acceptor channels using dichroic mirrors, and imaged on an EM-CCD camera (Andor Ixon).

Data Acquisition and Analysis:

Imaging fields containing 30-250 molecules were imaged using a 100 ms integration time for fast timescale traces and histograms, and 500 ms integration time for slow timescale traces. Individual traces were parsed out using custom written IDL software where they were corrected for background and dye-crosstalk signal (software available upon request). The individual traces were then filtered in MATLAB using a thresholding analysis where molecules containing no acceptor dye were discarded. FRET intensities were then calculated using the equation $I_A /(I_A + \gamma I_D)$ where I_A is the acceptor intensity, I_D is the donor intensity, and γ is the gamma correction factor. The gamma correction factor was determined as described[57] for a subset of ~100 molecules, averaged, and applied to the entire dataset. The first 5 seconds of individual FRET traces were then binned into FRET histograms. Zero peaks resulting from premature photobleaching were subsequently removed by either subtraction of a Gaussian fit from the data set, or deletion of the data entirely where applicable. The

center of the FRET distributions were determined by a non-linear Gaussian fitting algorithm and reported next to each fit.

HaMMy and Dwell Time Analysis:

Hidden Markov Modeling (HaMMy, http://bio.physics.illinois.edu/HaMMy.html) was used for identification and fitting of FRET states. Fits were brought into MATLAB where the idealized FRET traces generated by HaMMy were parsed into individual dwell times and binned in an array that was used to generate dwell time histograms. Dwell time histograms were fit with a single exponential decay function yielding a characteristic relaxation time, τ, which is equivalent to the inverse of the rate constant.

Exonuclease VII Experiments

Initially 1 μM primer was mixed with 10 μL of purified telomerase and incubated for 30 min at RT to promote binding. A master mixture was made and added to the telomerase solution providing final conditions of activity buffer, 1 μM α^{32}P-dGTP (Perkin Elmer BLU014Z500UC), and 200 μM ddTTP. In the case of Exonuclease VII digestion, telomerase activity reactions were allowed to proceed for the indicated length of time and 20 U of Exonuclease VII (Epicentre EN510250) was added and incubated at room temperature for 5 minutes. The reaction was quenched with 80 μL stopping buffer (10 mM Tris HCl pH 7.5, 1 mM EDTA, 0.1% w/v SDS). After phenol/chloroform extraction the products were ethanol precipitated, resolved on a 12% denaturing PAGE gel (19:1 cross-linking ratio), and exposed on a phosphorimager screen overnight. Quantification of the bands was accomplished by

using SAFA v11b[56]. The lane profiles seen in Figure 6b and c and Supplementary Figure 6b were generated by taking a cross section of each lane in ImageJ and normalizing the values to the total lane counts.

RESULTS

A single-molecule assay for human telomerase-DNA interactions

To analyze telomerase-DNA interactions by smFRET, enzyme complexes were reconstituted *in vitro* using a previously reported two-piece RNA system[29] (Pseudoknot nt 32-195 and CR4/CR5 nt 239-328), harboring a site-specific Cy3 (donor) modification at position U42, located 5' of the RNA template (Fig. 1b and Supplementary Fig.1). Total internal reflection fluorescence (TIRF) microscopy was used to measure the distance-dependent energy transfer between the Cy3 dye within hTR and a Cy5 (acceptor) label within the telomere DNA primer, reporting on the relative positions of the DNA primer and the hTR subunit within the complex (Fig. 1c). We recently reported a similar experimental design for the *T. thermophila* telomerase enzyme[30,31]; however, the capacity of human telomerase to bind certain DNA primer sequences with extremely high affinity ($t_{1/2}$ > 20 hours[32]) permits purification of the active telomerase fraction[33] and greatly facilitates the single-molecule experiments in the present study. Importantly, all FRET modified telomerase enzymes and telomere DNA primers supported near wild-type levels of activity and processivity as measured by direct DNA primer extension assays (Supplementary Fig. 2). For our initial measurements, we utilized an 18 nucleotide (nt) telomere DNA primer terminating in 'GGG' at the 3'-end (18GGG). The primer

was site-specifically modified with an internal Cy5 dye covalently coupled to a thymine at position 13 (from the 5' end) in the DNA primer (18GGG/T13, Fig. 2a). This DNA sequence permutation promotes a high affinity telomerase-DNA interaction[32] and permits efficient primer-dependent surface immobilization of telomerase complexes onto the streptavidin coated microscope slide using a 5' biotin moiety (Fig. 1c and Supplementary Fig. 3). Single-molecule FRET trajectories revealed a stable FRET = 0.59 state and exhibited single-step photobleaching of the FRET dyes (Fig. 1c and Fig. 2b), characteristic of a single telomerase complex bound to a DNA primer. Thus, human telomerase functions as a monomer under the conditions of our assay, consistent with previous single-molecule studies of *in vitro* reconstituted human telomerase enzymes[34], but in contrast to several reports that have demonstrated human telomerase can exist as a functional dimer when reconstituted *in vivo*[35,36].

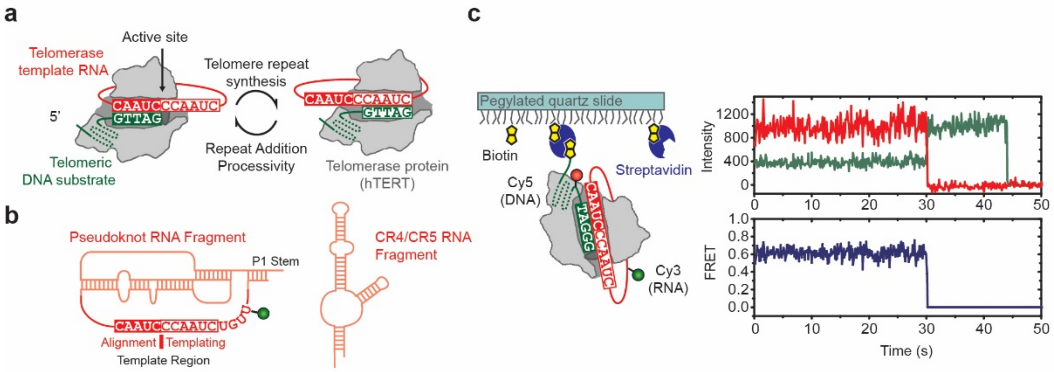


Figure 1. A single-molecule FRET assay for monitoring telomerase-DNA interactions. a, A model for telomere DNA repeat synthesis and repeat addition processivity (RAP) during the telomerase catalytic cycle. The precise location of the primer outside of the template hybrid is not well characterized and is therefore represented by a dashed line to denote an arbitrary path for the primer as it exits complex. **b**, Cartoon schematics of the two hTR fragments used in telomerase

reconstitutions. The pseudoknot fragment is shown with a U42-Cy3 modification. **c**, (left) Schematic illustration of experimental geometry during smFRET telomerase binding experiments using TIRF microscopy. (right) A representative single molecule trajectory of U42-Cy3 labeled telomerase bound to the 18GGG/T13-Cy5 primer. (top) Raw donor Cy3 (green) and acceptor Cy5 (red) dye intensities and (bottom) calculated FRET values are plotted as a function of time. The sudden drop in the dye intensity traces represents irreversible photobleaching of the FRET dyes.

We next tested whether the surface-immobilized telomerase-DNA complexes were competent to catalyze DNA synthesis. At time zero, a smFRET histogram constructed from several hundred telomerase-DNA complexes displayed a single dominant FRET population centered at FRET = 0.59 (Fig. 2b and Supplementary Fig. 4a), consistent with the individual FRET traces (Fig. 1b). Addition of dNTPs to the sample chamber promotes a time-dependent shift of the FRET distribution to lower FRET values which ultimately centers at FRET = 0.07 after a 60 minute incubation (Supplementary Fig. 4b and 4c). Since the donor Cy3 dye is incorporated into the telomerase enzyme, the shift to lower FRET values cannot be explained by enzyme dissociation over time. Moreover, smFRET histograms constructed from telomerase-DNA complexes incubated for 60 minutes in the absence of dNTPs, or with a catalytically dead telomerase mutant (D868A)[37], show no deviation from the initial FRET distribution (Supplementary Fig. 4d-e). These data indicate the majority of the surface-immobilized telomerase-DNA complexes in our experiments are competent to bind and catalytically extend a DNA primer. Telomerase-catalyzed DNA synthesis manifests itself as a gradual shift of the smFRET distribution to an extremely low FRET state, suggesting the DNA primer, and possibly telomerase itself, undergoes

considerable rearrangements during multiple rounds of processive telomere repeat synthesis.

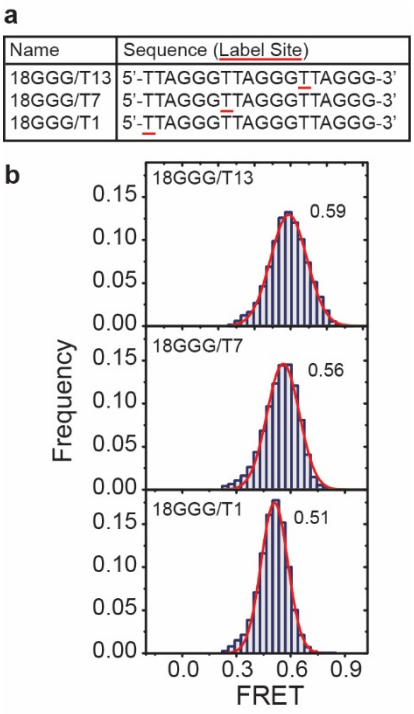


Figure 2. Telomere DNA substrate is compacted within the telomerase complex.
FRET histograms for the 5'-biotinylated 18GGG/T13-Cy5 (top), 18GGG/T7-Cy5 (middle), and 18GGG/T1-Cy5 (bottom) DNA primers. Histograms were fit with Gaussian functions (red lines) to determine the indicated centers of each FRET distribution. All histograms in the figure are constructed from at least 2000 individual molecules.

DNA rearrangements during telomere repeat synthesis

To further explore the conformation of the DNA bound within the telomerase-DNA complex, we next measured the smFRET distributions for a series of 18 nt 'GGG' terminating primers labeled with Cy5 at three different positions (T13, T7, and T1) (Fig. 2b). Again, all modified DNA primers supported near wild type level activity and processivity in DNA primer extension assays (Supplementary Fig. 2b). The smFRET histograms constructed from data collected with 18GGG/T13, 18GGG/T7, and 18GGG/T1 all yielded unimodal distributions, which were well-fit

by single Gaussian functions centered at FRET = 0.59, 0.56, and 0.51, respectively (Fig. 2b). These data suggest the entire 18 nt DNA primer assumes a compact conformation within the core telomerase complex. We next set out to analyze how the conformation of the DNA changes with the stepwise addition of dNTPs during telomere DNA repeat synthesis. To this end, surface-immobilized telomerase-DNA complexes were incubated for 30 minutes with specific combinations of dNTPs and ddNTPs to terminate the reaction at precisely controlled stages of repeat synthesis beginning with the high affinity 18GGG/T13 primer sequence (Supplementary Fig. 2b, lanes 8-11). To simplify interpretation of these smFRET measurements, we first tested whether the position of the U42-Cy3 label site within the hTR subunit is stationary throughout the repeat synthesis reaction. Doubly-labeled hTR molecules harboring U42-Cy3 and U29-Cy5 modifications were reconstituted into a catalytically active telomerase complex (Supplementary Figs. 1 and 2 lane 6). Notably, the immobilization of doubly-labeled telomerase complexes onto the microscope slide was DNA-primer dependent, as evidenced by the background level of telomerase complexes observed in the absence of biotinylated DNA primer (Supplementary Fig. 3). At each successive stage of telomere repeat synthesis, the FRET value measured between hTR positions U29 and U42 remained constant (Fig. 3a). This result is somewhat surprising in light of a similar experiment performed on the *Tetrahymena* telomerase enzyme which suggested that regions of RNA flanking the template undergo conformational rearrangements during the telomere repeat synthesis reaction[30]. Thus, the observation that hTR position U42 remains stationary

with respect to U29 (P1 stem) during telomere repeat synthesis may reflect a distinct structural property of the human telomerase enzyme (see Discussion).

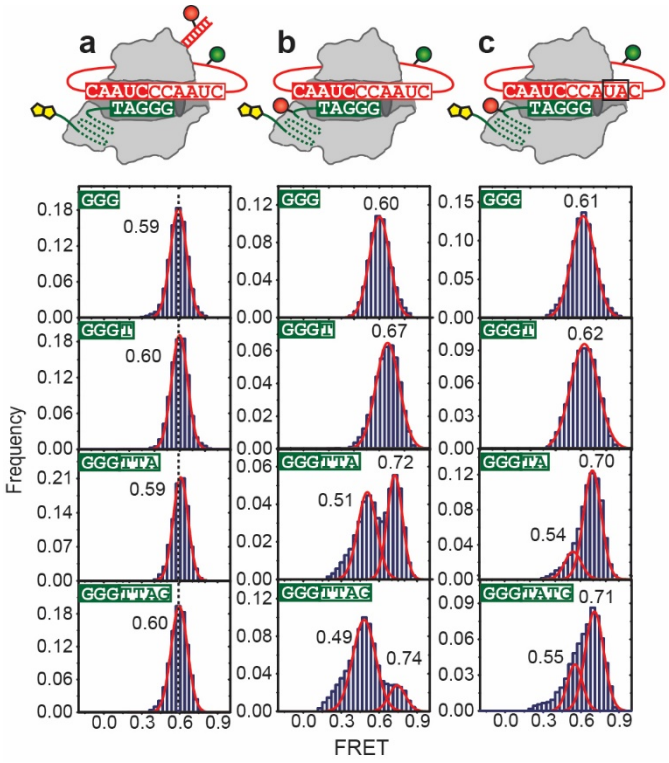


Figure 3. DNA conformational changes induced by completion of telomere DNA repeat synthesis.

a, FRET distributions constructed with doubly-labeled telomerase enzyme harboring hTR U29-Cy5 (P1 stem) and U42-Cy3 dyes initially bound to 5' biotinylated 18GGG primer. FRET histograms at each indicated step of the repeat synthesis reaction are shown and fit with single Gaussian functions (red lines) centered at FRET ~ 0.6 (dotted line). Initial 3' DNA sequence is highlighted in green and nascent DNA sequence is highlighted in white. **b**, FRET distributions constructed from U42-Cy3 labeled telomerase enzyme initially bound to 18GGG/T13-Cy5 DNA primer and extended to each indicated step of the repeat synthesis reaction are shown. Data were fit with either one or two Gaussian functions (red lines) to determine the indicated centers of each FRET distribution. **c**, FRET distributions constructed from U42-Cy3 labeled telomerase enzyme harboring the template mutations UA47-48AU, initially bound to the 18GGG/T13-Cy5 primer and extended to each indicated step of the repeat synthesis reaction are shown. Data were fit with either one or two Gaussian functions (red lines) to determine the indicated centers of each FRET distribution. All histograms in the figure are constructed from at least 500 individual molecules.

We next measured FRET between the U42-Cy3 and the DNA primer with the T13-Cy5 modification. The center of the initial FRET = 0.59 peak gradually increases, until the synthesized repeat terminates with 'TTA' or 'TAG', at which point a new significantly lower FRET = 0.49 state is observed (Fig. 3b). Analysis of individual smFRET trajectories revealed a dynamic interconversion between the FRET = 0.74 and FRET = 0.48 states for 18GGG/T13 primers extended to the complete telomere DNA repeat sequence (Fig. 4a). Single molecule FRET traces were fit using a hidden Markov model to yield idealized FRET trajectories[38], and the individual times spent in each FRET state (τ) were binned into dwell time distributions (Fig. 4b). Both distributions were well-fit by single exponential decay functions, yielding mean dwell times $< \tau >_{FRET=0.48} = 38 \pm 8.1$ seconds and $< \tau >_{FRET=0.74} = 10 \pm 1.6$ seconds, corresponding to the rate constants $k_{0.48->0.74} = 0.026 \pm 0.0059$ sec^{-1} and $k_{0.74->0.48} = 0.098 \pm 0.015$ sec^{-1}. Notably, the value we measure for $k_{0.74->0.48}$ is ~100-fold faster than previously reported estimates for the rate constant governing RAP measured by ensemble primer extension assays under similar experimental conditions[39].

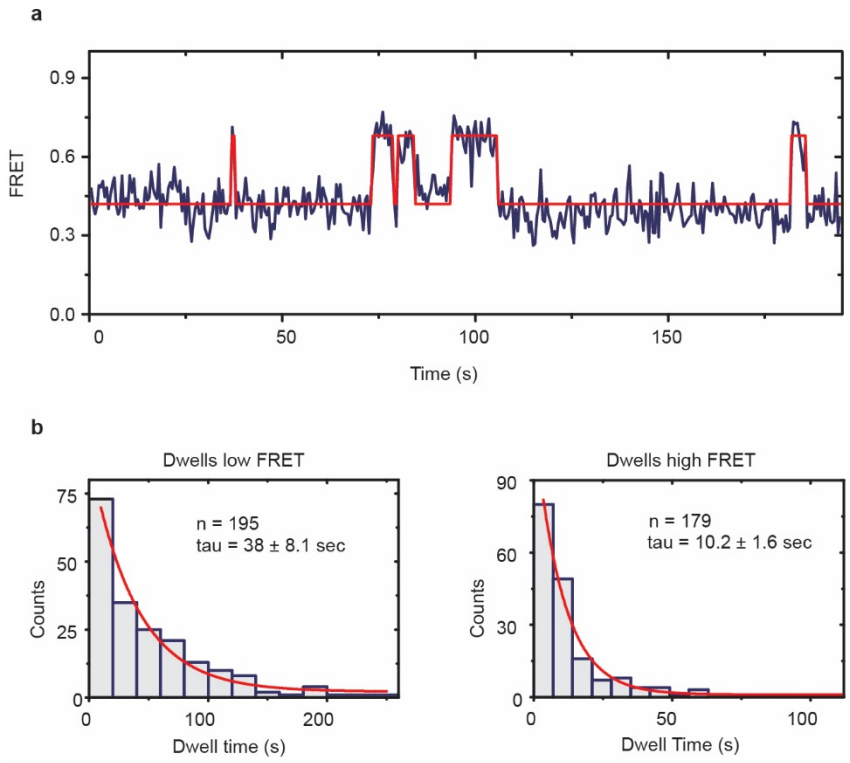


Figure 4. Kinetic analysis of DNA primer dynamics induced upon completion of a telomere DNA repeat. a, A representative single-molecule trace of a telomerase complex that has completed synthesis of a single telomere DNA repeat exhibiting multi-state FRET dynamics. The calculated FRET ratio is plotted as a function of time (blue line). Data were fit using the HaMMy software package to generate idealized FRET trajectories (red line). **b**, Dwell time histograms of the times spent in either the low FRET = 0.48 state (left) or high FRET = 0.74 state (right) are plotted and fit with single exponential decay functions yielding average dwell times $<\tau>$ of 38 ± 8.1 seconds and 10 ± 1.6 seconds for the low FRET and high FRET states, respectively. These mean dwell times correspond to the rate constants $k_{0.48->0.74}$ = 0.0026 ± 0.0059 sec^{-1} and $k_{0.74->0.48}$ = 0.098 ± 0.015 sec^{-1}. The sample sizes represent the number of individual transitions used in the analysis and error estimates represent standard error of the mean.

To further characterize the novel FRET = 0.48 conformation, we next utilized a mutant hTR (UA47-48AU) containing the altered template sequence, 3'-AAUCCCAUAC-5'. This mutant enzyme can bind the 18GGG/T13 primer with

wild-type affinity and extend to the end of the first telomere DNA repeat[40]. However, the mismatches introduced into the nascent DNA sequence disrupt base pairing with the downstream region of the hTR template and substantially abrogate RAP (Supplementary Fig. 2b, lane 7), as previously reported[40]. When the UA47-48AU enzyme was analyzed in the smFRET assay, we again observed an increase in the initial FRET = 0.61 state as seen with the wild type enzyme, but the stability of the low FRET conformation at the end of repeat synthesis was significantly reduced (Fig. 3c). These results provide evidence that the stability of the FRET = 0.48 intermediate relies upon correct base pairing between the nascent DNA product and the downstream region of the hTR template, and that formation of this previously uncharacterized translocation intermediate is not the rate-limiting step during RAP.

Telomere DNA movements during repeat addition processivity

To analyze DNA rearrangements required for the complete RAP cycle, we employed a single round translocation assay[39] to measure the DNA conformations during the initial binding, telomere DNA repeat synthesis, and RAP phases of the telomerase catalytic cycle (Fig. 5). For these experiments we used a 21 nt DNA primer with a T7-Cy5 modification terminating with the sequence 'TTA' (21TTA/T7, see Supplementary Table1). Upon initial binding we observed a bimodal FRET distribution consistent with our previous experiments performed with the 18GGG/T13 primer (compare Fig. 5a and Fig. 3b). Addition of a single ddGTP completes the first telomere repeat and further stabilizes the low FRET = 0.41 translocation intermediate, again consistent with our previous results (compare Fig.

5b and 3b). The presence of both dGTP and ddTTP should promote a single round of RAP during which telomerase completes the first telomere repeat, translocates, and then continues to add nucleotides until synthesis is terminated by incorporation of a ddTTP. However, RAP is not highly efficient in the absence of the *trans*-acting cofactors POT1-TPP1 (Supplementary Fig. 5a)[28,39]. Therefore, we established a technique to enrich for telomerase enzymes that have successfully completed one round of RAP by exploiting the difference in binding affinity for telomere DNA primers terminating in 'TAG' and 'GGT'[32]. Telomerase enzyme was pre-bound to the 5' biotinylated 21TTA/T7 primer in a centrifuge tube, and then activated to extend the primer by addition of dGTP and ddTTP in the presence of excess 18GGG chase primer (Supplementary Fig. 5b). In this way, complexes that successfully completed a RAP cycle converted the low-affinity 21TTA/T7 primer to a high affinity 25GGT/T7 primer and are selectively immobilized on the glass slide for smFRET measurements (Supplementary Fig. 5c). These single round RAP complexes yielded a unimodal FRET distribution centered at FRET = 0.49 (Fig. 5c), a value that is consistent with the conformation of the DNA primer established with the 18GGG series of primers (compare Fig. 5c with Fig. 2b 18GGG/T1). Taken together, these results suggest the translocation process during RAP is comprised of at least two distinct steps: (i) a rapid realignment of the 3'-end of the complete telomere DNA repeat with the downstream region of the hTR template (Fig. 5b); and (ii) a second, rate-limiting conformational rearrangement during which the realigned RNA:DNA hybrid is repositioned in the hTERT active site to prime the next round of repeat synthesis.

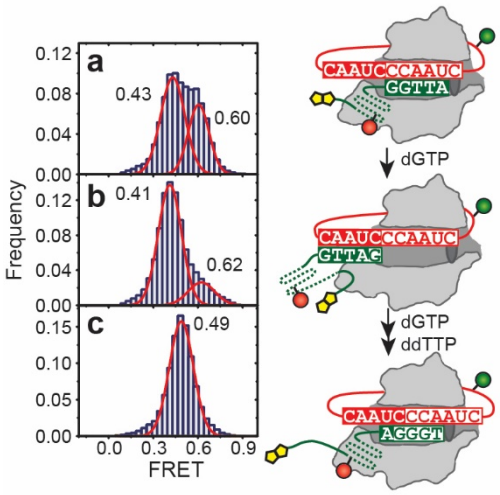


Figure 5. DNA conformational changes during repeat addition processivity. **a**, FRET histogram constructed from U42-Cy3 labeled telomerase bound to 21TTA/T7-Cy5 DNA primer. **b**, FRET histogram constructed from U42-Cy3 labeled telomerase initially bound to 21TTA/L2, surface immobilized on the microscope slide, and incubated in 200 μM ddGTP to complete the synthesis of a single telomere DNA repeat. **c**, FRET histogram for U42-Cy3 labeled telomerase enzyme initially bound to 21TTA/T7-Cy5, incubated in 200 μM dGTP and ddTTP to permit one round of repeat addition processivity in the presence of excess high-affinity chase 18GGG primer, followed by surface immobilization for FRET measurements as described in the main text and Supplementary Fig. 5. FRET distributions in all panels were fit with either single or double Gaussian functions (red lines) to determine the indicated position of each FRET distribution. All histograms in the figure are constructed from at least 700 individual molecules.

5' DNA primer extrusion during repeat addition processivity

The similarity of the observed FRET conformations for the 18GGG/T1 (Fig. 2) and the 25GGT/T7 (Fig. 5c) primers raised the intriguing possibility that during RAP, repositioning of the RNA:DNA hybrid is coupled with 5' DNA extrusion out of the enzyme complex. While suggestive, our smFRET results alone are not sufficient to firmly support this model for coordinated movements of the 3' and 5' ends of the DNA primer. Therefore, we used a recently established footprinting assay that

exploits the 5'→ 3' nuclease activity of ExoVII to probe how much telomere DNA is protected when bound to telomerase[41]. For these experiments, all telomere DNA primers possessed at least three copies of natural telomere DNA sequence and a 10 nt 5'-polyT extension. The use of a polyT sequence to extend the length of the primer was strategically chosen to avoid complications introduced by the formation of a stable intramolecular G-quadruplex in the DNA primer [42,43]. In addition, primers were added at ~200-fold greater concentrations than telomerase to minimize the possibility of multiple enzymes binding a single DNA substrate in our assays, thereby simplifying interpretation of the footprinting data. Telomere DNA primers pre-bound to telomerase were activated for extension by the addition of dGTP [α^{32}P] and ddTTP for the indicated period of time, followed by a five minute treatment with ExoVII to digest exposed 5'-single stranded DNA (Fig. 6a). ExoVII treatment was limited to five minutes to ensure the signal obtained was derived from stable telomerase-DNA complexes, rather than probing DNA that dissociated from the telomerase enzyme. Importantly, since telomerase incorporates radio-labeled dGTP into the cold DNA primer, all bands observed on the gel represent products of active telomerase complexes.

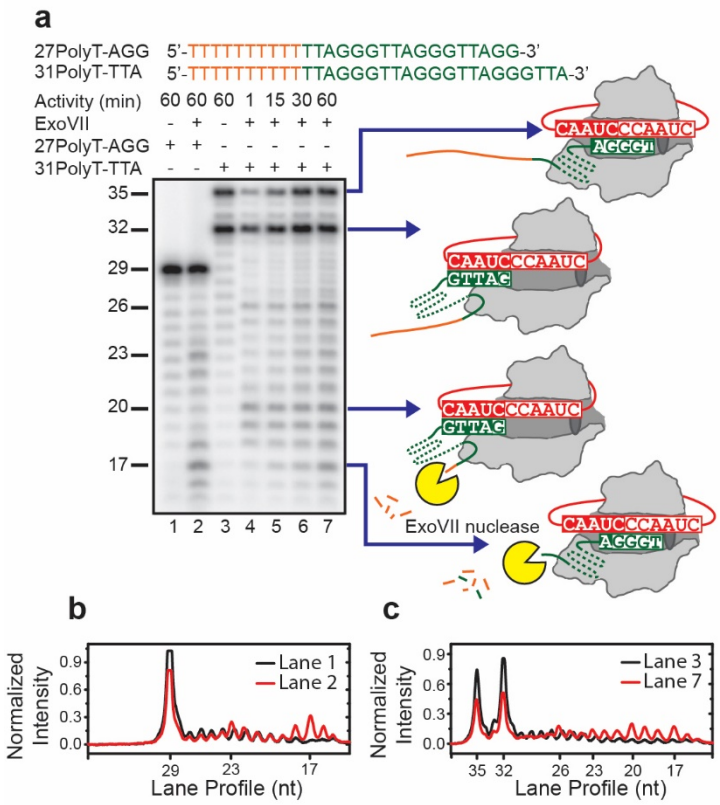


Figure 6. Exonuclease VII digestion reveals extrusion of 5' DNA during repeat addition processivity. a, Autoradiogram of a 12% denaturing PAGE gel resolving the products of telomerase-catalyzed primer extension of the indicated telomere DNA primers (27PolyT-AGG or 31PolyT-TTA). Telomerase extension products of the 27PolyT-AGG primer in the presence of 200 μM dGTP [α^{32}P] and ddTTP, untreated (Lane 1) or treated (Lane 2) with ExoVII. (Lane 3) Telomerase extension products of the 31PolyT-TTA primer in the presence of dGTP [α^{32}P] and ddTTP. (Lanes 4-7) Time-course of telomerase extension of 31PolyT-TTA in the presence of dGTP [α^{32}P] and ddTTP followed by ExoVII treatment. **b**, Lane profiles for 27PolyT-AGG extended primer untreated (Lane 1, black) and ExoVII treated (Lane 2, red). **c**, Lane profiles for 31PolyT-TTA extended primer untreated (Lane 3, black) and ExoVII treated (Lane 7, red). All lane profiles have been normalized by the total intensity of the lane.

We first used a 27 nt primer terminating with the sequence 'AGG' (27PolyT-AGG, Fig. 6a), which under the conditions of our experiments will incorporate a single dGTP and then terminate with ddTTP to yield a single 29 nt product (Fig. 6a,

lane 1). We observed a small amount of shorter non-specific degradation products not present in the untreated DNA primer sample (Supplementary Fig. 6a), which is likely a result of a minor nuclease contaminant. When DNA in this complex is partially digested with ExoVII, we observe two regions of protection centered about a predominant 17 nt band and a second 23 nt band (Fig. 6a, lane 2 and Fig. 6b). Interestingly, the length of the 23 nt protected fragment suggests the 5'-end of the DNA primer can adopt two conformations that maintain a 6 nt spacing. To address the question of what happens to the 5'-end of the DNA during a single round of RAP, we used a 31 nt primer terminating with the sequence 'TTA' (31PolyT-TTA, Fig. 6a), which is compatible with single round translocation experiments. In the absence of ExoVII, extension of this primer produces two predominant products: one 32 nt band corresponding to the addition of one dGTP which completes synthesis of the full telomere repeat, and a second 35 nt band representing successful RAP and extension of the next telomere repeat until termination by ddTTP incorporation (Fig. 6a, lane 3). At early stages of a reaction time course with the 31PolyT-TTA primer, the 32 nt product accumulates, and when treated with ExoVII, gives rise to protected fragments centered about a 20 nt and 26 nt (Fig. 6a, lane 4). These products precisely correspond to the 5'-protection pattern we observed for the 27PolyT-AGG primer, suggesting the 5' most contact is maintained throughout telomere repeat synthesis. As the reaction proceeds, the amount of the 35 nt RAP product increases, and when these complexes are treated with ExoVII, two additional protected fragments are observed centered about 17 nt and 23 nt (Fig. 6a, lane 7 and Fig. 6c). The accumulation of

these new cleavage products correlates with accumulation of the 35 nt extension product (Supplementary Fig. 6b), supporting the notion that these shorter Exo VII fragments are derived from digestion of primers that have undergone RAP. As described above, RAP is not highly efficient under the conditions of our assay, and therefore ExoVII products generated for the single round translocation reactions represent a superimposition of products derived from complexes that have stalled at the end of telomere repeat synthesis and those that have undergone a single RAP cycle. To de-convolute these cleavage patterns, we performed a control experiment with a 32 nt primer terminating with 'TAG' (32PolyT-TAG) to enrich for the 35 nt RAP product, which when treated with ExoVII exclusively produced the protected fragments centered about 17 nt and 23 nt (Supplementary Fig. 6c). These results provide direct physical evidence that extrusion of the 5' end of the DNA out of the enzyme is coupled to completion of a RAP cycle.

DISCUSSION

We present a model for the dynamic handling of DNA during telomerase catalysis that integrates our results together with existing biochemical data (Fig. 7). In this model, telomerase initially binds its telomere DNA substrate via a combination of base pairing interactions with the template region of the integral hTR subunit, together with distinct contacts in the 5'-region of the DNA substrate. Once bound, a single telomere DNA repeat is rapidly synthesized. Interestingly, our smFRET approach revealed that upon completion of the telomere DNA repeat synthesis reaction, a dynamic equilibrium exists, wherein the 3'-end of the nascent DNA primer

transiently samples different alignment registers of the hTR template (ii). Our novel single-molecule assay provided the first direct quantitative characterization of this obligatory step during telomerase repeat addition processivity (RAP). Existing models for telomerase catalysis suggest the denaturation of the nascent RNA:DNA hybrid is the kinetically slow step during RAP[12,44–46] ; however, our data demonstrate this step is fast compared to the overall rate for RAP that has been reported in the literature[39]. Therefore, to reconcile our smFRET results with ensemble assays for RAP, we propose a subsequent rate-limiting conformational rearrangement of the telomerase complex serves to reposition the re-aligned RNA:DNA hybrid into the hTERT active site to prime the subsequent round of repeat addition (iv). During this slow conformational rearrangement the 5' end of the DNA primer is extruded out of the enzyme complex. This model is consistent with the recent report by Qi et al. that inferred rebinding of the RNA:DNA duplex in the hTERT active site may be required during RAP from *trans* template complementation experiments[47].

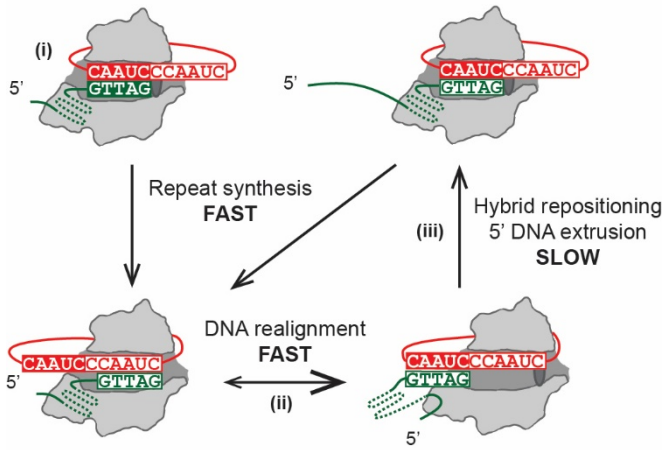


Figure 7. Model for DNA conformational rearrangements during telomerase catalytic cycle. (i)Telomerase initially binds a telomere DNA substrate through RNA:DNA hybrid interactions with the template and additional 5'-DNA contacts. Rapid synthesis of a complete telomere DNA repeat extends the RNA:DNA hybrid to the template boundary, at which point a dynamic equilibrium of DNA primer states exists (ii), favoring the formation of a re-aligned RNA:DNA hybrid. (iii) In a second, rate-limiting, step of repeat addition processivity the re-aligned RNA:DNA hybrid is repositioned in the telomerase active site and the 5'-end of the DNA primer is extruded from the enzyme complex.

A recent smFRET analysis of *Tetrahymena* telomerase revealed the RNA template and/or flanking regions undergo conformational rearrangements during telomere repeat synthesis[30]. These previous experiments would appear to contradict the present finding that hTR-U42 is stationary throughout the repeat synthesis reaction. However, it is possible that the template region, rather than the 5' flanking sequence, is the major source of structural compliance within telomerase RNA during telomere repeat synthesis. An alternative, but not mutually exclusive, explanation for the observed smFRET results may relate to differences in the mechanism of template boundary definition for the human and *Tetrahymena* telomerase enzymes. This notion is in accordance with recent biochemical evidence from Wu et al. which demonstrated that template definition in human telomerase is an intrinsic property of RNA:DNA hybrid recognition[41], rather than being externally defined by adjacent protein-RNA contacts as described for the *Tetrahymena* system[48].

The telomerase enzyme used for the experiments in the present work was reconstituted using the well-established rabbit reticulocyte lysate system[37,49]. This telomerase reconstitution approach has been reported to yield monomeric enzyme complexes[34], consistent with our single molecule FRET trajectories which exhibit

single-step photobleaching behavior. In contrast, human telomerase reconstituted in human cell lines can produce a dimeric telomerase complex[33], though the structural basis for dimerization remains unknown. Despite this difference in oligomerization state, a recent study showed that *in vitro* and *in vivo* reconstituted telomerase complexes exhibit similar functional properties with respect to DNA primer handling[41]. In addition, a recent analysis of many disease-associated alleles of telomerase components led to the conclusion that the individual monomers within the dimeric telomerase complex act independently[36]. However, a separate study that analyzed the catalytic properties of telomerase complexes harboring one active and one catalytically dead subunit found that a non-functional mutant telomerase monomer can act in a dominant-negative fashion and abrogate the catalytic activity of the wild type monomer[35]. While the source of this discrepancy remains unknown, the majority of available experimental evidence supports the notion that the DNA handling properties characterized for the monomeric telomerase complexes used in the present study will generally apply to the individual monomers within a telomerase dimer.

The regulated extrusion of the DNA primer out of the telomerase complex has direct implications for the molecular mechanism of the *trans*-acting telomerase cofactor POT1/TPP1, which dramatically increases the rate and efficiency of RAP[28,39]. Previous experiments utilizing a mutant hTR template that produced a telomere DNA repeat sequence incapable of binding POT1/TPP1 demonstrated that a DNA primer harboring a single 5' wild-type POT1/TPP1 binding site could be

processively extended, suggesting POT1/TPP1 may remain stably associated with telomerase during multiple rounds of RAP[39]. Subsequently, the so-called 'TEL patch' was shown to mediate the interaction of POT1/TPP1 with telomerase[50], and single-molecule experiments revealed that POT1/TPP1 can dynamically slide along telomere DNA primers[51]. Our results suggest that six nucleotides of DNA are extruded from the telomerase complex with each round of RAP, an amount of telomere DNA that is smaller than the established ten nucleotide binding site size of POT1[28,39] . Thus, as DNA is released with each round of RAP, the POT1/TPP1 dimer associated with telomerase would be expected to dynamically slide along the growing DNA chain before another POT1/TPP1 complex can bind.

REFERENCES

1. De Lange, T. Protection of mammalian telomeres. *Oncogene* **21,** 532–540 (2002).

2. Blackburn, E. H. Switching and signaling at the telomere. *Cell* **106,** 661–73 (2001).

3. Counter, C. M. *et al.* Telomerase activity is restored in human cells by ectopic expression of hTERT (hEST2), the catalytic subunit of telomerase. *Oncogene* **16,** 1217–22 (1998).

4. Kim, N. *et al.* Specific association of human telomerase activity with immortal cells and cancer. *Science* **266,** 2011–2015 (1994).

5. Greider, C. W. & Blackburn, E. H. The telomere terminal transferase of Tetrahymena is a ribonucleoprotein enzyme with two kinds of primer specificity. *Cell* **51,** 887–98 (1987).

6. Morin, G. B. The human telomere terminal transferase enzyme is a ribonucleoprotein that synthesizes TTAGGG repeats. *Cell* **59,** 521–9 (1989).

7. Wyatt, H. D. M., West, S. C. & Beattie, T. L. InTERTpreting telomerase structure and function. *Nucleic Acids Res.* **38,** 5609–22 (2010).

8. Theimer, C. & Feigon, J. Structure and function of telomerase RNA. *Curr. Opin. Struct. Biol.* **16,** 307–18 (2006).

9. Greider, C. W. & Blackburn, E. H. A telomeric sequence in the RNA of Tetrahymena telomerase required for telomere repeat synthesis. *Nature* **337,** 331–7 (1989).

10. Lingner, J. *et al.* Reverse Transcriptase Motifs in the Catalytic Subunit of Telomerase. *Science* **276,** 561–567 (1997).

11. Blackburn, E. H. *et al.* Recognition and elongation of telomeres by telomerase. *Genome* **31,** 553–60 (1989).

12. Greider, C. Telomerase is processive. *Mol. Cell. Biol.* **11,** (1991).

13. Zhao, Y. *et al.* Processive and distributive extension of human telomeres by telomerase under homeostatic and nonequilibrium conditions. *Mol. Cell* **42,** 297–307 (2011).

14. Morin, G. B. Recognition of a chromosome truncation site associated with alpha-thalassaemia by human telomerase. *Nature* **353,** 454–6 (1991).

15. Collins, K. & Greider, C. W. Tetrahymena telomerase catalyzes nucleolytic cleavage and nonprocessive elongation. *Genes Dev.* **7,** 1364–1376 (1993).

16. Hammond, P. W., Lively, T. N. & Cech, T. R. The anchor site of telomerase from Euplotes aediculatus revealed by photo-cross-linking to single- and double-stranded DNA primers. *Mol. Cell. Biol.* **17,** 296–308 (1997).

17. Romi, E. *et al.* High-resolution physical and functional mapping of the template adjacent DNA binding site in catalytically active telomerase. *Proc. Natl. Acad. Sci. U. S. A.* **104,** 8791–6 (2007).

18. Peng, Y., Mian, I. S. & Lue, N. F. Analysis of telomerase processivity: mechanistic similarity to HIV-1 reverse transcriptase and role in telomere maintenance. *Mol. Cell* **7,** 1201–11 (2001).

19. Wyatt, H. D. M., Lobb, D. & Beattie, T. L. Characterization of physical and functional anchor site interactions in human telomerase. *Mol. Cell. Biol.* **27,** 3226–40 (2007).

20. Gillis, A. J., Schuller, A. P. & Skordalakes, E. Structure of the Tribolium castaneum telomerase catalytic subunit TERT. *Nature* **455,** 633–7 (2008).

21. Collins, K. Single-stranded DNA repeat synthesis by telomerase. *Curr. Opin. Chem. Biol.* **15,** 643–8 (2011).

22. Stratmann, S. a & van Oijen, a M. DNA replication at the single-molecule level. *Chem. Soc. Rev.* **43,** 1201–20 (2014).

23. Dangkulwanich, M., Ishibashi, T., Bintu, L. & Bustamante, C. Molecular Mechanisms of Transcription through Single-Molecule Experiments. *Chem. Rev.* **114,** 3203–23 (2014).

24. Kapanidis, A. N. *et al.* Initial transcription by RNA polymerase proceeds through a DNA-scrunching mechanism. *Science* **314,** 1144–7 (2006).

25. Andrecka, J. *et al.* Single-molecule tracking of mRNA exiting from RNA polymerase II. *Proc. Natl. Acad. Sci. U. S. A.* **105,** 135–40 (2008).

26. Liu, S., Abbondanzieri, E. A., Rausch, J. W., Le Grice, S. F. J. & Zhuang, X. Slide into action: dynamic shuttling of HIV reverse transcriptase on nucleic acid substrates. *Science* **322,** 1092–7 (2008).

27. Abbondanzieri, E. *et al.* Dynamic binding orientations direct activity of HIV reverse transcriptase. *Nature* **453,** 184–9 (2008).

28. Wang, F. *et al.* The POT1-TPP1 telomere complex is a telomerase processivity factor. *Nature* **445,** 506–10 (2007).

29. Chen, J. & Greider, C. W. Template boundary definition in mammalian telomerase. *Genes Dev.* **17,** 2747–52 (2003).

30. Berman, A. J., Akiyama, B. M., Stone, M. D. & Cech, T. R. The RNA accordion model for template positioning by telomerase RNA during telomeric DNA synthesis. *Nat. Struct. Mol. Biol.* **18,** 1371–5 (2011).

31. Wu, J. Y., Stone, M. D. & Zhuang, X. A single-molecule assay for telomerase structure-function analysis. *Nucleic Acids Res.* **38,** 1–11 (2010).

32. Wallweber, G., Gryaznov, S., Pongracz, K. & Pruzan, R. Interaction of human telomerase with its primer substrate. *Biochemistry* **42,** 589–600 (2003).

33. Cohen, S. B. *et al.* Protein composition of catalytically active human telomerase from immortal cells. *Science* **315,** 1850–3 (2007).

34. Alves, D. *et al.* Single-molecule analysis of human telomerase monomer. *Nat. Chem. Biol.* **4,** 287–9 (2008).

35. Sauerwald, A. *et al.* Structure of active dimeric human telomerase. *Nat. Struct. Mol. Biol.* **20,** 454–60 (2013).

36. Zaug, A. J., Crary, S. M., Jesse Fioravanti, M., Campbell, K. & Cech, T. R. Many disease-associated variants of hTERT retain high telomerase enzymatic activity. *Nucleic Acids Res.* **41,** 8969–78 (2013).

37. Weinrich, S. L. *et al.* Reconstitution of human telomerase with the template RNA component hTR and the catalytic protein subunit hTRT. *Nat. Genet.* **17,** 498–502 (1997).

38. McKinney, S., Joo, C. & Ha, T. Analysis of single-molecule FRET trajectories using hidden Markov modeling. *Biophys. J.* **91,** 1941–51 (2006).

39. Latrick, C. M. & Cech, T. R. POT1-TPP1 enhances telomerase processivity by slowing primer dissociation and aiding translocation. *EMBO J.* **29,** 924–33 (2010).

40. Drosopoulos, W. C., Direnzo, R. & Prasad, V. R. Human telomerase RNA template sequence is a determinant of telomere repeat extension rate. *J. Biol. Chem.* **280,** 32801–10 (2005).

41. Wu, R. A. & Collins, K. Human telomerase specialization for repeat synthesis by unique handling of primer-template duplex. *EMBO J.* 1–16 (2014).

42. Lipps, H. J. & Rhodes, D. G-quadruplex structures: in vivo evidence and function. *Trends Cell Biol.* **19,** 414–22 (2009).

43. Burge, S., Parkinson, G. N., Hazel, P., Todd, A. K. & Neidle, S. Quadruplex DNA: sequence, topology and structure. *Nucleic Acids Res.* **34,** 5402–15 (2006).

44. Romero, D. P. & Blackburn, E. H. A conserved secondary structure for telomerase RNA. *Cell* **67,** 343–53 (1991).

45. Blackburn, E. H. & Collins, K. Telomerase: an RNP enzyme synthesizes DNA. *Cold Spring Harb. Perspect. Biol.* **3,** 1–9 (2011).

46. Lue, N. F. Adding to the ends: what makes telomerase processive and how important is it? *Bioessays* **26,** 955–62 (2004).

47. Qi, X. *et al.* RNA/DNA hybrid binding affinity determines telomerase template-translocation efficiency. *EMBO J.* **31,** 150–61 (2012).

48. Lai, C. K., Miller, M. C. & Collins, K. Template boundary definition in Tetrahymena telomerase. *Genes Dev.* **16,** 415–20 (2002).

49. Beattie, T. L., Zhou, W., Robinson, M. O. & Harrington, L. Reconstitution of human telomerase activity in vitro. *Curr. Biol.* **8,** 177–80 (1998).

50. Nandakumar, J. *et al.* The TEL patch of telomere protein TPP1 mediates telomerase recruitment and processivity. *Nature* (2012).

51. Hwang, H., Buncher, N., Opresko, P. L. & Myong, S. POT1-TPP1 Regulates Telomeric Overhang Structural Dynamics. *Structure* 1–9 (2012).

52. Tesmer, V. M. *et al.* Two inactive fragments of the integral RNA cooperate to assemble active telomerase with the human protein catalytic subunit (hTERT) in vitro. *Mol. Cell. Biol.* **19,** 6207–16 (1999).

53. Mitchell, J. R. & Collins, K. Human telomerase activation requires two independent interactions between telomerase RNA and telomerase reverse transcriptase. *Mol. Cell* **6,** 361–71 (2000).

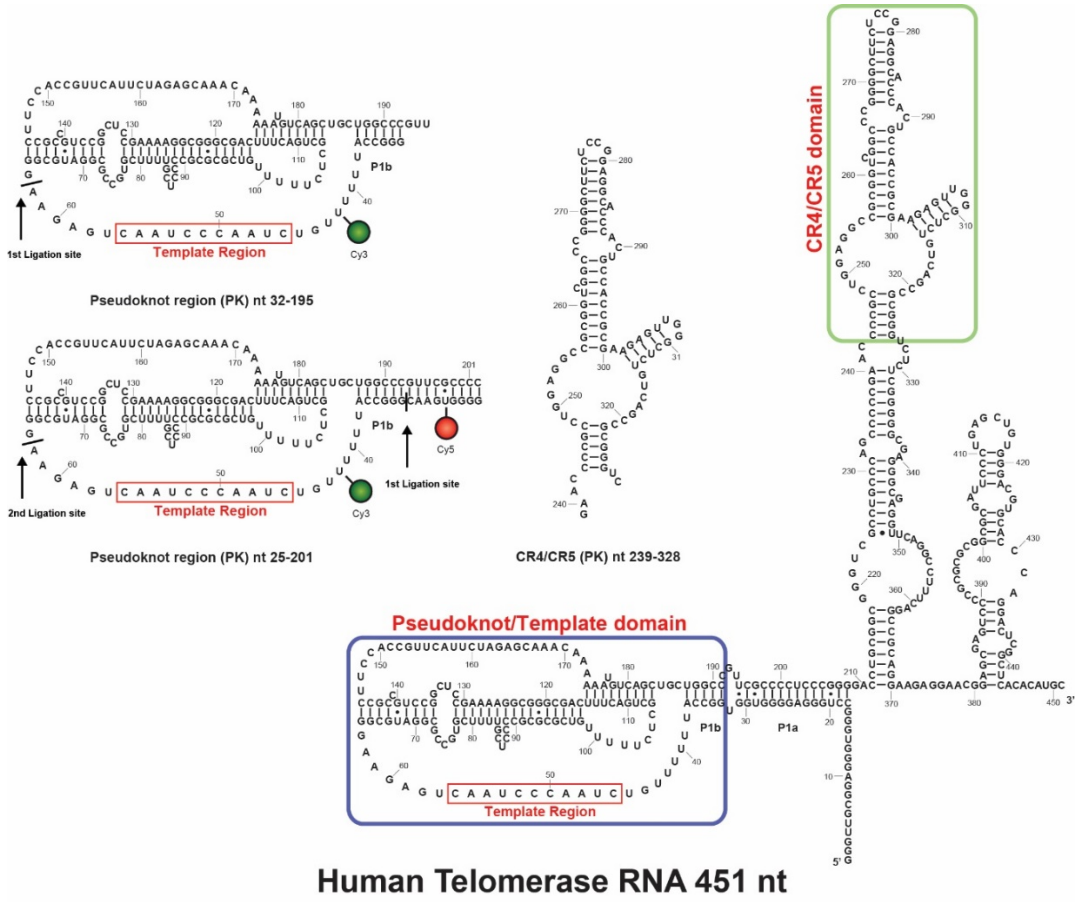
54. Akiyama, B. M. & Stone, M. D. *Assembly of complex RNAs by splinted ligation. Methods Enzymol.* **469,** 27–46 (2009).

55. Hengesbach, M., Kim, N. K., Feigon, J. & Stone, M. D. Single-molecule FRET reveals the folding dynamics of the human telomerase RNA pseudoknot domain. *Angew. Chem. Int. Ed. Engl.* **51,** 5876–9 (2012).

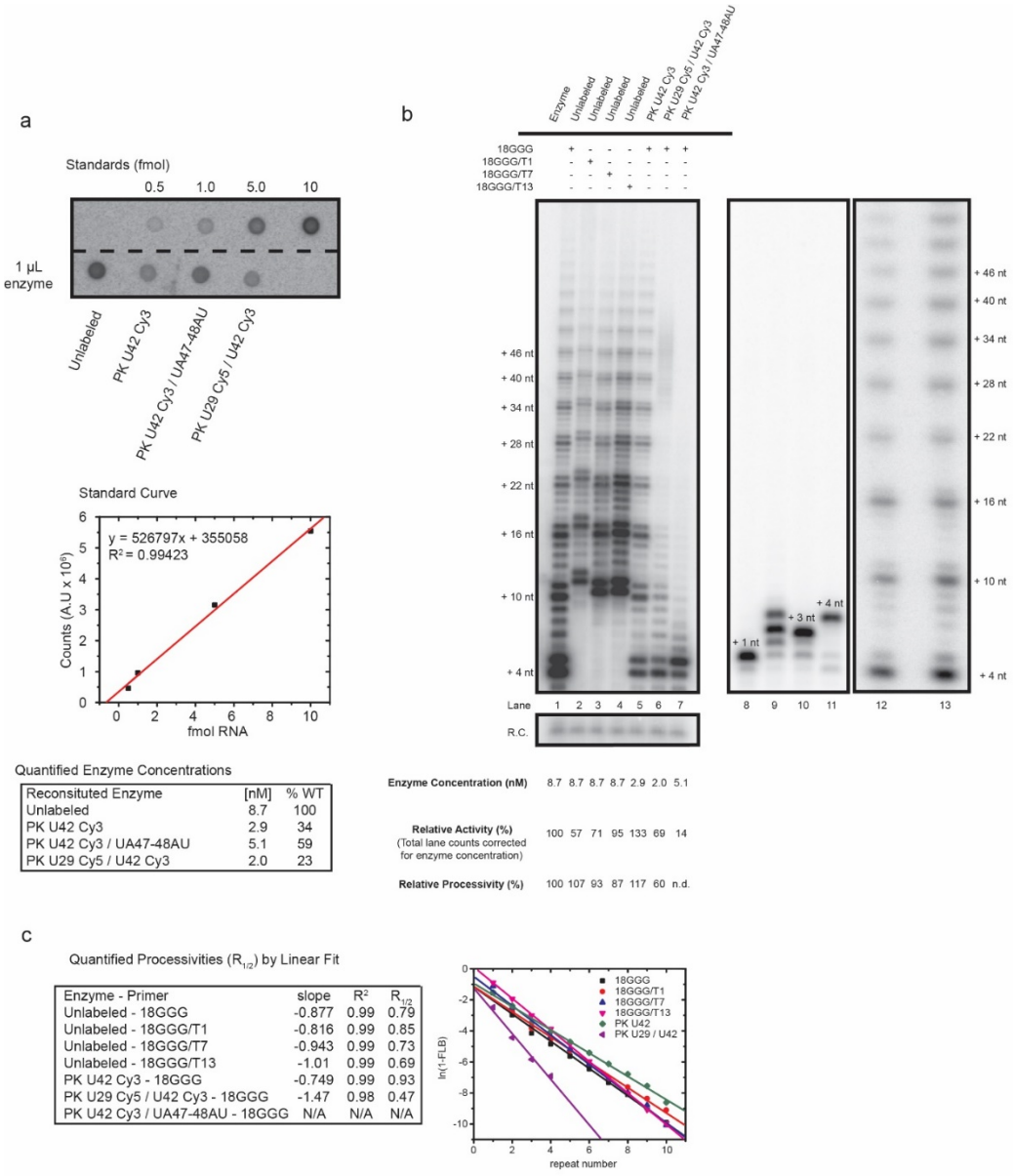
56. Das, R., Laederach, A., Pearlman, S. M., Herschlag, D. & Altman, R. B. SAFA: semi-automated footprinting analysis software for high-throughput quantification of nucleic acid footprinting experiments. *RNA* **11,** 344–54 (2005).

57. Roy, R., Hohng, S. & Ha, T. A practical guide to single-molecule FRET. *Nat. Methods* **5,** 507–16 (2008).

SUPPLEMENTARY INFORMATION

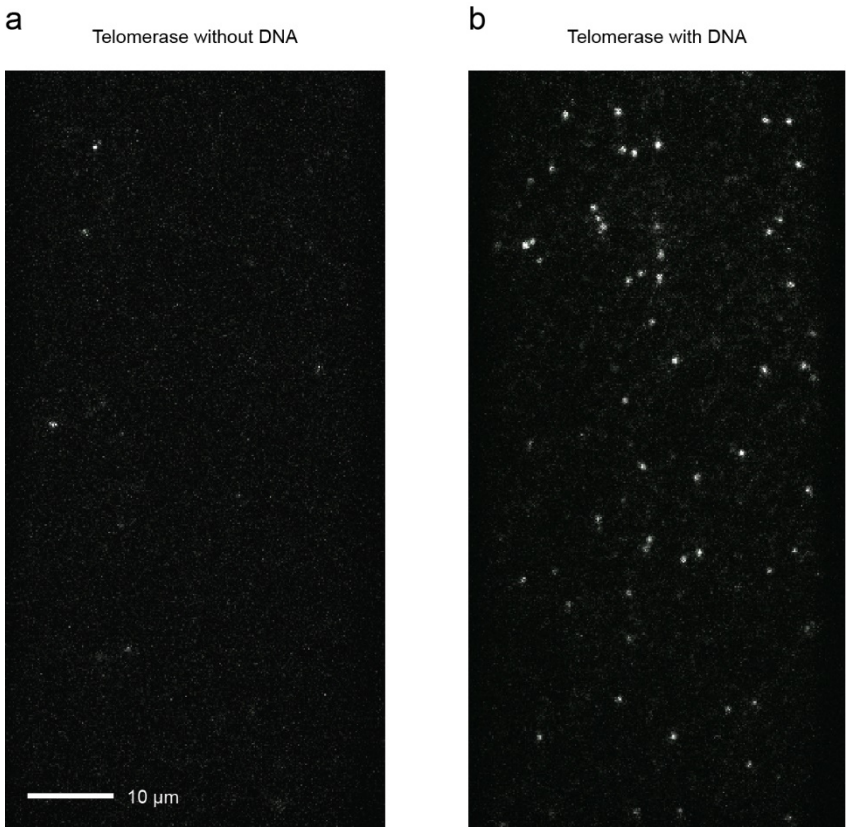


Supplementary Figure 1 | Telomerase RNA fragments used in single-molecule FRET experiments. A pseudoknot fragment (nts 32-195) labeled at position U42 with Cy3 (green circle) was constructed by a two piece DNA-splinted RNA ligation at the indicated site. Based upon previously reported work, a U32G mutation was made to enhance P1 stem stability[1]. A doubly-labeled pseudoknot fragment with an extended P1 stem harboring U42-Cy3 and U29-Cy5 labels was constructed by two sequential two piece DNA-splinted RNA ligation reactions as described in the methods. All telomerase reconstitution reactions included a separate CR4/CR5 (nts 239-328) fragment generated by *in vitro* transcription as described in the methods. The secondary structure models are modified versions of a previous representation of human telomerase RNA[2].

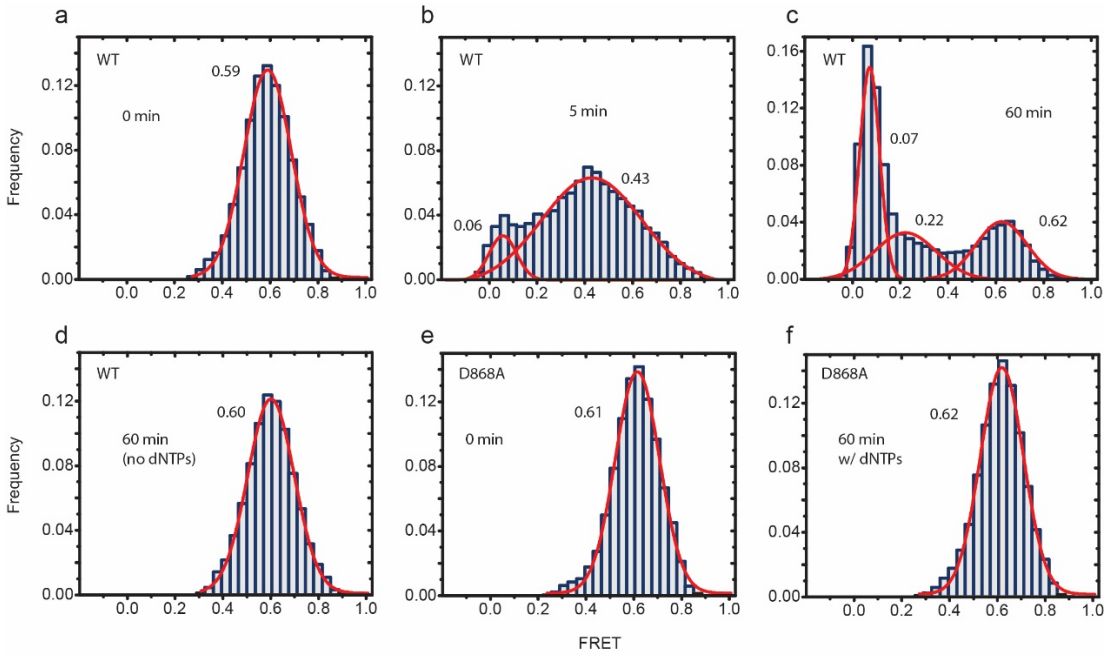


Supplementary Figure 2 | Quantification of telomerase enzyme reconstitution by RNA dot blot and direct DNA primer extension. a, (top) Scanned image of RNA dot blot quantification. (middle) Plot of internal standards from top panel quantified in ImageJ and fit by a linear regression. (bottom) Concentrations of telomerase extracted by standard curve. **b,** Direct DNA primer extension assays for all telomerase enzymes and DNA primers used in the study. Activity assays were performed as described in the detailed methods section unless otherwise noted. Concentrations of individual enzyme reconstitution reactions determined by RNA dot

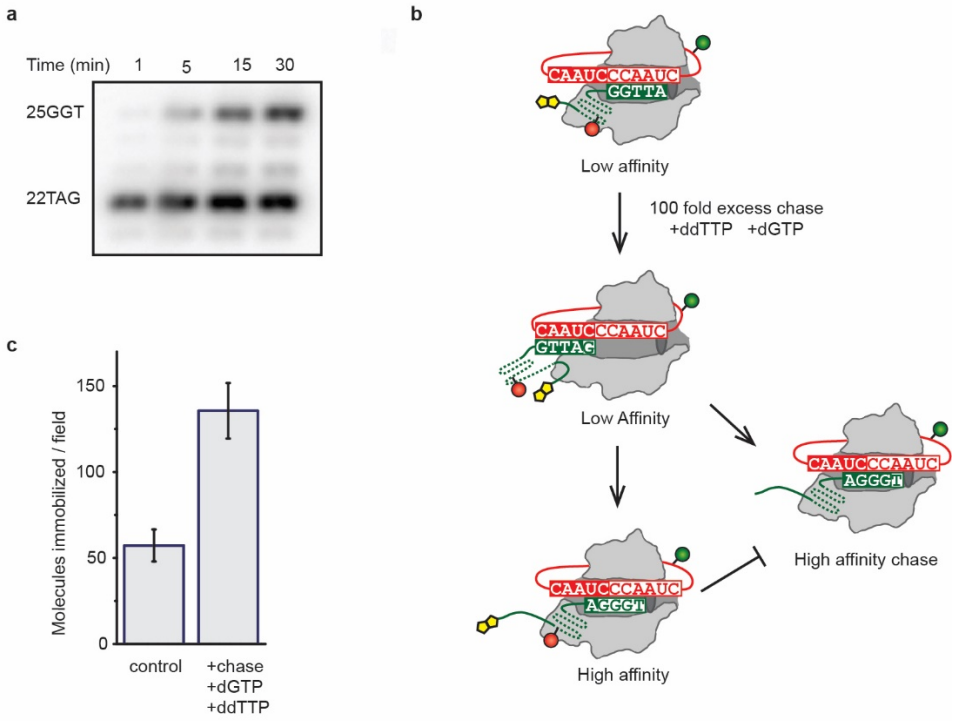
blot analysis were used to calculate total enzyme activity, defined as the total counts per lane corrected for the recovery control (RC), relative to the wild type (WT) enzyme. (**1**) Unlabeled telomerase extension of 18GGG DNA primer. (**2**) Unlabeled telomerase extension of 5'-biotinylated 18GGG/T1 DNA primer. The retardation of products in this lane is likely due to the Cy5 label site being in close proximity to the 5'-biotin moiety. (**3**) Unlabeled telomerase extension of 18GGG/T7 DNA primer. (**4**) Unlabeled telomerase extension of 18GGG/T13 DNA primer. (**5**) U42-Cy3 labeled telomerase extension of 18GGG DNA primer. (**6**) Doubly-labeled U42-Cy3 and U29-Cy5 telomerase extension of 18GGG DNA primer. (**7**) U42-Cy3 labeled telomerase with UA47-48AU template mutation extension of 18GGG DNA primer demonstrates knockdown of repeat addition processivity. (**Lanes 8-11**) Single nucleotide addition and chain termination experiments which were accomplished by extension for 30 minutes to match smFRET experiments described in Fig. 3 (main text). (**8**). Unlabeled telomerase extension of 17AGG primer in the presence of dATP, dGTP [$\checkmark^{32}$P], and ddTTP yielding the product 19GGT. (**9**). Unlabeled telomerase extension of 17AGG primer in the presence of dGTP [$\checkmark^{32}$P] and dTTP yielding the product 20GTT. This experimental condition did not induce a clean 20GTT product and was therefore excluded from single-molecule FRET experiments. (**10**). Unlabeled telomerase extension of 17AGG primer in the presence of dGTP [$\checkmark^{32}$P], dTTP, ddATP yielding the product 21TTA. (**11**). Unlabeled telomerase extension of 5'-biotinylated 17AGG primer. For this experiment an initial incubation in the presence of dGTP [$\checkmark^{32}$P] for 2 minutes was followed by the addition of dTTP, dATP, and ddGTP yielding the product 22TAG. (**12**) Unlabeled telomerase extension of 5'-biotinylated 18GGG in the presence of saturating triplet state quencher, TROLOX. (**13**) Unlabeled telomerase extension of 5'-biotinylated 18GGG in the absence of the triplet state quencher, TROLOX. **c,** After quantification of the bands using SAFA 11b, the processivities were calculated using a previously reported method[3]. In short, the fraction of DNA left bound (FLB) at each repeat band was determined. A linear fit of ln(1-FLB) versus repeat number was plotted and fit by linear regression using the expression $\ln(1\text{-FLB})=(0.693/R_{1/2})*(\text{repeat number})$. Dividing 0.693 by the slope of each line provides $R_{1/2}$, the processivity value, for the respective enzyme. All the processivity values were quite similar except for a slight knockdown in processivity for the doubly labeled telomerase and the expected knockdown in the UA47-48AU mutant enzyme.



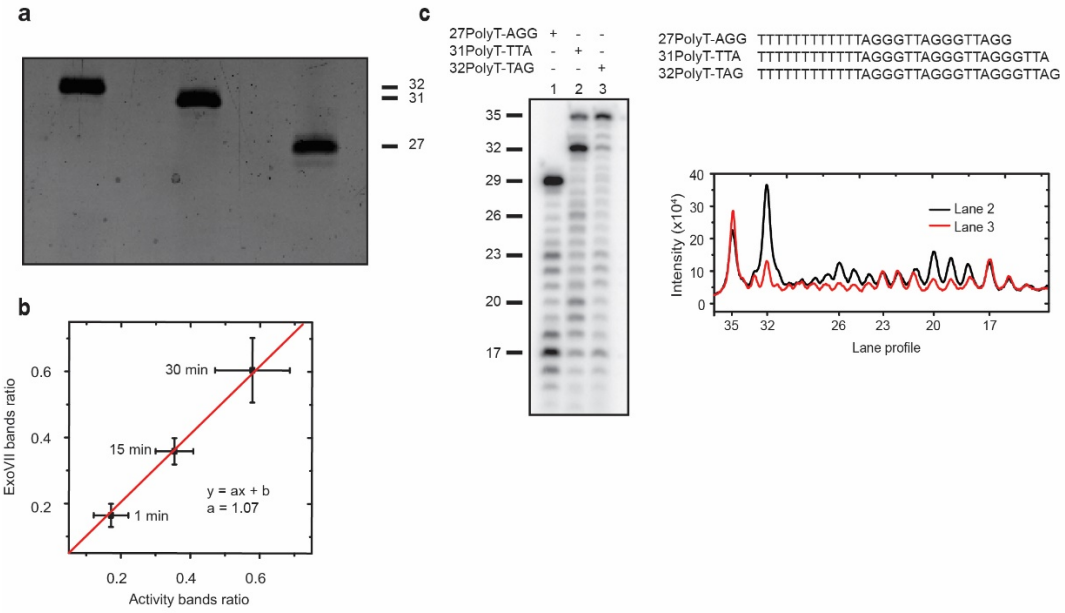
Supplementary Figure 3 | Telomerase immobilization is DNA-primer dependent. **a**, FRET-labeled telomerase was added to the sample chamber in the absence of 5'biotinylated DNA primer, showing only a background level of complexes imaged by a TIRF microscope. **b**, Pre-incubation of FRET-labeled telomerase enzyme with 5'biotinylated DNA primer increases immobilization efficiency via the biotin interaction with the streptavidin coated surface.



Supplementary Figure 4 | In situ telomerase activity alters the observed distribution of smFRET values. a, The U42-Cy3 labeled telomerase was bound to 18GGG/T13-Cy5, producing a unimodal FRET distribution (left most panel). Addition of dATP, dTTP, and dGTP, initiates telomerase DNA primer extension activity. After incubation times of 5 minutes (**b**) and 60 minutes (**c**) the FRET distributions trend toward progressively lower values. Notably, the very low FRET value in panel (**c**) is not due to photo-bleaching or enzyme dissociation since the telomerase enzyme is labeled with the Cy3 donor dye. Therefore, this non-zero FRET value implies the Cy5 dye on the primer is moving away from the RNA template region. **d**, If the experiment is performed in the absence of dNTPs no change in the FRET distribution is observed. **e**, Telomerase enzymes reconstituted with a catalytically dead mutant TERT (D868A) show a similar initial FRET distribution to the wild type enzyme centered at FRET = 0.61. **f**, As expected, a 60 minute incubation in the presence of dNTPs does not alter the FRET distribution for the D868A mutant, supporting the interpretation that a shift in the observed FRET distribution is due to telomerase catalytic activity on the slide. All histograms in the figure are constructed from at least 750 individual molecule



Supplementary Figure 5 | Single-round repeat addition processivity experiments. **a**, The translocation efficiency of our reconstituted telomerase is not quantitative as demonstrated by the ratio of the 22TAG (single repeat) vs. 25GGT (single round translocation). **b**, We therefore devised a method to specifically enrich for complexes that have successfully completed a single round of translocation for smFRET analysis. U42-Cy3 labeled telomerase was pre-bound to the low affinity 5' biotinylated 21TTA/L2-Cy5 DNA primer. Addition of dGTP and ddTTP initiates completion of the first telomere DNA repeat, translocation, and additional synthesis of the next repeat until chain termination by ddTTP incorporation. Note that only those complexes that have successfully translocated and extended to the high affinity 25GGT state will remain bound to the biotinylated primer in the presence of 100-fold excess of non-biotinylated high affinity 18GGG chase primer. In contrast, complexes that stall in the low affinity 22TAG state will be effectively competed by the non-biotinylated 18GGG chase primer, and will therefore not be efficiently immobilized on the microscope slide for smFRET analysis. **c**, This approach is validated by the observed ~2.5-fold increase in immobilization efficiency of complexes permitted to extend to the high-affinity 25GGT state compared to the low affinity 21TTA control. Bar plot shows average number of immobilized molecules per field of view in the microscope, and error bars represent the standard deviation across 20 independent fields of view.



Supplementary Figure 6 | 5'-DNA extrusion is coupled to translocation completion. a, The purity of synthetic primers 27PolyT-AGG, 31PolyT-TTA, and 32PolyT-TAG was tested by a 12% denaturing PAGE gel and ethidium bromide staining. The absence of impurities past the -1 band suggests the background non-specific products seen in Figure 6 Lanes 1 and 3 (main text), are due to a nuclease contaminant pulled down with telomerase. **b**, In Figure 5 (main text), we quantified the activity related bands (32 and 35 nt) and Exo related bands (17 and 20 nt) in Lanes 4-7, and compared the ratio of 35nt/32nt extension products to the ratio of 17nt/20nt ExoVII products. A correlation plot of these ratios as a function of the activity time course demonstrates a high level of correlation, strongly suggesting the 17 nt Exo VII cleavage product is derived from the 35 nt RAP product band. Error bars represent the standard deviation of the experiment performed in triplicate. **c**, To more directly verify the band at 17 nt was associated with the 35 nt translocated species, we used a primer that hybridizes to the full template repeat (32polyT-TAG). Extension of this primer in the presence of dGTP [α^{32}P] and ddTTP for 60 minutes dramatically enriches for the 35GGT RAP product band. When treated with Exo VII, we again observe two regions of protection centered about 17nt and 23nt (Lane 3). (Inset) Comparison of the ExoVII cleavage patterns from the 32 nt single repeat band (Lane 2, black line) and the 35 nt RAP product band (Lane 3, red line) de-convolutes the two sets of cleavage products that are superimposed in Figure 5c (main text).

Supplementary Table 1 | DNA and RNA oligonucleotides

Name	Sequence (Label site*)
18GGG/T13	5'-Bio-TTAGGGTTAGGGTTAGGG-3'
18GGG/T7	5'-Bio-TTAGGGTTAGGGTTAGGG-3'
18GGG/T1	5'-Bio-TTAGGGTTAGGGTTAGGG-3'
21TTA/T7	5'-Bio-TTAGGGTTAGGGTTAGGGTTA-3'
18GGG	5'-TTAGGGTTAGGGTTAGGG-3'
17AGG	5'-Bio-TTAGGGTTAGGGTTAGG-3'
U42 Labeled RNA	5'-GGGCCAUUUUUUGUCUAACCCUAACUGAGAA-3'
U42 Labeled UA47-48-AU Mutant RNA	5'-GGGCCAUUUUUUGUCAUACCCUAACUGAGAA-3'
U29 Labeled RNA	5'-GGGGUGAAC-3'
DNA Splint for 32-195	5'-CAGCGCGCGGGGAGCAAAAGCACGGCGCCTACGCCCTTCTCAGTTAGGGTTAGACAAAAAATGGCCACCACCCCTCCCAGG-3'
DNA Splint for 25-201	5'-CAGCGCGCGGGGAGCAAAAGCACGGCGCCTACGCCCTTCTCAGTTAGGGTTAGACAAAAAATGGCCCGTTCACCCC-3'
27PolyT-AGG	5'-TTTTTTTTTTTTAGGGTTAGGGTTAGG-3'
31PolyT-TTA	5'-TTTTTTTTTTTTAGGGTTAGGGTTAGGGTTA-3'
32PolyT-TAG	5'-TTTTTTTTTTTTAGGGTTAGGGTTAGGGTTAG-3'

hTR PCR Primer 32-195 Forward	5'-TAATACGACTCACTATAGGGCCATTTTTTGTCTAACCC-3'
hTR PCR Primer 32-195 Reverse	5'-AACGGGCCAGCAG-3'
hTR PCR Primer 63-201 Forward	5'-TAATACGACTCACTATAGAGAAGGGCGTAGGC-3'
hTR PCR Primer 63-201 Reverse	5'-GGGGCGAACGGG-3'
hTR PCR Primer 63-195 Forward	5'-TAATACGACTCACTATAGAGAAGGGCGTAGGC-3'
hTR PCR Primer 63-195 Reverse	5'-AACGGGCCAGCAG-3'
hTR PCR Primer 239-328 Forward	5'-TAATACGACTCACTATAGAACCCCGCCTGG-3'
hTR PCR Primer 239-328 Reverse	5'-GACCCGCGGCTG-3'
Dot Blot Probe Primer	5'-CGGTGGAAGGCGGCAGGCCGAGGC-3'

*Labeling sites in DNA or RNA oligonucleotides represent internal amino modifiers C6-dT (IDT) or 5-Amino-allyl-uridine (Dharmacon) respectively. These modified sites were then coupled to the appropriate FRET dye and purified by reverse phase HPLC.

Supplementary References

1. Chen, J. & Greider, C. W. Template boundary definition in mammalian telomerase. *Genes Dev.* **17,** 2747–52 (2003).

2. Podlevsky, J. D., Bley, C. J., Omana, R. V, Qi, X. & Chen, J. J.-L. The telomerase database. *Nucleic Acids Res.* **36,** D339–43 (2008).

3. Latrick, C. M. & Cech, T. R. POT1-TPP1 enhances telomerase processivity by slowing primer dissociation and aiding translocation. *EMBO J.* **29,** 924–33 (2010).

CHAPTER III

Single molecule FRET-Rosetta reveals conformational changes in the essential telomerase RNA pseudoknot domain during catalysis

*In review at eLife

ABSTRACT

Ribonucleoprotein (RNP) complexes mediate diverse and essential cellular processes. However, generating homogeneous and functional RNP samples often represents a significant barrier to developing structural and functional models. By example, the telomerase RNP is critical for rapidly dividing cells, including stem cells and the majority of human cancers. However, a complete understanding of telomerase mechanism necessitates improved structural models of the active RNP complex. The recent availability of several high-resolution structures has provided insight into telomerase subdomains, yet the core human telomerase architecture remains unknown. Here we integrate single molecule Forster Resonance Energy Transfer (smFRET) distance constraints and existing high-resolution structural data with Rosetta modeling to map the conformation of the essential telomerase RNA pseudoknot domain within the catalytically active RNP. Our modeling places the evolutionarily conserved RNA pseudoknot fold on a protein surface opposite the enzyme active site. In this model, the pseudoknot fold interfaces with the C-Terminal Element (CTE), a crucial protein domain that shares structural homology with the thumb domain of other polymerases. Interestingly, we find the telomerase pseudoknot

domain exhibits nanometer-scale rearrangements during enzyme catalysis. We present and critically test a model wherein RNA pseudoknot movement is physically coupled to template hybrid synthesis in the active site. Finally, the architecture of the human telomerase RNP established by smFRET-Rosetta is in excellent agreement with an independently solved medium-resolution cryoEM structure of the *Tetrahymena thermophia* telomerase RNP, providing a powerful blind test of our modeling methodology.

INTRODUCTION

Sparse biophysical and biochemical data can be integrated with Rosetta modeling as an effective strategy for generating nanometer scale pseudo-atomic models of complex biological systems. Models that emerge from these studies provide a conceptual framework for designing additional experiments to reiteratively refine structure-function relationships. Structural modeling is particularly important for systems refractory to traditional structure determination techniques, such as x-ray crystallography or NMR. Such is the case with many ribonucleoprotein (RNP) complexes, such as telomerase or components of the spliceosome, which are difficult to reconstitute and exhibit a high degree of structural dynamics [1,2].

Here, we have integrated single molecule Forster Resonance Energy Transfer (smFRET) with a Rosetta-based modeling approach to establish the architecture of the human telomerase core RNP. FRET is a biophysical structure-probing tool that monitors the distance-dependent energy transfer between a donor dye and an acceptor dye, strategically coupled to the structure of interest [3-5]. The dynamic range of FRET

is approximately 25-80 angstroms, a range well suited for most structural biology applications [6]. Moreover, FRET at the single molecule level permits dissection of heterogeneous and dynamic enzyme populations. To date, most FRET studies of the telomerase RNP have focused on gaining qualitative insight about structure and molecular dynamics [7-14]. In cases where FRET derived distances were employed, modeling uncertainty remained high due to the difficulty of converting FRET to precise distance values [15,16]. Much of this difficulty results from the uncertainty introduced by the unique dye environments in each FRET sample [17-19]. However, recent smFRET studies of the RNA polymerase II holoenzyme have demonstrated that these challenges can be overcome by several methods, including triangulation of dye positions through multiple FRET measurements and Bayesian parameter estimation for statistical modeling of dye location [17,20-23]. We expand on this approach by imposing compliant distance ranges for each FRET value rather than discrete distances, which devalues precise calibration of individual distances and emphasizes redundancy in the FRET distance network. This modeling decision abrogates the requirement for complex biophysical parameterization that is currently not feasible for telomerase due to its low reconstitution efficiency.

The telomerase catalytic core consists of the telomerase reverse transcriptase (TERT) protein subunit and an integral RNA component that carries the template for telomerase reverse transcription (hTR or hTER) [24,25]. Over the past decade, the structures of several conserved mammalian TR fragments have been reported, including two RNA motifs required for catalytic activity: CR4/5 and the pseudoknot

fold [26-28]. The CR4/5 domain is a conserved RNA element of unknown function that has recently been structurally characterized at high-resolution in the context of the telomerase RNA binding domain [29]. Similar to CR4/5, the functional role of the RNA pseudoknot fold remains poorly understood despite its importance for telomerase activity and disease states [30-32]. Although the structure of the native human pseudoknot fold is known, its position within the larger hTR pseudoknot domain and, more generally, the telomerase RNP remains unknown [10,27].

Telomerase differs from other reverse transcriptases in its ability to add multiple template repeats to its substrate during a single binding event, an activity termed repeat addition processivity (RAP) [24,25]. Although the precise molecular mechanism underlying RAP remains poorly understood, a series of protein, DNA, and RNA rearrangements are required. First, telomerase binds the substrate DNA through hybridization with the hTR template, reverse transcribes a single template repeat and halts at the template boundary [33-35]. Subsequent hybrid denaturation facilitates an enzymatic sub-step known as translocation, where the nascent DNA realigns with the downstream template region and rebinds in the active site for the next round of DNA synthesis (Fig. S1) [9,36-38]. The smFRET network established in the present study is uniquely capable of dissecting dynamic and heterogeneous structural rearrangements during the telomerase catalytic cycle.

In the present work, we report 3D models of the hTR pseudoknot domain architecture in the assembled RNP complex, along with evidence for a dynamic motion of the pseudoknot domain during telomerase catalysis. These models were

generated using known telomerase RNA and protein-structures and a novel Rosetta energy function that integrates loose distance constraints derived from single molecule FRET measurements. Surprisingly, the modeling reveals a convergent solution to pseudoknot domain architecture where the catalytically indispensable pseudoknot fold contacts the protein surface opposite the template hybrid and active site. Upon stimulation of telomerase activity, the pseudoknot exhibits a conformational rearrangement, suggesting that pseudoknot motion is physically coupled to synthesis at the active site. In order to critically test this model, we generated a circularly permuted hTR pseudoknot domain (cp-PK) that breaks the RNA tether connecting the pseudoknot fold to the template hybrid at the active site. The cp-PK construct supports telomerase catalysis in ensemble activity assays with a severe knockdown in enzyme processivity, reiterating the functional significance of PK connectivity for telomerase processivity [39]. When observed at the single molecule level, the cp-PK decouples the pseudoknot conformational change from catalytic activity, supporting the hypothesis that pseudoknot motion is coupled to template hybrid synthesis during processive telomerase activity.

METHODS

Telomerase reconstitution and activity

In-vitro RNA preparation

In vitro reconstitution of telomerase enzyme was performed using the established two RNA fragment system in which the essential RNA pseudoknot domain and the CR4/CR5 domain are added *in trans*. Unlabeled CR4/CR5 RNA (hTR 239-328) was

generated by *in vitro* transcription using T7 RNAP and polyacrylamide gel electrophoresis (PAGE) purified. Site-specific RNA labeling within the RNA pseudoknot and CR4/5 domains was achieved using DNA-splinted RNA ligation techniques as described previously [9]. Unlabeled fragments of the RNA pseudoknot domain (hTR 63–195, hTR 120-195, and cphTR 120-195+32-62) were prepared by *in vitro* transcription using T7 RNAP and purified using a 6% denaturing PAGE gel. Subsequently, the RNA was phosphatase (CIP, NEB) treated to remove the 5′ triphosphate, phenol/chloroform extracted to remove CIP, T4 PNK (NEB) treated to add a 5′ monophosphate, phenol/chloroform extracted to remove PNK and PAGE purified to remove excess nucleotide. Both enzymatic reactions were carried out using NEB specifications.

Dye labeling of synthetic RNA

Synthetic RNA fragments were purchased from Dharmacon containing internal aminoallyl uridines used for dye labeling reactions. Specific RNA sequences are listed in Table S2. RNAs were generally processed as previously described [10]. RNAs from Dharmacon were directly dissolved in 50 μL 0.1 M $NaHCO_3$ and then used to solubilize a single mono-reactive dye pack (GE). The labeling reaction was held at 37 °C for two hours. The RNA was precipitated and deprotected according to manufacturer protocol (Dharmacon). RNA was again precipitated and then dissolved in 60 μL of 0.1 M TEAA pH 7.5. An analytical C-8 reverse column was used to purify the labeled from unlabeled RNA fractions as previously described [40]. After

purification the RNAs were precipitated, pooled, and stored at -70 °C until use in ligation.

Ligation of RNA fragments

Modifications of the previously reported procedure included addition of 1 U μL^{-1} RNasin and carrying out the ligation overnight at 30 °C [10,40]. The DNA splints for each RNA ligation are listed in Table S2. Note for the doubly labeled pseudoknot domain RNA constructs, the 32-62 fragment was first ligated to the 63-120 fragment, denaturing PAGE purified, then ligated to the 120-195 fragment, and finally purified again by denaturing PAGE gel. During the PK ligations, when the required RNA fragment did not contain a label, the unlabeled fraction of the synthetic oligo was used.

Telomerase reconstitution and purification

Telomerase reconstitution was performed using the TnT Quick Coupled Transcription/Translation system (Promega) [10]. For every 100 μL of TnT Quick mix, 2 μg of pNFLAG-hTERT was added. RNAs were then added to the reaction mixture where dye-labeled RNAs were added to a final concentration of 0.05 μM, whereas all unlabeled RNAs were added to 1 μM. As an exception, the circular permutant dye labeled RNA was reconstituted at 1 μM. The reconstitution mixture was held at 30 °C for 3 h.

Directly after reconstitution, telomerase enzymes were purified using an N-terminal FLAG tag on hTERT. To pull down the enzyme, Sigma Anti-FLAG M2-agarose bead slurry was used at 50 μL bead slurry per 200 μL lysate. The beads were

first washed four times with wash buffer 1 (50 mM Tris HCl pH 8.3, 3 mM $MgCl_2$, 2 mM DTT, 1 mM spermidine, 100 mM KCl), spinning at 2350, r.c.f., 4 °C, for 1 min between washes. The beads were blocked while being agitated in 750 µL blocking buffer (50 mM Tris HCl pH 8.3, 3 mM $MgCl_2$, 2 mM DTT, 1 mM spermidine, 0.5 mg ml^{-1} BSA, 0.05 mg ml^{-1} glycogen, 0.1 mg ml^{-1} yeast tRNA) for 15 min at 4 °C. After spinning the beads at 2,350 r.c.f., 4 °C, for 1 min the supernatant was discarded. The blocking step was repeated (without discarding supernatant) and then the beads were added to the lysate. The bead-lysate mixtures were agitated at 4 °C for 2 h allowing the prepared enzyme to associate with the anti-FLAG beads. The mixture was spun at 2,350 r.c.f., 4 °C, for 1 min and the supernatant was flash frozen and stored for analytical purposes. The beads were then washed with 750 µL of wash buffer 2 (50 mM Tris HCl pH 8.3, 3 mM $MgCl_2$, 2 mM DTT, 1 mM spermidine, 300 mM KCl) and spun at 2,350 r.c.f., 4 °C, for 1 min. This process was repeated three times to remove residual lysate. The beads were washed three more times with 750 µL of wash buffer 1. To elute the enzyme, 60 µL of elution buffer (50 mM Tris HCl pH 8.3, 3 mM $MgCl_2$, 1 mM spermidine, 0.5 mg ml^{-1} BSA, 20% glycerol, 0.75 mg ml^{-1} 3 × FLAG peptide) was added for every 200 µL of lysate and the slurry was incubated at 4 °C for one hour. Beads were removed by centrifugation filtration using a Nanosep MF 0.45 µm filter. From this solution, 10-µL aliquots were prepared, flash frozen in liquid nitrogen and stored at −70 °C until use.

Telomerase activity assay

Activity assays were performed using 5 μL of purified enzyme diluted to a final volume of 10 μL in 1 × activity buffer. Each reaction maintained a final concentration of 1 μM primer substrate, 250 μM dATP, 250 μM dTTP, 1 μM α^{32}P-dGTP (PerkinElmer BLU514Z500UC), and 1 × activity buffer (50 mM Tris HCl pH 8.3, 3 mM $MgCl_2$, 2 mM DTT, 1 mM spermidine). The reactions were initiated by addition of dNTPs and incubated at 30 °C for 90 min. Reactions were terminated by addition of 190 μL stopping buffer containing the loading control (10 mM Tris HCl pH 7.5, 1 mM EDTA, 0.1% w/v SDS). The reactions were mixed vigorously with 100 μL of phenol to remove protein components followed by a 10 min spin at 15,000 r.c.f. The aqueous layer was further extracted with 100 μL of chloroform and spun at 15,000 r.c.f. for 2 min. After isolation of the aqueous phase, the radiolabeled DNA was precipitated and separated on a 12% denaturing PAGE gel (19:1 crosslinking ratio). The gel was dried, exposed on a phosphorimager screen overnight and scanned using a Typhoon scanner. Quantification of the bands was accomplished by using SAFA v11b [41].

Single molecule FRET experiments

Slide preparation

Quartz slides (Finkenbeiner Inc.) were cleaned and pegylated using the following protocol [3]. The slides were boiled in water to remove parafilm and coverslips from previous experiments. The slides were then scrubbed with alconox (Alconox Inc.), rinsed and placed into a solution containing 10% alconox. After sonication for 20 min, the slides were rinsed and placed in water and sonicated for 5 min. Slides

were then sonicated in acetone for 15 min. The slides were directly transferred into 1 M KOH and sonicated for 20 min. After thorough rinsing, the slides surfaces were flame dried and cleaned using a butane torch (BernzOmatic). While the slides were cooling, a silanizing solution containing 100 ml of methanol, 5 ml of glacial acetic acid and 1 ml of N-(2-aminoethyl)-3-aminopropyltrimethoxysilane (UCT) was prepared. The cooled slides were placed in the solution, sonicated for 1 min and allowed to stand in solution at room temperature for at least 20 min. While incubating slides, 200 mg of mPEG-Succinimidyl Valerate MW 5,000 (Laysan Bio, Inc.), was resuspended in 400 μL of 0.1 M $NaHCO_3$. Also 2 mg of Biotin-PEG-Succinimidyl Valerate MW 5,000 (Laysan Bio, Inc.) was resuspended in 200 μL of 0.1 M $NaHCO_3$. After rinsing the slides with dH_2O and drying with nitrogen, the solutions were mixed and applied to each slide surface (150 μL per slide). A coverslip was used to cover the solution and incubated overnight at room temperature in a humidor box. The following day, channels were assembled using Paramfilm as a spacer on the pegylated quartz slides and plasma-cleaned coverslips were used as the second chamber face.

Single molecule experimental procedure

To prepare the slides for molecule deposition, the channels were incubated with 60 μL of 10 mg ml^{-1} BSA (NEB) for 20 min. The slide was then washed with 100 μL of T50 (10 mM Tris HCl pH8, 50 mM NaCl) and then incubated with 60 μL of 0.2 mg ml^{-1} streptavidin in T50 for 5 min. After coating the slide with streptavidin, the excess protein was washed away with 200 μL of T50. The slide surface was

coated with immobilization primer by washing 100 nM primer over the surface in T50. Excess primer was then rinsed away with 200 μL of T50. A total of 10 μL telomerase solution from the immuno-precipitation was diluted with 20 μL of imaging buffer (50 mM Tris HCl pH 8.3, 3 mM $MgCl_2$, 1 mM spermidine, 0.5 mg ml^{-1} BSA, 4% glucose) and flowed over the slide surface. For labeled RNAs, each pseudoknot construct was diluted to 1 nM and applied to the slide surface. The number of molecules on the surface was observed in real time to reach the desired density for smFRET measurements. Once telomerase molecules/RNAs were deposited on the surface, the stable complexes were retained and imaged after multiple buffer exchanges. Imaging of the FRET-labeled enzymes required an oxygen-scavenging system and a triplet state quencher to increase dye lifetime. Imaging buffer does not substantially impact telomerase activity [7,9]. The buffer was saturated with Trolox (triplet state quencher), passed through a 0.22-μm filtration unit and brought to pH 8.3. Once ready for imaging, the solution was brought to 2 μg ml^{-1} catalase and 1 mg ml^{-1} glucose oxidase. The pH was found to be stable for more than 30 min with these conditions (data not shown). In the case of telomerase FRET activity, dATP, dTTP, and dGTP were added as described in the text at a concentration of 200 μM in imaging buffer.

Microscope setup

Data were collected using a custom built prism-type microscope, allowing hundreds of molecules to be studied in parallel [40]. Cy3 donor dyes were directly excited with a 532 nm laser (Laserglow). Depending on the proximity of Cy5 acceptor dye to the

donor dye, a variable degree of energy transfer will occur from the excited donor dye to the acceptor dye. Due to this FRET process, both dyes are capable of fluorescing and emit at distinct wavelengths. Emitted fluorescence was split into separate donor and acceptor channels using dichroic mirrors, and imaged on an EM-CCD camera (Andor Ixon).

Data acquisition and analysis

Imaging fields containing 30–250 molecules were imaged using a 100 ms integration time for fast timescale traces and histograms and 1000 ms integration time for slow timescale traces. Individual traces were parsed out using custom written IDL software where they were corrected for background and dye-crosstalk signal (software available on request). The individual traces were then filtered in MATLAB manually where only traces that showing FRET, acceptor bleaching and then donor bleaching pathway were accepted. This specific photo-bleaching process was required for correction of the gamma factor at the single molecule level. FRET intensities were then calculated using the equation $(I_A - \beta I_D)/ (I_A + \gamma I_D)$ where I_A is the acceptor intensity, I_D is the donor intensity and γ is the gamma correction factor and β accounts for the leakage of the donor fluorescence into the acceptor channel. The gamma correction factor was determined as described for individual molecules [42]. The first five seconds of the corrected, individual FRET traces were then binned into FRET histograms. The center of the FRET distributions were determined by a non-linear Gaussian fitting algorithm.

Determining distance from FRET

FRET peak centers were converted from FRET efficiency to distance using the Equation 1. In this equation, FRET represents the observed peak center of the corrected smFRET data and R_0 is the tabulated Förster radius (5.6 nm) specified by the dye manufacturer (Amersham Biosciences). Typically R_0 has been calculated using a number of tabulated parameters and an experimentally determined quantum yield of the donor [17,20,22,23]. In our experiments, it was not possible to experimentally determine the quantum yield of the donor in the assembled complexes due to low sample concentrations. Therefore we accounted for errors in R_0 through compliance in the FRET scoring parameter in the modeling function.

Rosetta modeling

Generation of modeling scaffold

An initial modeling scaffold was generated using known sub-structures of human telomerase. The previously published homology model of human TERT was used with an updated model of the RBD, to reflect information from the recently solved structure of the RBD from the vertebrate Takifugu rubripes [43]. The new homology model of the human RBD was built using HHPRED and Modeller [44,45] using 4LMO, 2R4G, and 3KYL as templates [43,46,47]. This was integrated back into the original homology model by aligning corresponding residues. The template hybrid was modeled as an ideal A-form helix, and because interactions between the template hybrid and the protein were not modeled explicitly (see below), and the template hybrid was fixed relative to the protein. The DNA in the template hybrid was modeled as RNA for simplicity. The template hybrid was positioned in the human

homology model by aligning the end of the template hybrid (5' end of the template and 3' end of the primer) to the template hybrid in the active site in the homology model, which was built using information from the *T. castaneum* crystal structure of the RBD, RT, and CTE domains bound to the DNA/RNA template hybrid (PDB ID 3KYL) [47]. The structure of the P6.1 helix, residues 301-305 and 311-315, was taken from the crystal structure of the *O. latipes* RBD bound to CR4/5 RNA, PDB ID 4O26 [29]. To position it relative to the human TERT model, the O. latipes RBD and CR4/5 structure was aligned to the homology model of the human RBD, and the resulting coordinates of residues 199-203 and 209-213 from the O. latipes CR4/5, which correspond to the human P6.1 stem, were used. For models built with an enforced P1 stem, the recently solved structure of the Tetrahymena RBD bound to the template boundary element (TBE), was aligned to the human RBD model, then the P1b helix was aligned to the TBE helix [48]. The final scaffold contained the TERT homology model bound to the template hybrid, the P6.1 stem, and the P1b helix.

In addition to the modeling scaffold, other known RNA structures were treated as rigid bodies that were free to move relative to the modeling scaffold. The solution structure of human P2ab, PDB ID 2L3E, was used for residues 78-93 and 121-131, and the pseudoknot, PDB ID 2KID, for residues 96-118 and 170-183. P2a.1 (residues 64-66, 143-145, 69-72, 136-139) was modeled as an ideal A-form helix.

General algorithm and scoring

Models of the telomerase RNA were generated using a modified version of fragment assembly of RNA (FARFAR), a method for building de novo structures of RNA by

sampling known fragments of RNA structures from the protein data bank in a Monte Carlo simulation. Structures are scored with a statistically derived low-resolution potential meant to approximate the free energy of the RNA ($Score_{RNA}$). Here, we modified this method to include smFRET data and protein-RNA sterics in two additional score terms, $Score_{FRET}$ and $Score_{clash}$, respectively. Now, the total score of an RNA structure is given by the sum of the original $Score_{RNA}$ with $Score_{FRET}$ and $Score_{clash}$. $Score_{FRET}$ takes the form of a smoothed square well potential. First, the peak measured FRET efficiencies +/- approximately one standard deviation were converted into distances, as described above. Then, for each pair of residues for which there is smFRET data, a penalty of zero is assigned if they are within a distance range of each other that corresponds to the center of the FRET distribution plus or minus approximately one standard deviation. The potential is then smoothed up to a penalty of 200 Rosetta units over 5Å above and below this distance range (Fig. 3). When there were multiple peaks in the FRET distributions, each peak was converted into a separate score term. Each of these terms was calculated for a given model, and the lowest was applied. In this way, if a pair of residues agreed with one of the FRET peaks, it would not get a penalty if it failed to agree with any other FRET peak for the same pair of residues. $Score_{clash}$ was introduced to account for steric clashes between the RNA and TERT protein, because the original fragment assembly method does not model interactions between RNA and protein residues. For computational efficiency, the coordinates of the atoms of the protein structure, taking into account their van der Waals radii, were placed into a three dimensional grid with

bin spacing set at 0.1 Å. Here, we accounted only for clashes with the backbone atoms of the protein structure because the TERT structure was a homology model and we do not allow for any conformational changes of the protein. Each time the RNA structure was scored, it was aligned to this grid using the fixed RNA components of the initial modeling scaffold, then $Score_{clash}$ was computed by counting the number of clashes and multiplying by a scaling factor, here 0.3. In this way, conformations of the RNA that had many steric clashes with the protein were penalized. This algorithm is available as part of the Rosetta software package.

Analysis

For each of the modeling runs, approximately 2500 structures were generated, 10,130 in total. Convergence of the runs was assessed by clustering the 400 best scoring RNA models, after alignment of all the models to the modeling scaffold, with a range of different radii. The smallest radius that placed at least 25% of these models in the largest cluster estimates the extent of convergence and provides an estimate of the resolution of the models.

RESULTS

Mapping telomerase RNA structure and dynamics with a single molecule FRET network

To investigate the relative positioning of essential RNA structural elements within the assembled telomerase RNP complex we devised a network of smFRET pairs including five distinct labeling sites within the hTR pseudoknot and CR4/5 domains (Fig. 1A). Four label sites in the pseudoknot domain were strategically

chosen to analyze the pseudoknot domain, as well as the relative orientation of the pseudoknot fold and template region. Knowledge of the high-resolution position of the CR4/5 domain label site (U312) relative to the telomerase protein subunit provided an important reference point for smFRET measurements to all other label sites in the pseudoknot domain [29,49]. For all measurements used in our study, the pseudoknot and CR4/5 domains were prepared as two separate RNA fragments and reconstituted *in trans* into active telomerase RNP complexes (Fig. S2). This approach simplifies preparation of the dye labeled RNAs and has previously been shown to support reconstitution of catalytically active telomerase (Fig. 1)[9] . We prepared telomerase enzymes harboring one of ten unique smFRET pairs, across the two RNA domains, using established telomerase reconstitution and immuno-purification techniques [50,51]. Each labeled telomerase enzyme was then surface-immobilized for smFRET analysis via binding to a 5'-terminal biotinylated DNA primer with sequence $(TG)_6$TTAGGG (Fig. 1B). This DNA primer possesses native telomere DNA sequence exclusively at the 3' terminus, ensuring that all complexes are bound through the same seven nucleotide DNA/RNA hybrid in the telomerase RNP [7,8,13]. We note that the repeating (TG) dinucleotide sequence present in this telomere DNA primer does not alter the telomerase reaction product profile when assayed *in vitro* (Fig. S2A lanes 1 and 2). Moreover, while unassembled smFRET RNA can, in principle, form the same seven base pair hybrid with the modified primer, control experiments revealed that RNA alone did not form a stable hybrid with the primer,

demonstrating the specificity of the primer-telomerase RNP interactions in our experiments (Fig. S3).

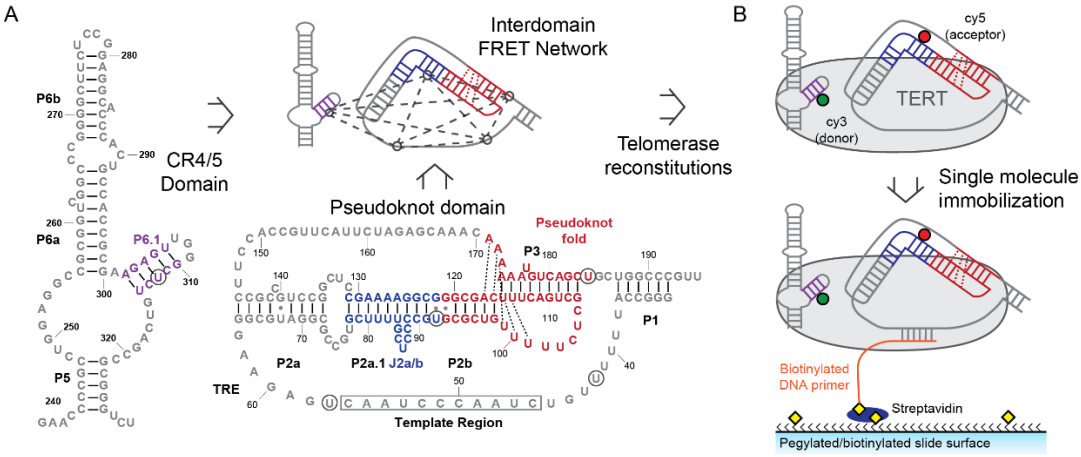


Figure 1. Single molecule FRET network in the telomerase RNP. (A) Secondary structures of pseudoknot and CR4/5 fragments used in telomerase reconstitution. Five separate label sites (denoted by black circles) span the two domains and create a network of FRET pairs. (B) Dye labeled RNA fragments are reconstituted with TERT to generate catalytically active RNPs. Each complex is surface immobilized through a 5' biotinylated DNA primer $(TG)_6TTAGGG$ that hybridizes to the template region of telomerase RNA.

Single-molecule FRET measurements were made using a prism-type total internal reflection fluorescence (TIRF) microscope, and each experiment included data collected from several hundred telomerase complexes. The immobilized telomerase complexes revealed predominantly unimodal smFRET distributions, consistent with a stably assembled and homogeneous RNA structure within the RNP. All smFRET values fell within the sensitive range of the FRET response (FRET = 0.22-0.76) (Fig. 2A-J, Table S1). FRET measurements between U92-U184 yielded a similar distribution to previous work [10], suggesting the pseudoknot triplex is also formed in the assembled RNP (Fig 2H, red). FRET measurements between the

template and CR4/5 unambiguously place the U312 position, within the P6.1 stem, closer to the 3' template region than the 5' template region (Fig 2D vs G, red). FRET across the template region is consistent with hybrid formation, further supporting the notion that the short, relatively weak RNA:DNA hybrid is stabilized in the active site of the telomerase complex (Fig 2A, red). These FRET results are generally consistent with the distances previously proposed in a model of the protein-free hTR pseudoknot domain (Fig. S4). Our smFRET data therefore supports the hypothesis of a triangular domain architecture of pseudoknot RNA [52], which is preserved within the assembled telomerase RNP. This result also corroborates recent chemical mapping data that demonstrates telomerase assembly minimally effects hTR secondary structure [53].

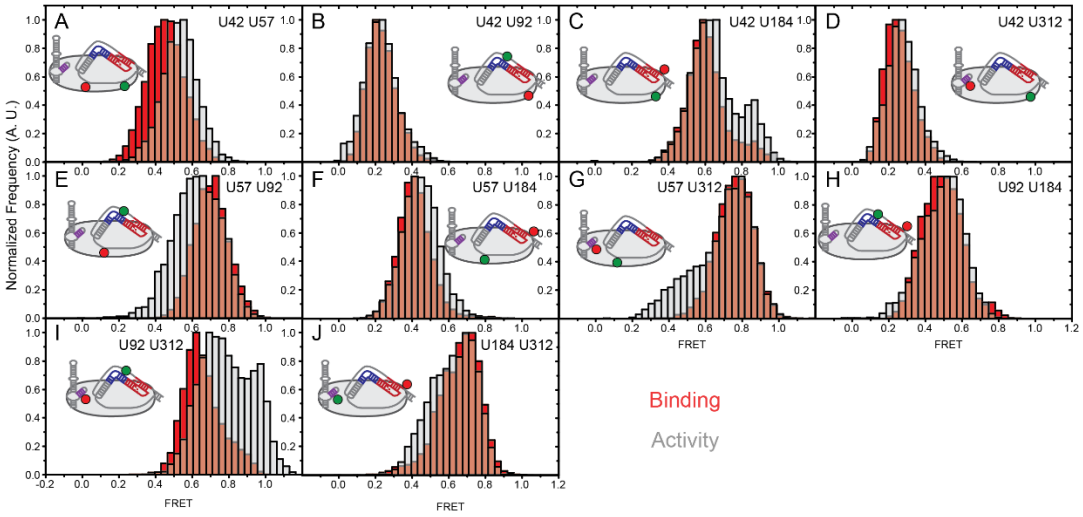


Figure 2. **Telomerase activity modulates RNA conformation inducing multiple FRET changes.** (A-J) Each panel represents reconstructed FRET histograms generated by compiling single molecule traces. The FRET label sites are listed in plain text in the panel and depicted in the cartoon version of telomerase. Red distributions were obtained by binding individual telomerase complexes to 5'-Bio-$(TG)_6$TTAGGG-3', capturing time resolved traces and then binning the corrected data into a histogram. Grey histograms represent exposure of the telomerase complexes to telomerase activity buffer, followed by similar data collection and analysis. (A) U42-

U57 FRET pair, (B) U42-U92 FRET pair, (C) U42-U184 FRET pair, (D) U42-U312 FRET pair, (E) U57-U92 FRET pair, (F) U57-U312 FRET pair, (G) U57-U312 FRET pair, (H) U92-U184 FRET pair, (I) U92-U312 FRET pair and (J) U184-U312 FRET pair. All histograms were generated from at least 50 molecules.

To specifically test whether the pseudoknot architecture is modified during telomerase RAP activity, we initiated telomerase catalysis by introducing activity buffer containing dGTP, dTTP, and dATP to the surface-immobilized complexes. Previous smFRET studies employing this same approach revealed the majority (>90%) of the telomerase complexes are catalytically active under our experimental conditions [9]. Addition of the telomerase activity buffer specifically promotes rapid extension of the DNA substrate to the template boundary, followed by the relatively slow translocation sub-step [54]. Therefore, our smFRET experiments enrich for telomerase complexes paused at a pre-translocation state, capturing the conformational rearrangements that are critical for this rate-limiting step of the telomerase catalytic cycle (Fig. S1).

Addition of telomerase activity buffer resulted in significant changes in five of the ten smFRET distributions, suggesting a conformational rearrangement of the RNA pseudoknot domain during telomerase catalysis (Fig 2, compare red and grey histograms). The shift in the distribution measuring template hybrid conformation indicates the template RNA undergoes a conformational rearrangement during telomere repeat synthesis (Fig 2A, red vs grey). FRET dyes probing the structure of the template recognition element (TRE), a crucial RNA motif located 3' of the RNA template, revealed shifted distributions supporting the notion of TRE dynamics during telomerase catalysis (Fig 2E and G, red vs. grey) [8]. Measurements at the 5' end

of the template (Fig 2C, red vs grey) revealed a new peak at ~0.85 FRET suggesting that the P1 stem may exist in multiple distinct conformations during the repeat addition cycle [50,55]. Surprisingly, the most pronounced FRET change was observed in the distribution of the U92-U312 smFRET pair, which probed the relative positions of the CR4/5 domain and the RNA pseudoknot fold (Fig 2I, red vs grey). Interestingly, the CR4/5 domain did not show activity-dependent structural changes with relation to the U42 position, which was shown to be static during telomerase catalysis (Fig 2D, red vs grey) [9]. The notion of a static CR4/5 domain is also consistent with the extensive contacts made between the CR4/5 RNA and the RNA binding domain of TERT [29,49,56]. Thus, the prominent shift in the U92-U312 FRET distribution likely represents movement of the RNA pseudoknot fold during catalysis. Finally, we did not detect any significant FRET change across the RNA pseudoknot, confirming that the human pseudoknot is stably folded throughout the catalytic cycle (Fig 2H, red vs grey) [10]. These qualitative interpretations of the smFRET data provide useful information about RNA domain structure and dynamics within the catalytically active telomerase complex. However, we sought a more detailed perspective of the RNP structure and dynamics by using automated molecular modeling in conjunction with the experimentally derived structural constraints provided by our smFRET approach.

Modeling telomerase RNP organization with smFRET-Rosetta

Prior to modeling the hTR pseudoknot domain within the telomerase RNP, we first combined known RNA and protein sub-structures of telomerase to create a modeling scaffold. The smFRET results were consistent with a homogenous domain

architecture and folded hTR pseudoknot; therefore, we integrated known structures of hTR pseudoknot domain motifs into our scaffold. Solution structures of the J2a/b and pseudoknot fold were treated as rigid bodies during modeling (Fig 3A). The template-primer hybrid was modeled as an ideal A-form helix, based on our smFRET binding data and the crystal structure of the *T. castaneum* TERT bound to a model DNA/RNA hybrid [47,57]. The P2a.1 region was also modeled as A-form helix based on previous chemical mapping experiments [58]. Lastly, while smFRET data conclusively demonstrated stable P1 stem formation in RNA alone, data taken within the assembled RNP could be consistent with either formation or remodeling of the P1 helix (Fig. S5). To account for the uncertainty of P1 formation during RNP modeling, we performed modeling runs both with and without an enforced P1 stem (see below).

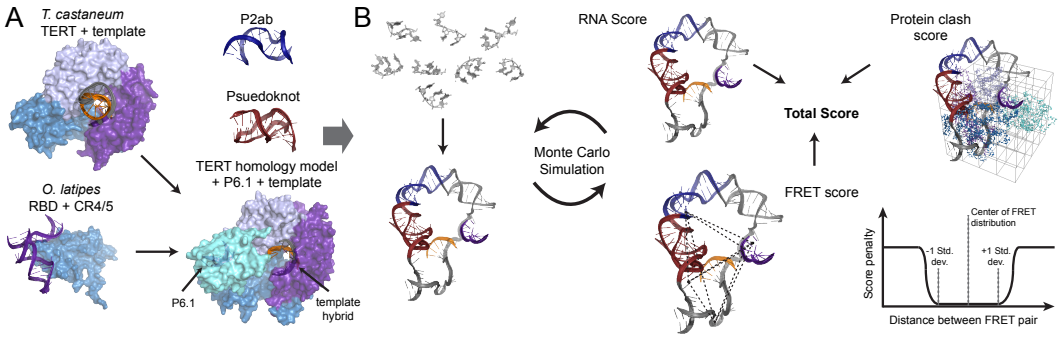


Figure 3. **Single molecule FRET-Rosetta modeling workflow.** (A) Structures of TERT bound to a template hybrid (PDB:3KYL, top left) and the RBD domain bound to the CR4/5 (PDB:4026, bottom left) allowed positioning of the template hybrid in the active site and the CR4/5 on the RBD of our updated homology model (bottom right). Solution structures of the pseudoknot fold (PDB:2K95) and the P2ab bulge (PDB:2L3E) were used as static RNA structures during the modeling process. Pseudoknot domain RNA was modeled around the protein homology model using these steric and structural constraints (B) RNA conformation was modeled by a previously defined algorithm that substitutes individual three nucleotide fragments of varying structure into the unknown regions of the RNA (see methods). The scoring function of the algorithm was modified to include a protein clash term which prevents RNA:protein steric clashing (top right) and a FRET term which enforces the distance

constraints generated by smFRET (bottom right). The FRET scoring term is defined by a smooth square potential that penalizes unsatisfied distance constraints (see methods).

Next we assembled a static structure of the telomerase reverse transcriptase protein component (hTERT). A homology model of hTERT was built by modifying the published homology model with an updated homology model of the RBD (Fig. 3A) [59]. This protein structure was used for all RNP models in the present study (Fig 3A). To complete our modeling scaffold, we used the three published RNP structures of telomerase sub-complexes to fix known RNA/protein interaction sites (Fig 3A). The template hybrid was positioned by aligning the 3' end of the DNA in the protein active site based on the *T. castaneum* TERT structure bound to an A-form helix [47]. A recent structure of the *O. latipes* RBD in complex with CR4/5 allowed us to position the P6.1 helix directly on the homology model RBD through alignment [29]. Finally, for modeling runs including an enforced P1 stem, we docked the P1 stem in accordance with the recent *T. thermophila* stem II-RBD structure [48]. This modeling decision was supported by previous biochemical data that established functional homology between the human P1 stem and *T. thermophila* stem II [50]. Combining all of this prior knowledge, the initial modeling point can be envisioned as a telomerase RNP complex where the P6.1 stem, template hybrid, and P1 stem were fixed with respect to the hTERT homology model. The P2a.1, P2ab, and the pseudoknot were free to move relative to TERT and the remaining RNA (47 nucleotides) was structurally unconstrained.

Models of the entire pseudoknot domain RNA within the RNP were generated using a version of the previously reported fragment assembly approach [60]. This method was modified to incorporate smFRET distance constraints and steric clashes between the RNA and the protein (Fig. 3B). Four separate modeling runs were performed, using either the binding or activity FRET data, each with or without enforcing P1 stem formation. Modeling runs consisted of approximately 2500 independent trials, each resulting in a final model. To utilize the FRET constraints in the modeling procedure, it was necessary to convert our measured FRET efficiencies to physical distances using the standard FRET expression (Equation 1). Previous work has highlighted the importance of obtaining appropriate photo-physical correction terms to reliably convert FRET efficiency measurements into accurate distances [23]. While our study permitted us to make several of the necessary corrections by analyzing and adjusting individual FRET trajectories (see methods), accurate experimental determination of the Forster radius (R_o) for each experiment was not feasible due to difficulties generating large quantities of each smFRET-labeled telomerase sample. However, several points suggest that the approximation R_o = 5.6 nm for the Cy3/Cy5 pair used in our modeling is reasonable [61-63]. First, previous work characterizing the observed experimental variation of R_o has shown that, on average, R_o deviates from the standard values by ~+/- 0.2 nm [20]. Second, we conservatively set the $Score_{FRET}$ term to only introduce a penalty if the modeled distance deviates more than ~0.5 nm from the calculated distance, which encompasses even the worst reported errors in R_o (Table S1). Such a high level of

compliance in the distance constraints might be expected to prohibit convergence during modeling; however, the intrinsic redundancy within the FRET network helped localize each dye position within the network (Fig. 1 and 2).

Equation 1. FRET relation to intramolecular distances.

$$FRET = \frac{1}{1 + (R/R_0)^6}$$

Modeling of the pseudoknot domain around the initial RNP scaffold yielded convergent models in each of the four runs with an RNA RMSD of 1.0 nm for runs with the P1 stem formed (Fig. 4A) and 1.4 nm for runs without enforcing the P1 stem (Fig. 4B). Such convergence supports the notion that our network of compliant distance constraints is capable of producing a specific RNP architecture. Indeed, control modeling runs in which FRET constraints were not enforced failed to achieve convergence (RMSD = 3.1 nm, Fig. 4C), demonstrating the importance of experimentally-derived distance constraints in guiding the modeling process. Interestingly, smFRET-Rosetta modeling performed without enforcing P1 structure or location converged to a similar structure (RMSD = 1.3nm between the centers of the most populated clusters. This result implies that positioning of an individual RNA element is not absolutely critical for modeling precision and convergence and suggests that the systematic error of this method is approximately 1.3 nm. Comparison of predicted FRET values (extrapolated from the converged models) versus the experimental FRET values revealed a high degree of correlation, indicating

that the experimental FRET distance constraints were satisfied during the modeling process (Fig. S6). In both sets of convergent models, the pseudoknot domain forms a ring-like structure around TERT. Surprisingly, the catalytically essential pseudoknot fold resides on a protein face opposite the active site, contacting the C-terminal extension (CTE) domain (Fig 4A and 4B). The J2a/b bulge induces a bend in the P2 stem structure wrapping the RNA around the CTE, guiding the free template strand toward the active site. In the case of P1 enforcement, the RNA ring is closed by P1, creating a continuous RNA structure that encircles the telomerase protein component. This pseudoknot conformation is also compatible with the full length RNA as models including the stems P1, P4, and P5 produced convergent solutions (Fig. S7).

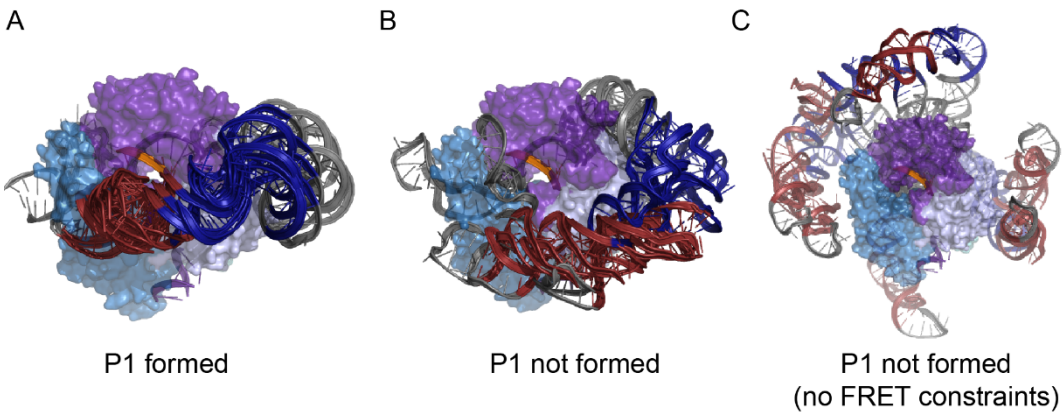


Figure 4. FRET-Rosetta modeling produces convergent solutions to the pseudoknot domain architecture. (A) Superposition of three representative models of the pseudoknot domain RNA modeled in the context of the TERT homology model. These models contain a formed P1 stem at the recently characterized RNA binding site of the RBD (PDB:5C9H). (B) Three representative RNP models generated with an unformed P1 stem, showing similar positioning of the pseudoknot fold relative to P1 formed. (C) Lowest energy models for RNP architecture when the $Score_{FRET}$ term is removed from the scoring function.

Modeling the RNA pseudoknot domain using the distance constraints generated in the presence of activity buffer returned two classes of models (Fig 5A ii and iii). The two distinct model classes are a result of the two peaks in the U92-U312 activity histogram (Fig 2I, grey). Comparison of these activity-dependent models (state ii and state iii) with the primer bound model (state i) reveals a repositioning of the RNA pseudoknot fold around the CTE, consistent with the increase in FRET associated with telomerase activity (Fig 2I and Fig 5A, compare peaks and models). When considering possible sources of the observed RNA pseudoknot conformational change, we recognized that each round of telomere DNA synthesis likely induces template hybrid rotation and extrusion out of the active site (ref). Interestingly, during this putative hybrid rotation, the 3' end of the RNA template closely tracks the proposed pseudoknot domain motion (Fig 5B). This correlation raised the possibility that telomere DNA repeat synthesis at the telomerase active site may be physically coupled to pseudoknot motion through the intervening single stranded TRE (Fig 1, nucleotides 57-63).

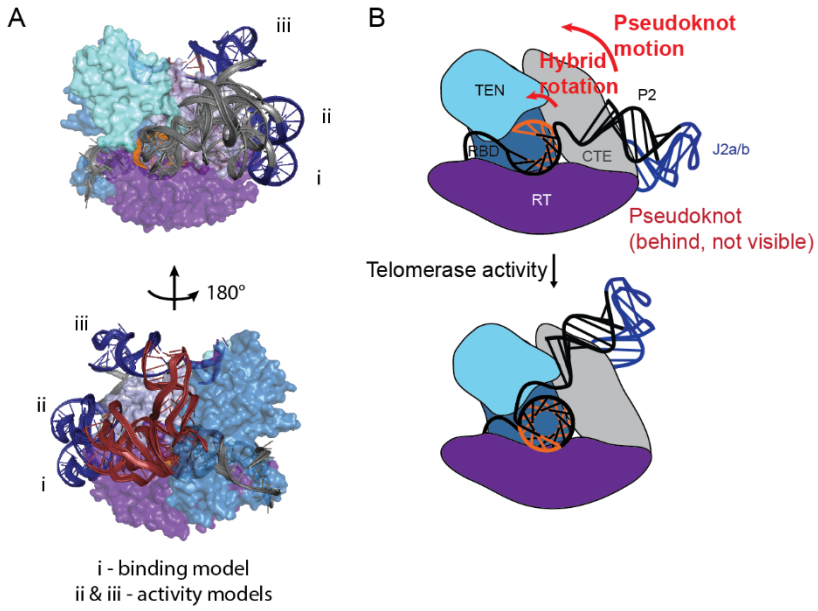


Figure 5. **Telomerase activity induces motion of pseudoknot domain around CTE.** (A) Three separate RNA states observed in the modeling when using binding and activity data. State i is produced when telomerase is modeled with the binding data (red histograms Fig. 2). State ii and iii are a result of modeling the RNP with the activity based data (grey histograms Fig. 2). Both state ii and iii show a rotation of the pseudoknot around the complex away from its original binding position of state I. (B) The cartoon representation of the telomerase complexes shows that the observed motion of the pseudoknot domain tracks the template hybrid rotation that occurs during telomerase activity.

Comparison of our human RNP models with the recently reported 9 Å cryoEM structure of *Tetrahymena* telomerase provided an excellent blind test for the accuracy of our modeling [64]. Excitingly, the human telomerase models place the pseudoknot fold on the same protein face and in a similar position as in the 9 Å *Tetrahymena* telomerase structure (Fig S8). Out of these models, the activity based model is most similar to the *Tetrahymena* structure, which is consistent with the notion that the recent structure represents the pre-translocated state (Fig S8D). This

comparison represents an ideal test case, shedding light on the resolution and accuracy of our modeling methodology, as well as the shared domain organization of telomerase enzymes from diverse species.

Pseudoknot motion tracks telomerase catalysis through the TRE

The hypothesis that hTR conformational changes are coupled to hybrid movement in the active site suggested that physical connectivity between the template region and the pseudoknot fold may be essential for telomerase processivity. We set out to directly test this notion by physically decoupling the template hybrid from the pseudoknot fold using a circular permutation of the pseudoknot domain (cp-hTR) that introduces a physical break between hTR nucleotides 62-63. We reconstituted telomerase enzymes with the cp-hTR pseudoknot domain, and elected to use the FRET pair (U92-U312) that was most sensitive to conformational rearrangements induced by telomerase catalysis. When cp-hTR telomerase was surface-immobilized and analyzed at the single molecule level, we observed a very similar FRET distribution in comparison to wild-type telomerase (Fig 6A and B). Strikingly, addition of activity buffer to the cp-hTR telomerase completely failed to shift the major FRET peak to the higher value that gave rise to the 'state ii' conformation of the RNA pseudoknot in our modeling (compare Fig 6C and 6D). In contrast, the cp-hTR was competent to sample the minor high FRET population that coincided with the 'state iii' conformation of the RNA pseudoknot. Taken together, these results demonstrate that physical connectivity between the template and the pseudoknot fold is strictly required for the activity-dependent conformational change into the 'state ii'

conformation of the RNA pseudoknot (Fig. 6E). Next, we performed bulk telomerase activity assays comparing wild-type and cp-hTR telomerase complexes. Both enzymes were competent to bind telomere DNA primer and extend to the end of the first repeat, but the cp-hTR telomerase was unable to perform the translocation step that is required for processive addition of multiple telomere DNA repeats (Fig. 6F). This result is consistent with a recent report analyzing activity defects of circularly permuted full-length hTR constructs [39], and reveals that physical connectivity between the template hybrid and pseudoknot fold couples telomere DNA synthesis to the hTR pseudoknot motion during telomerase processivity (Fig 6E).

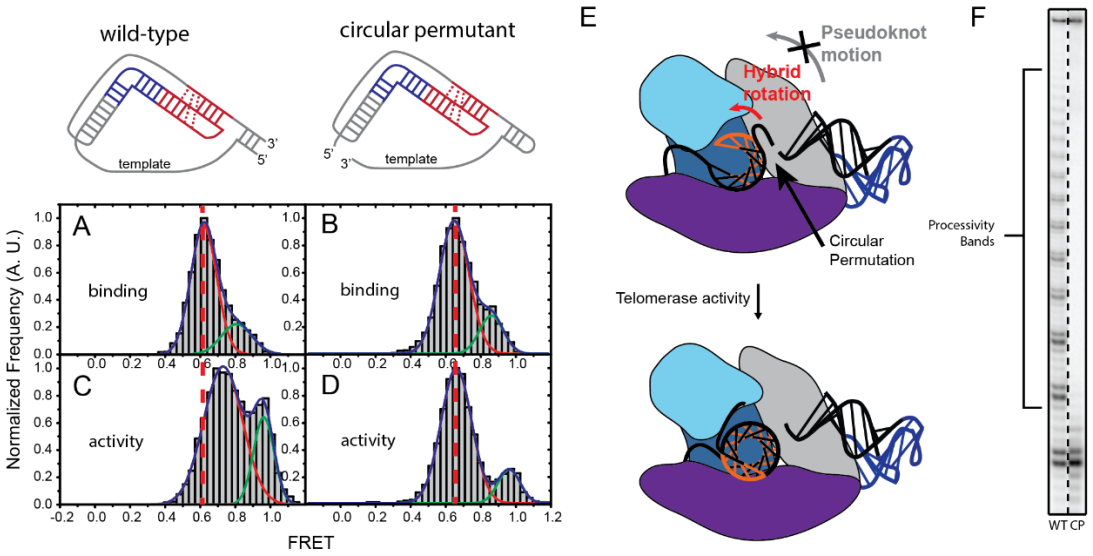


Figure 6. **Template-pseudoknot connectivity is crucial for motion of the pseudoknot domain and telomerase processivity.** (A, C) Single molecule FRET histograms generated by a WT RNP containing a U92-U312 FRET pair before (A) and after (C) addition of activity buffer. The vertical dotted line delineates the observed change in FRET upon addition of activity buffer. (B, D) Single molecule FRET histograms of RNPs containing a circular permutant RNA at hTR 62-63 RNP and a U92-U312 FRET pair, before (B) and after (D) addition of activity buffer. The vertical dotted line delineates the lack of FRET change observed upon addition of activity buffer. (E) Cartoon model depicting the circular permutation at hTR 62-63

decouples pseudoknot motion from the activity-induced rotation of the template hybrid. (F) Activity assay of WT and circular permutant enzymes with a $(TG)_6TTAGGG$ primer substrate. WT products show binding, extension, and translocation, resulting in processivity bands. Circular permutant complexes exhibit binding and extension, but failed translocation and therefore a lack of processivity bands.

DISCUSSION

Our FRET experiments and modeling suggest the dynamic motion of the pseudoknot domain is coupled to synthesis at the active site through the TRE. The pseudoknot domain is indispensable for human telomerase activity, yet there is no experimental evidence as to how the pseudoknot contributes to catalysis. One interesting possibility integrates a recent proposal that suggests the CTE domain is mobile during the processive action of telomerase [65,66]. This concept arises from the homology between TERT and other polymerases, where the CTE of TERT is structurally homologous to the thumb domain of other polymerases. In this mechanism, telomerase CTE undergoes a conformational change at the end of the catalytic cycle, which destabilizes hybrid binding in the active site and promotes telomerase translocation. Interestingly, our model of the pseudoknot domain reveals that the pseudoknot fold resides on the CTE protein face. Pseudoknot guiding of CTE conformation could resolve the quandary of how an RNA motif distal from the active site is crucial for enzyme function. Furthermore the observed motion of the pseudoknot could act to specifically modulate CTE conformation, aiding in disengagement of the template hybrid and therefore the processive action of telomerase.

Pseudoknot motion is physically coupled to template hybrid rotation through the TRE, yet it remains unclear what produces the driving force for this motion. One possibility is that catalytically induced template hybrid rotation could physically pull the pseudoknot around the CTE through the single stranded TRE RNA. A second, but not mutually exclusive possibility, is that the TRE itself plays an important role in telomerase catalysis. Biochemical mapping in Tetrahymena telomerase showed that the TEN domain binds to telomerase RNA, and more specifically to the TRE [67]. Initial substrate binding rotates the template hybrid in a way that positions the TRE distal from the TEN domain. Extension of the template hybrid drives extrusion and rotation of the template hybrid, forcing the TRE closer to the TEN domain, allowing TRE-TEN interaction. In this way, hybrid rotation could govern TRE-TEN interaction, thereby influencing pseudoknot location.

Rosetta modeling integrated with FRET distance constraints provided the human telomerase RNP architecture, paving the way for future studies involving other biological systems for use with other physical constraints. An example of a biological system primed for study with our methodology is the spliceosome. The recent structural characterization of the spliceosome integrated with smFRET-Rosetta could act to delineate the conformational changes associated the multitude of spliceosome sub-states [68]. Although the modeled telomerase assembly integrated known RNA structures, this modeling method could be applied to completely de novo RNA structure prediction. This is especially true for low-mid resolution density maps generated by cryoEM, as there is no standardized algorithm for de novo RNA

modeling. Similarly, the modeling algorithm can be refined to include other physical parameters such as electrostatic interactions of RNA and protein surfaces and chemical mapping data [69]. A few well-tested enhancements of the modeling algorithm could lead to a simple tool that provides accurate RNP architectures through interchangeable physical constraints.

REFERENCES

1 Weinrich, S. L. *et al.* Reconstitution of human telomerase with the template RNA component hTR and the catalytic protein subunit hTRT. *Nature genetics* **17**, 498-502 (1997).

2 Galej, W. P., Nguyen, T. H., Newman, A. J. & Nagai, K. Structural studies of the spliceosome: zooming into the heart of the machine. *Current opinion in structural biology* **25**, 57-66 (2014).

3 Roy, R., Hohng, S. & Ha, T. A practical guide to single-molecule FRET. *Nature methods* **5**, 507-516 (2008).

4 Zhuang, X. *et al.* A single-molecule study of RNA catalysis and folding. *Science* **288**, 2048-2051 (2000).

5 Weiss, S. Fluorescence spectroscopy of single biomolecules. *Science* **283**, 1676-1683 (1999).

6 Stryer, L. Fluorescence energy transfer as a spectroscopic ruler. *Annual review of biochemistry* **47**, 819-846 (1978).

7 Akiyama, B. M., Parks, J. W. & Stone, M. D. The telomerase essential N-terminal domain promotes DNA synthesis by stabilizing short RNA-DNA hybrids. *Nucleic acids research* **43**, 5537-5549 (2015).

8 Berman, A. J., Akiyama, B. M., Stone, M. D. & Cech, T. R. The RNA accordion model for template positioning by telomerase RNA during telomeric DNA synthesis. *Nature structural & molecular biology* **18**, 1371-1375 (2011).

9 Parks, J. W. & Stone, M. D. Coordinated DNA dynamics during the human telomerase catalytic cycle. *Nature communications* **5**, 4146 (2014).

10 Hengesbach, M., Kim, N. K., Feigon, J. & Stone, M. D. Single-molecule FRET reveals the folding dynamics of the human telomerase RNA pseudoknot domain. *Angewandte Chemie* **51**, 5876-5879 (2012).

11 Mihalusova, M., Wu, J. Y. & Zhuang, X. Functional importance of telomerase pseudoknot revealed by single-molecule analysis. *Proceedings of the National Academy of Sciences of the United States of America* **108**, 20339-20344 (2011).

12 Stone, M. D. *et al.* Stepwise protein-mediated RNA folding directs assembly of telomerase ribonucleoprotein. *Nature* **446**, 458-461 (2007).

13 Wu, J. Y., Stone, M. D. & Zhuang, X. A single-molecule assay for telomerase structure-function analysis. *Nucleic acids research* **38**, e16 (2010).

14 Hengesbach, M., Akiyama, B. M. & Stone, M. D. Single-molecule analysis of telomerase structure and function. *Current opinion in chemical biology* **15**, 845-852 (2011).

15 Gavory, G., Symmons, M. F., Krishnan Ghosh, Y., Klenerman, D. & Balasubramanian, S. Structural analysis of the catalytic core of human telomerase RNA by FRET and molecular modeling. *Biochemistry* **45**, 13304-13311 (2006).

16 Cole, D. I. *et al.* New models of Tetrahymena telomerase RNA from experimentally derived constraints and modeling. *Journal of the American Chemical Society* **134**, 20070-20080 (2012).

17 Muschielok, A. & Michaelis, J. Application of the nano-positioning system to the analysis of fluorescence resonance energy transfer networks. *The journal of physical chemistry. B* **115**, 11927-11937 (2011).

18 Dale, R. E., Eisinger, J. & Blumberg, W. E. The orientational freedom of molecular probes. The orientation factor in intramolecular energy transfer. *Biophysical journal* **26**, 161-193 (1979).

19 Stephenson, J. D., Kenyon, J. C., Symmons, M. F. & Lever, A. M. Characterizing 3D RNA Structure by Single Molecule FRET. *Methods* (2016).

20 Andrecka, J. *et al.* Nano positioning system reveals the course of upstream and nontemplate DNA within the RNA polymerase II elongation complex. *Nucleic acids research* **37**, 5803-5809 (2009).

21 Kalinin, S. *et al.* A toolkit and benchmark study for FRET-restrained high-precision structural modeling. *Nature methods* **9**, 1218-1225 (2012).

22 Andrecka, J. *et al.* Single-molecule tracking of mRNA exiting from RNA polymerase II. *Proceedings of the National Academy of Sciences of the United States of America* **105**, 135-140 (2008).

23 Muschielok, A. *et al.* A nano-positioning system for macromolecular structural analysis. *Nature methods* **5**, 965-971 (2008).

24 Greider, C. W. & Blackburn, E. H. The telomere terminal transferase of Tetrahymena is a ribonucleoprotein enzyme with two kinds of primer specificity. *Cell* **51**, 887-898 (1987).

25 Morin, G. B. The human telomere terminal transferase enzyme is a ribonucleoprotein that synthesizes TTAGGG repeats. *Cell* **59**, 521-529 (1989).

26 Kim, N. K., Zhang, Q. & Feigon, J. Structure and sequence elements of the CR4/5 domain of medaka telomerase RNA important for telomerase function. *Nucleic acids research* **42**, 3395-3408 (2014).

27 Kim, N. K. *et al.* Solution structure and dynamics of the wild-type pseudoknot of human telomerase RNA. *Journal of molecular biology* **384**, 1249-1261 (2008).

28 Zhang, Q., Kim, N. K. & Feigon, J. Architecture of human telomerase RNA. *Proceedings of the National Academy of Sciences of the United States of America* **108**, 20325-20332 (2011).

29 Huang, J. *et al.* Structural basis for protein-RNA recognition in telomerase. *Nature structural & molecular biology* **21**, 507-512 (2014).

30 Qiao, F. & Cech, T. R. Triple-helix structure in telomerase RNA contributes to catalysis. *Nature structural & molecular biology* **15**, 634-640 (2008).

31 Chen, J. L. & Greider, C. W. Telomerase RNA structure and function: implications for dyskeratosis congenita. *Trends in biochemical sciences* **29**, 183-192 (2004).

32 Vulliamy, T. *et al.* The RNA component of telomerase is mutated in autosomal dominant dyskeratosis congenita. *Nature* **413**, 432-435 (2001).

33 Brown, A. F. *et al.* A self-regulating template in human telomerase. *Proceedings of the National Academy of Sciences of the United States of America* **111**, 11311-11316 (2014).

34 Lai, C. K., Miller, M. C. & Collins, K. Template boundary definition in Tetrahymena telomerase. *Genes & development* **16**, 415-420 (2002).

35 Miller, M. C., Liu, J. K. & Collins, K. Template definition by Tetrahymena telomerase reverse transcriptase. *The EMBO journal* **19**, 4412-4422 (2000).

36 Greider, C. W. Telomerase is processive. *Molecular and cellular biology* **11**, 4572-4580 (1991).

37 Blackburn, E. H. *et al.* Recognition and elongation of telomeres by telomerase. *Genome / National Research Council Canada = Genome / Conseil national de recherches Canada* **31**, 553-560 (1989).

38 Qi, X. *et al.* RNA/DNA hybrid binding affinity determines telomerase template-translocation efficiency. *The EMBO journal* **31**, 150-161 (2012).

39 Mefford, M. A. & Zappulla, D. C. Physical Connectivity Mapping by Circular Permutation of Human Telomerase RNA Reveals New Regions Critical for Activity and Processivity. *Molecular and cellular biology* **36**, 251-261 (2015).

40 Akiyama, B. M. & Stone, M. D. Assembly of complex RNAs by splinted ligation. *Methods in enzymology* **469**, 27-46 (2009).

41 Das, R., Laederach, A., Pearlman, S. M., Herschlag, D. & Altman, R. B. SAFA: semi-automated footprinting analysis software for high-throughput quantification of nucleic acid footprinting experiments. *Rna* **11**, 344-354 (2005).

42 McCann, J. J., Choi, U. B., Zheng, L., Weninger, K. & Bowen, M. E. Optimizing methods to recover absolute FRET efficiency from immobilized single molecules. *Biophysical journal* **99**, 961-970 (2010).

43 Harkisheimer, M., Mason, M., Shuvaeva, E. & Skordalakes, E. A motif in the vertebrate telomerase N-terminal linker of TERT contributes to RNA binding and telomerase activity and processivity. *Structure* **21**, 1870-1878 (2013).

44 Soding, J., Biegert, A. & Lupas, A. N. The HHpred interactive server for protein homology detection and structure prediction. *Nucleic acids research* **33**, W244-248 (2005).

45 Eswar, N. *et al.* Comparative protein structure modeling using MODELLER. *Current protocols in protein science / editorial board, John E. Coligan ... [et al.]* **Chapter 2**, Unit 2 9 (2007).

46 Rouda, S. & Skordalakes, E. Structure of the RNA-binding domain of telomerase: implications for RNA recognition and binding. *Structure* **15**, 1403-1412 (2007).

47 Mitchell, M., Gillis, A., Futahashi, M., Fujiwara, H. & Skordalakes, E. Structural basis for telomerase catalytic subunit TERT binding to RNA template and telomeric DNA. *Nature structural & molecular biology* **17**, 513-518 (2010).

48 Jansson, L. I. *et al.* Structural basis of template-boundary definition in Tetrahymena telomerase. *Nature structural & molecular biology* **22**, 883-888 (2015).

49 Bley, C. J. *et al.* RNA-protein binding interface in the telomerase ribonucleoprotein. *Proceedings of the National Academy of Sciences of the United States of America* **108**, 20333-20338 (2011).

50 Chen, J. L. & Greider, C. W. Template boundary definition in mammalian telomerase. *Genes & development* **17**, 2747-2752 (2003).

51 Wu, R. A., Dagdas, Y. S., Yilmaz, S. T., Yildiz, A. & Collins, K. Single-molecule imaging of telomerase reverse transcriptase in human telomerase holoenzyme and minimal RNP complexes. *eLife* **4** (2015).

52 Zhang, Q., Kim, N. K., Peterson, R. D., Wang, Z. & Feigon, J. Structurally conserved five nucleotide bulge determines the overall topology of the core domain of human telomerase RNA. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 18761-18768 (2010).

53 Zemora, G., Handl, S. & Waldsich, C. Human telomerase reverse transcriptase binds to a pre-organized hTR in vivo exposing its template. *Nucleic acids research* **44**, 413-425 (2016).

54 Latrick, C. M. & Cech, T. R. POT1-TPP1 enhances telomerase processivity by slowing primer dissociation and aiding translocation. *The EMBO journal* **29**, 924-933 (2010).

55 Wu, R. A. & Collins, K. Human telomerase specialization for repeat synthesis by unique handling of primer-template duplex. *The EMBO journal* **33**, 921-935 (2014).

56 Chen, J. L., Opperman, K. K. & Greider, C. W. A critical stem-loop structure in the CR4-CR5 domain of mammalian telomerase RNA. *Nucleic acids research* **30**, 592-597 (2002).

57 Gillis, A. J., Schuller, A. P. & Skordalakes, E. Structure of the Tribolium castaneum telomerase catalytic subunit TERT. *Nature* **455**, 633-637 (2008).

58 Ly, H., Blackburn, E. H. & Parslow, T. G. Comprehensive structure-function analysis of the core domain of human telomerase RNA. *Molecular and cellular biology* **23**, 6849-6856 (2003).

59 Steczkiewicz, K. *et al.* Human telomerase model shows the role of the TEN domain in advancing the double helix for the next polymerization step. *Proceedings of the National Academy of Sciences of the United States of America* **108**, 9443-9448 (2011).

60 Das, R. & Baker, D. Automated de novo prediction of native-like RNA tertiary structures. *Proceedings of the National Academy of Sciences of the United States of America* **104**, 14664-14669 (2007).

61 Ishikawa-Ankerhold, H. C., Ankerhold, R. & Drummen, G. P. Advanced fluorescence microscopy techniques--FRAP, FLIP, FLAP, FRET and FLIM. *Molecules* **17**, 4047-4132 (2012).

62 Ha, T. *et al.* Initiation and re-initiation of DNA unwinding by the Escherichia coli Rep helicase. *Nature* **419**, 638-641 (2002).

63 Yuan, F. *et al.* Use of a novel Forster resonance energy transfer method to identify locations of site-bound metal ions in the U2-U6 snRNA complex. *Nucleic acids research* **35**, 2833-2845 (2007).

64 Jiang, J. *et al.* Structure of Tetrahymena telomerase reveals previously unknown subunits, functions, and interactions. *Science* **350**, aab4070 (2015).

65 Yang, W. & Lee, Y. S. A DNA-hairpin model for repeat-addition processivity in telomere synthesis. *Nature structural & molecular biology* **22**, 844-847 (2015).

66 Bryan, C. *et al.* Structural Basis of Telomerase Inhibition by the Highly Specific BIBR1532. *Structure* **23**, 1934-1942 (2015).

67 O'Connor, C. M., Lai, C. K. & Collins, K. Two purified domains of telomerase reverse transcriptase reconstitute sequence-specific interactions with RNA. *The Journal of biological chemistry* **280**, 17533-17539 (2005).

68 Yan, C. *et al.* Structure of a yeast spliceosome at 3.6-angstrom resolution. *Science* **349**, 1182-1191 (2015).

69 Cheng, C. Y. *et al.* Consistent global structures of complex RNA states through multidimensional chemical mapping. *eLife* **4**, e07600 (2015).

SUPPLEMENTARY INFORMATION

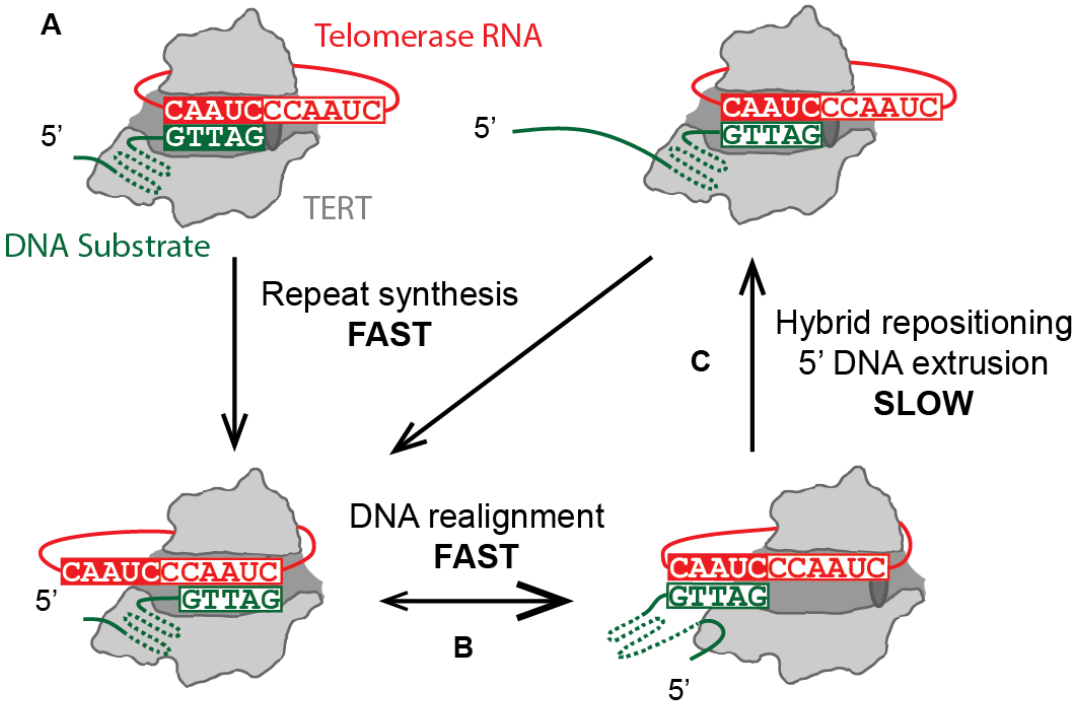
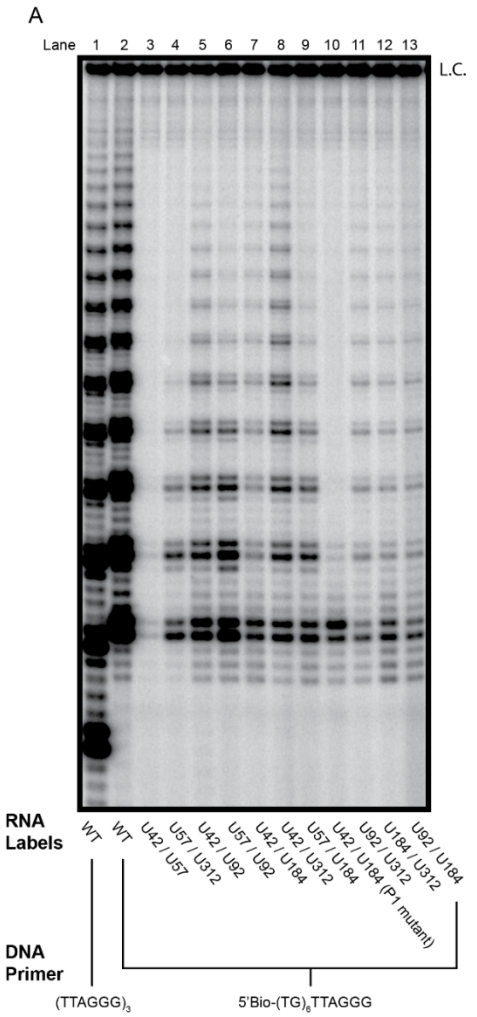
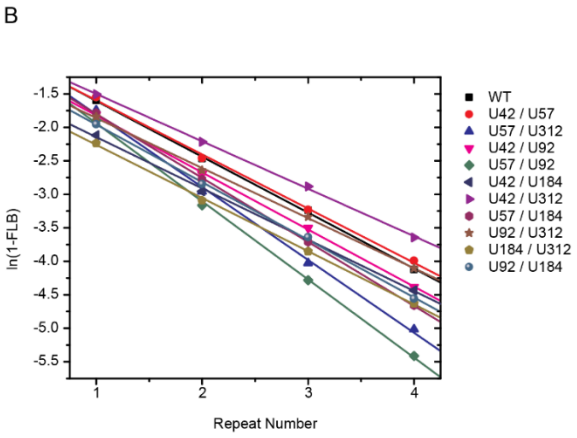


Figure S1. The telomerase catalytic cycle. (A) Telomerase initially binds a telomere DNA substrate through RNA:DNA hybrid interactions with the template. Rapid synthesis of a complete telomere DNA repeat extends the RNA:DNA hybrid to the template boundary, at which point a dynamic equilibrium of DNA primer states exists (B), favoring the formation of a realigned RNA:DNA hybrid. (C) In a second, rate-limiting, step of telomerase translocation, the realigned RNA:DNA hybrid is repositioned in the telomerase active site and the 5′-end of the DNA primer is extruded from the enzyme complex.





C

(relative to WT)

Enzyme	slope	R^2	$R_{1/2}$	Processivity	Activity
WT	-0.83	0.99	0.83	1.00	1.00
U42 / U57	-0.81	0.99	0.86	1.04	0.09
U57 / U312	-1.09	0.99	0.64	0.77	0.16
U42 / U92	-0.85	0.99	0.82	0.99	0.29
U57 / U92	-1.17	0.99	0.59	0.71	0.31
U42 / U184	-0.77	0.99	0.90	1.08	0.21
U42 / U312	-0.71	0.99	0.98	1.18	0.29
U57 / U184	-0.96	0.99	0.72	0.87	0.20
U92 / U312	-0.75	0.99	0.92	1.11	0.20
U184 / U312	-0.79	0.99	0.88	1.06	0.14
U92 / U184	-0.86	0.99	0.81	0.98	0.17
U42 / U184 P1mutant	N/A	N/A	N/A	N/A	0.15
U92 / U312 Circular Permutant	N/A	N/A	N/A	N/A	1.16

Figure S2. Activity of modified telomerase complexes. (A) Direct DNA primer extension assays for telomerase enzymes and DNA primers used in the study. Activity assays were performed as described in the detailed methods section unless otherwise noted. Total enzyme activity was defined as the total counts per lane corrected for the loading control (LC). (1) Unlabeled telomerase extension of $(TTAGGG)_3$ DNA primer. (2) Unlabeled telomerase extension of 5'- biotinylated $(TG)_6TTAGGG$ DNA primer. The retardation of products in this lane is due to the 5'-biotin moiety. (3-13) Extension reactions performed with the biotinylated $(TG)_6TTAGGG$ DNA primer. The difference in total activity for these enzymes versus WT can likely be account for by the difference in RNA concentration used in the reconstitution reaction, where lower concentrations were used for dye labeled enzymes to preserve RNA. Another possibility, that is not mutually exclusive, is different lysate batches produce varying reconstitution efficiencies. Note that the

level of total activity for these enzymes mirrored the efficiency of their single molecule immobilizations (data not shown). (3) U42-U57 labeled telomerase. (4) U57-U312 labeled telomerase. (5) U42-U92 labeled telomerase. (6) U57-U92 labeled telomerase. (7) U42-U184 labeled telomerase. (8) U42-U312 labeled telomerase. (9) U57-U184 labeled telomerase. (10) U42-U184 labeled telomerase, P1 mutated. The lack of processivity bands from this lane seems to be a product of template read-through, as observed by the higher intensity band at +1 relative to all other telomerase products. (11) U92-U312 labeled telomerase. (12) U184-U312 labeled telomerase. (13) U92-U184 labeled telomerase. See main text for circular permutant activity assay. (B) After quantification of the bands using SAFA 11b, the processivities were calculated using a previously reported method (1). In short, the fraction of DNA left bound (FLB) at each repeat band was determined. A linear fit of ln(1-FLB) versus repeat number was plotted and fit by linear regression using the expression ln(1-FLB)=(0.693/$R_{1/2}$)*(repeat number). Dividing 0.693 by the slope of each line provides $R_{1/2}$, the processivity value, for the respective enzyme. (C) All the processivity values were very similar except for the P1 mutant and circular permutant enzymes, which showed no processivity. The similarity in processivity argues that the active enzyme complexes that were observed at the single molecule level maintained WT level activity.

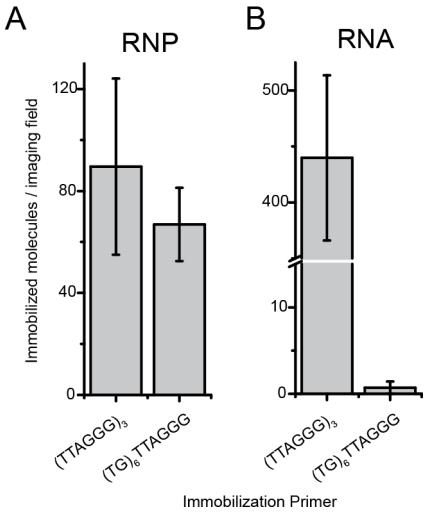


Figure S3. A modified telomere primer allows immobilization of telomerase RNP but not protein-free telomerase RNA. (A) U42-U92 labeled telomerase was immobilized through a primer with the sequence $(TTAGGG)_3$ or $(TG)_6TTAGGG$. Immobilization with either primer produced similar levels of immobilized complexes, depicting that the $(TG)_6TTAGGG$ primer, which forms a 7nt hybrid with the RNA template, is sufficient for high affinity pulldown of reconstituted telomerase complexes. (B) A $(TTAGGG)_3$ primer allowed robust immobilization of protein-free dye labeled RNA by forming a 11 nucleotide hybrid with the template region. No RNA was immobilized when a primer with the sequence $(TG)_6TTAGGG$ was used. The lack of immobilization is expected as the relatively short 7nt hybrid with the template region is insufficient for robust RNA pulldown in a protein-free context. In these experiments the biotinylated primers were surface immobilized at a concentration of 100 nM. Telomerase RNPs were immobilized after immuno-precipitation with a 1:3 dilution. In the context of protein-free RNA experiments, 1 nM dye labeled RNA was used for single molecule immobilization. A standard threshold analysis of the single molecule traces was used to parse out the number of molecules containing acceptor fluorescence per imaging field.

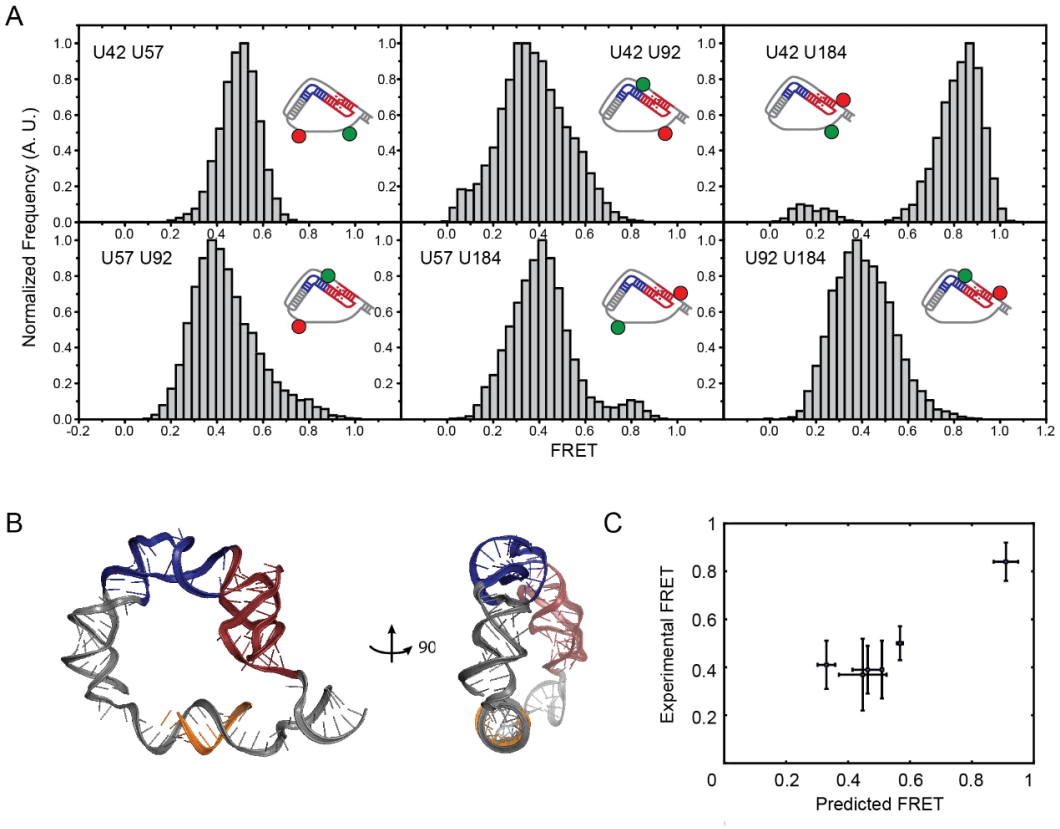


Figure S4. FRET-Rosetta modeling reveals the protein-free pseudoknot domain is consistent with triangular RNA architecture. (A) Single molecule FRET histograms of dye labeled pseudoknot domains immobilized on a $(TTAGGG)_3$ primer. (Top-left) U42-U57 labeled telomerase. (Top-middle) U42-U92 labeled telomerase. (Top-right) U42-U184 labeled telomerase. (Bottom-left) U57-U92 labeled telomerase. (Bottom-middle) U57-U184 labeled telomerase. (Bottom-right) U92-U184 labeled telomerase. (B) Representative pseudoknot domain RNA model returned from FRET-Rosetta modeling algorithm using $Score_{FRET}$ and $Score_{RNA}$ terms. (C) Plotting the predicted FRET (based on distances in models) versus the experimental FRET reveals a high degree of correlation, showing that the converged models satisfied the FRET distance constraints during the modeling process.

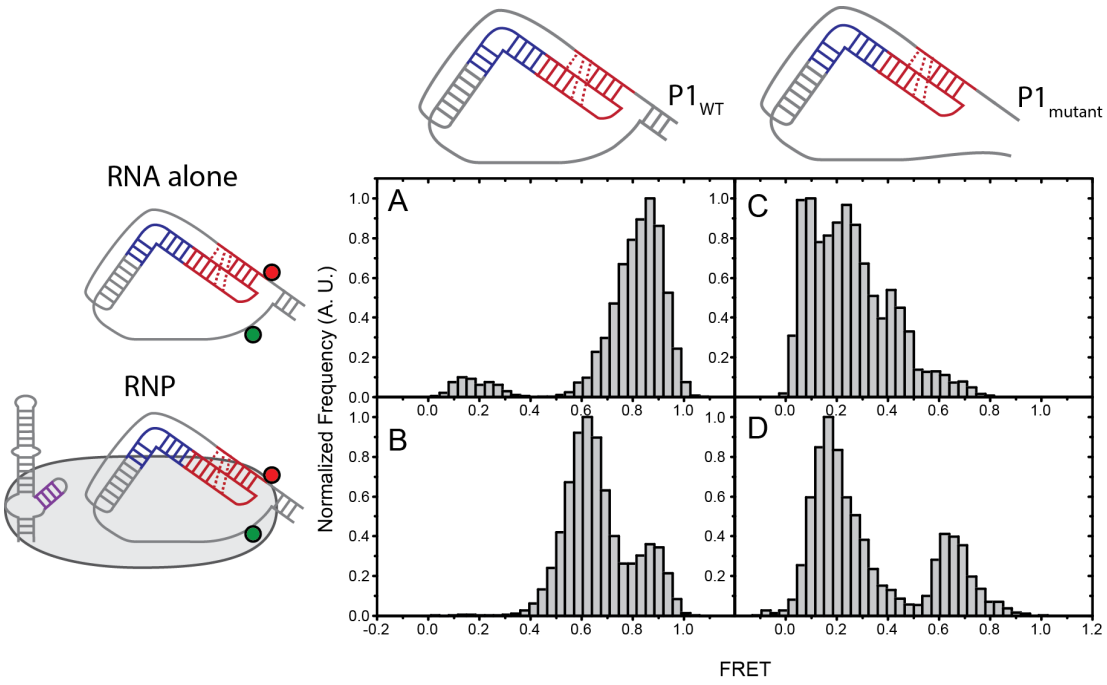


Figure S5. Disruption of P1 stem completely deforms protein-free RNA architecture and partially deforms RNA architecture in the RNP. Each histogram was generated by immobilizing either U42-U184 FRET labeled RNA with $(TTAGGG)_3$ primer. This allowed immobilization of both protein-free RNA as well as assembled RNPs. (A) WT protein-free pseudoknot domain RNA shows a single predominant distribution at high FRET. (B) Mutation of the P1 stem reveals that protein-free RNA without P1 stem formation adopts an entirely different conformation, where the main FRET peak splits and shifts towards lower FRET. This result highlights the importance of the P1 stem for capturing the pseudoknot domain in a circular architecture. (C) When the WT RNP is immobilized, two predominant distributions are present. The major distribution represents the assembled RNP whereas the high FRET peak likely represents misassembled or unassembled RNP/RNA. (D) Mutation of the P1 in the context of the RNP results in the major distribution shifting toward low FRET, suggesting most of the complexes are misassembled. The second minor distribution likely represents properly assembled RNP even in the absence of the P1 stem, suggesting the P1 is important but not absolutely required for RNP assembly.

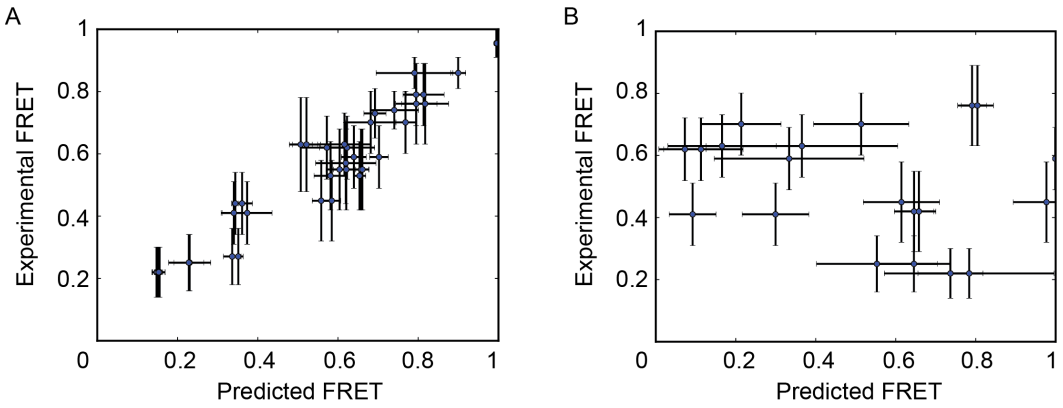


Figure S6. FRET distance constraints are satisfied in the convergent models when the $Score_{FRET}$ term is employed. (A) Distances from the Rosetta models utilizing the FRET distance constraints are converted to predicted FRETs and plotted versus the corresponding experimental FRET values. The high degree of correlation confirms that the modeling process is satisfying all distance constraints during the modeling process. (B) Removal of the FRET distance constraints during Rosetta modeling results in predicted FRETs that are totally uncorrelated from the experimental FRET values, suggesting the $Score_{RNA}$ and $Score_{clash}$ are insufficient for producing an accurate representation of the experimental FRET values.

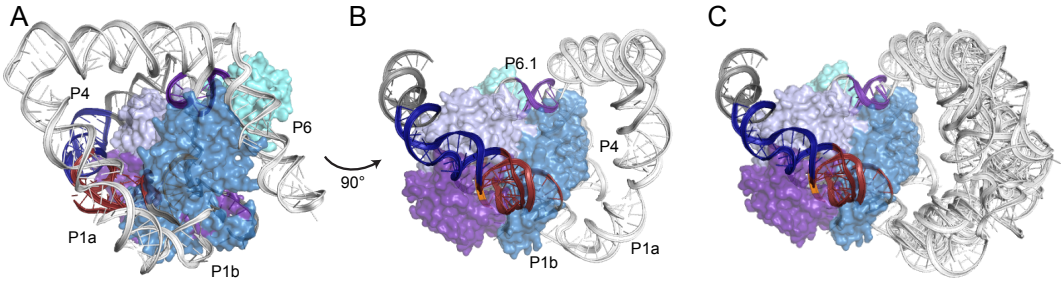


Figure S7. Native stem structures are capable of bridging the pseudoknot and CR4/5 domains. White RNA structures are P1,P4, and P5 bridging stems.

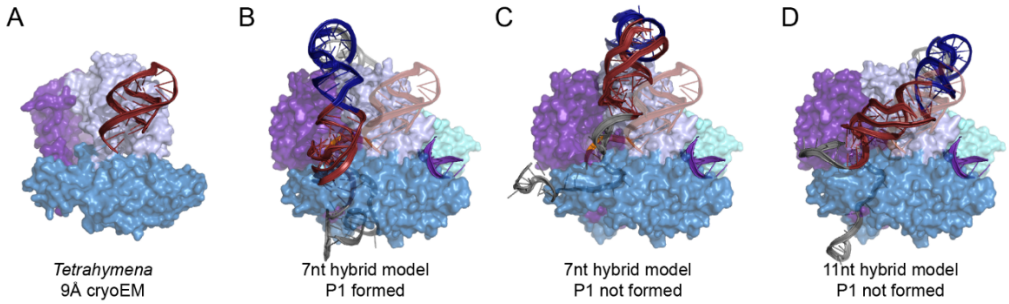


Figure S8. Convergent human RNP models reveal striking homology between human and *Tetrahymena* RNA architectures. (A) The *Tetrahymena* cryoEM structure is reduced to depiction of the TERT protein component and the pseudoknot fold. (B) Comparison of P1 formed pseudoknot (red) model versus the *Tetrahymena* pseudoknot (opaque red ovelay) (C) Comparison of P1 not formed pseudoknot (red) model versus the *Tetrahymena* pseudoknot (opaque red ovelay). (D) Comparison of state iii activity based RNA architecture versus the *Tetrahymena* pseudoknot (opaque red ovelay).

Table S1. FRET distance constraints and errors used for Rosetta modeling

7 nucleotide hybrid "binding" models

Pair of atoms (C5)	Lower bound (mean - std. dev.- 5 Å)	Upper bound (mean + std. dev. + 5 Å)
U42-U57	48.1	68.5
U42-U92	59.5	80.8
U42-U184	44.0	61.4
U42-U312	57.5	78.8
U57-U92	39.4	57.3
U57-U184	50.6	69.0
U57-U312	34.5	56.2
U92-U184	49.2, 44.4	70.0, 64.1
U92-U312	42.8	60.3
U184-U312	42.4, 39.4	59.9, 54.4

11 nucleotide hybrid "activity" models

Pair of atoms (C5)	Lower bound (mean - std. dev.- 5 Å)	Upper bound (mean + std. dev. + 5 Å)
U42-U57	45.9	64.1
U42-U92	59.5	80.8
U42-U184	43.2, 33.1	60.6, 49.0
U42-U312	56.6	77.1
U57-U92	40.3	61.8
U57-U184	49.5	67.5
U57-U312	43.6, 34.5	64.9, 54.0
U92-U184	49.2, 44.4	70.0, 64.1
U92-U312	36.4, -5.0	57.3, 43.1
U184-U312	43.6, 39.0	63.3, 55.5

Table S2. DNA and RNA oligos used for preparation of RNA fragments.

RNA /DNA	Name	Sequence (label site)	Function	hTR segment
RNA	U42 labeled RNA	5'-GGGCCAUUUUUUGUCUAACCCUAACUGAGAA-3'	Dye labeling pseudoknot RNA	32-62
RNA	U57 labeled RNA	5'-GGGCCAUUUUUUGUCUAACCCUAACUGAGAA-3'	Dye labeling pseudoknot RNA	32-62
RNA	U42 short labeled RNA	5'-GGGCCAUUUUUUGUCUAAC-3'	Dye labeling pseudoknot RNA	32-50
RNA	U57 short labeled RNA	5'-P-CCUAACUGAGAA-3'	Dye labeling pseudoknot RNA	51-62
RNA	U92 labeled RNA	5'-P-GGGCGUAGGCGCCGUGCUUUUGCUCCCCGUG CGCUGUUUUUCUCGCUGACUUUCAGC-3'	Dye labeling pseudoknot RNA	63-119
RNA	U184 labeled RNA	5'-P-GGGCGGAAAAGCCUCGGCCUGCCGCCUUCCAC CGUUCAUUCUAGAGCAAACAAAAAAUGUCAGCUGC UGGCCCGUU-3'	Dye labeling pseudoknot RNA	120-195
RNA	U312 labeled RNA	5'-P-CCACCGCGAAGAGUUGGGCUCUGUCAGCCGCG GGUC-3'	Dye labeling CR4/5 RNA	293-329
RNA	CR4/5 unlabeled RNA	5'-GAACCCCGCCUGGAGGCCGCGGUCGGCCCGGGC UUCUCCGGAGGCACCCACUG-3'	Dye labeling CR4/5 RNA	239-292
DNA	DNA splint 1 for PK ligation	5'-CAGCGCGCGGGGAGCAAAAGCACGGCGCCTACG CCCTTCT CAGTTAGGGTTAGACAAAAAATGGCCACC ACCCCTCCCAGG-3'	Splint 5' pseudoknot fragments for ligation	
DNA	DNA splint 2 for PK ligation	5'-CAGCTGACATTTTTTGTTTGCTCTAGAATGAACGG TGGAAGGCGGCAGGCCGAGGCTTTTCCGCCCGCTGA AAGTCAGCGAGAAAAACAG-3'	Splint 3' pseudoknot fragments for ligation	
DNA	Telomere primer	5'-Bio-TTAGGGTTAGGGTTAGGG	Immobilization of RNA and RNPs for single molecule experiments	
DNA	Modified telomere primer	5'-Bio-TGTGTGTGTGTTAGGG	Specific immobilization of RNPs for single molecule experiments	
DNA	hTR PCR Primer 63-195 Forward	5'-TAATACGACTCACTATAGAGAAGGGCGTAGGC-3'	Template for in-vitro transcription of 63-195 for PK construction	
DNA	hTR PCR Primer 63-195 Reverse	5'-AACGGGCCAGCAG-3'		
DNA	hTR PCR Primer 120-195+32-63 (CP)	5'-TTCTAATACGACTCACTATAGGGCGGAAA-3'	PCR assembly of circular permutant RNA (generated by Primerize)	
DNA	hTR PCR Primer 120-195+32-63 (CP)	5'-GGTGGAAGGCGGCAGGCCGAGGCTTTTCCGCCCT ATAGTGAGTCGTA3'		
DNA	hTR PCR Primer 120-195+32-63 (CP)	5'-GCCGCCTTCCACCGTTCATTCTAGAGCAAACAAAA AATGTCAGCTGCTGGCCCGTTG-3'		
DNA	hTR PCR Primer 120-195+32-63 (CP)	5'-TTCTCAGTTAGGGTTAGACAAAAAATGGCCCACCA ACGGGCCAGCAGCTGACAT-3'		

SUPPLEMENTARY REFERENCES

1. Latrick CM & Cech TR (2010) POT1-TPP1 enhances telomerase processivity by slowing primer dissociation and aiding translocation. *The EMBO journal* 29(5):924-933.

CHAPTER IV

Mechanical unfolding of human telomere G-quadruplex DNA probed by integrated fluorescence and magnetic tweezers spectroscopy.

*Originally published in Nucleic Acids Research 2013

ABSTRACT

Single-molecule techniques facilitate analysis of mechanical transitions within nucleic acids and proteins. Here, we describe an integrated fluorescence and magnetic tweezers instrument that permits detection of nanometer-scale DNA structural rearrangements together with the application of a wide range of stretching forces to individual DNA molecules. We have analyzed the force-dependent equilibrium and rate constants for telomere DNA G-quadruplex (GQ) folding and unfolding, and have determined the location of the transition state barrier along the well-defined DNA stretching reaction coordinate. Our results reveal the mechanical unfolding pathway of the telomere DNA GQ is characterized by a short distance (< 1 nm) to the transition state for the unfolding reaction. This mechanical unfolding response reflects a critical contribution of long-range interactions to the global stability of the GQ fold, and suggests that telomere-associated proteins need only disrupt a few base pairs to destabilize GQ structures. Comparison of the GQ unfolded state to a single stranded polyT DNA revealed the unfolded GQ exhibits a compacted non-native conformation reminiscent of the protein molten globule. We expect the capacity to interrogate macromolecular structural transitions with high spatial resolution under

conditions of low forces will have broad application in analyses of nucleic acid and protein folding.

INTRODUCTION

Model RNA and DNA hairpins have been extensively characterized by optical trapping methods, providing a detailed view of the folding energy landscapes of these fundamental nucleic acid structures [1-5]. However, the ability to achieve high spatial resolution using force spectroscopy relies upon the application of relatively large stretching forces (>10 pN) to suppress the measurement noise introduced by the long, flexible DNA handles used to attach the structure of interest to micron-scale beads held in the optical trap. Recently, a study of the force-dependent structural dynamics of single Holliday junctions was reported, which paired an optical trap to apply precisely calibrated stretching forces with single molecule Förster resonance energy transfer (smFRET) to monitor DNA structural dynamics [6]. This fluorescence-force method provided a powerful tool for probing sub-nanometer scale structural rearrangements within single Holliday junctions at very low stretching forces (< 1pN) over short periods of time. In addition, several groups have reported measurements which combine smFRET with a magnetic tweezers apparatus [7,8]. The use of magnets to apply mechanical loads to individual DNA molecules has several potential advantages over optical traps. For example, the combination of a high power optical trapping laser with single molecule fluorescence is technically difficult to implement, typically requiring interlacing of the FRET excitation and optical trapping beams in order to avoid rapid photo-damage of the FRET probes by the high power trapping

laser [9,10]. Moreover, the sensitivity of optical traps to mechanical drift makes the application of low stretching forces (< 1 pN) over extended periods of time far more challenging than with a simple magnetic tweezers system. Here we describe an integrated fluorescence and magnetic tweezers microscope capable of measuring nanometer scale structural transitions in single DNA molecules at low stretching forces. We demonstrate the utility of this approach by analyzing the mechanical unfolding pathway of a model human telomere DNA substrate.

Telomeres are specialized chromatin structures that protect linear ends of eukaryotic chromosomes from aberrant DNA processing by DNA damage repair machinery [11]. The foundation of human telomere structure is a long stretch of double-stranded DNA comprised of a hexa-nucleotide DNA repeat sequence (T_2AG_3). In addition, all telomeres terminate with a 3' single-stranded G-rich DNA tail which has the capacity to fold into a unique secondary structure called a G-quadruplex (GQ). Human telomere DNA GQs are proposed to play a central role in telomere homeostasis and small molecule ligands that selectively bind and stabilize telomere DNA GQs have shown promise as anti-cancer drugs [12,13]. Thus, intensive efforts have been made to better understand the structure and function of telomere DNA GQs.The first solution structure of a human telomere GQ revaled a fundamental structural architecture in which guanine bases are hydrogen bonded in a planar quartet geometry and may coordinate a single centrally located monovalent metal ion (Figure 1*A*, top left) [14].Three adjacent intra-molecular G-quartets may interact via stacking interactions and are topologically linked by short intervening DNA loop sequences (Figure 1*A*,

bottom left). Moreover, the folding properties of telomere DNA GQs vary with the presence of different monovalent cations. Na^+ ions predominantly promote the formation of an anti-parallel GQ conformation[14], whereas GQ DNA crystals formed in the presence of K^+ ions revealed a distinct parallel GQ folding topology[15]. More recent solution studies have demonstrated that multiple GQ topologies coexist in the presence of K^+, including the anti-parallel, parallel, and several hybrid forms [16-19]. The structural and dynamic properties of telomere DNA GQs have been studied using smFRET[20-22]. These experiments revealed that a particular GQ topological fold must transit through an obligatory unfolded intermediate in order to isomerize into a distinct GQ fold. More recently, the rupture force distribution of single telomere DNA GQs has been analyzed using atomic force microscopy and optical trapping, providing a direct measurement of telomere DNA GQ mechanical stability [23-26]; however, these force spectroscopy studies did not analyze the force dependence of telomere DNA GQ folding/unfolding at equilibrium. Here, we report a detailed analysis of the force-dependence of telomere DNA GQ folding and unfolding at equilibrium under Na^+ folding conditions. Our results demonstrate that the structural equilibrium between the unfolded and Na^+-induced telomere DNA GQ structure is highly sensitive to forces between ~1-8 pN. Analysis of the force-dependent rate constants for folding and unfolding provide a direct measurement of the position of the transition state barrier for telomere DNA GQ unfolding along the DNA stretching reaction coordinate. Interestingly, in contrast to other DNA secondary structures characterized by force spectroscopy, we find telomere DNA GQs exhibit a very short

distance (< 1 nm) to the transition state barrier for unfolding. This unfolding behavior indicates telomere DNA GQ structure is significantly stabilized by long range contacts, and once these contacts are disrupted the entire GQ fold readily dissolves.

MATERIAL AND METHODS

DNA Oligonucleotides

All DNA oligonucleotides were purchased from Integrated DNA Technologies, Inc. The sequences of all DNAs used in the study are listed in Supplementary Table 1.

Dye labeling of DNA oligonucleotides

The biotin DNA Handle was labeled at the amino modification on C6 of T3 with mono-reactive Cy5 (GE Healthcare) and the EcoRI DNA Handle was labeled at the amino modification on C6 of T16 with mono-reactive Cy3 (GE Healthcare). Dye labeled fragments were EtOH precipitated and purified by reverse-phase chromatography using a C8 column (Agilent, Eclipse XDB-C8) on an AKTA purifier. Following HPLC purification samples were EtOH precipitated, and resuspended in ddH_2O. DNA concentrations were determined using a Nanodrop.

DNA annealing reactions

The biotinylated,Cy5-labeled, and Cy3-labeled strands were annealed to either the Tel24 abasic, 15R60/T8 hairpin abasic, or polyT abasic containing sequence by heating to 95 °C for 4 minutes followed by slow cooling to room temperature (over several hours) in the presence of a buffer (50mM Tris-HCl pH 8) containing either 100mM NaCl or 100mM KCl.

DNA molecule for integrated fluorescence and magnetic tweezers measurements

DNA molecules were constructed by ligation of three precursor DNA fragments: the annealed biotinylated smFRET fragment, a digoxigenin modified DNA linker fragment, and lambda DNA handle. The preparation of desired Cy5 and Cy3 labeled DNA insert is described above. The digoxigenin-labeled DNA linker is synthesized by PCR reaction with a pUC19 template using primers flanking the multiple cloning site. PCR reaction were set up using a dNTP mixture containing a 1:4 molar ratio of Digoxigenin-11-dUTP (Roche) : dTTP and digested with BamHI. The 15721 base pair lambda DNA handle was prepared by BamHI and EcoRI enzymatic digestion and agarose gel purification. Ligation reactions were set up with a 1:1:1 molar ratio of the desired Cy5 and Cy3 labeled DNA insert harboring an EcoRI sticky end, the purified EcoRI/BamHI lambda DNA handle, and the BamHI digoxigenin modified DNA linker fragment. Ligation reactions were run overnight at 16^{o}C in the presence of T4 DNA ligase (NEB). The final ligation product was heat inactivated at 65^{o} C to ensure the T4 DNA ligase did not remain associated with the DNA.

Single molecule FRET measurements in the absence of force

For these experiments, quartz slides (Finkenbeiner Inc.) were cleaned by sonicating for 20 minutes in 10% w/v Alconox, 20 minutes in ddH_2O, 20 minutes in acetone, 20 minutes 1M KOH, then 20 minutes in fresh 1M KOH. Slides are then rinsed with ddH2O and dried under nitrogen, followed by flame cleaning with a propane torch for ~2 minutes. Sample chambers were prepared by sandwiching pieces of parafilm between the quartz slide and a plasma cleaned coverglass (Harrik Plasma Cleaner), cover glasses from TED Pella, Inc. (Prod.# 260146). Chambers were heated on a hot

block (95^{o}C) for 1 minute to seal the parafilm to the glass. Channels (~10 uL total volume) were treated with 35 uL of 1 mg/mL of biotinylated BSA (Sigma Cat.#A8549) for 5 minutes, washed with 100 uL of T50 buffer (10mM Tris-HCl pH 8, 50mM NaCl), then incubated with 50uL of 0.2 mg/mL streptavidin (Invitrogen Cat. #S888), then washed with 100 uL of T50. Channels were then equilibrated with buffer matching the desired experimental condition (see main text).

Next, 100 uL of 5-10 pM annealed fluorescent and biotin-labeled smFRET DNA was deposited onto the streptavidin coated quartz slide in the desired buffer condition. After 5 minutes incubation, desired buffer was flushed to the slide to remove the unbound DNA. Data was collected in imaging buffer containing 50mM Tris pH 8, 100 mM KCl or NaCl, 0.4% (w/v) D-glucose, 0.1 mg/ml glucose oxidase (Sigma Cat.# G2133-250KU), 0.02 mg/ml catalase (CalBiochem Cat.#219001) and saturated with Trolox (Aldrich Cat.#23,881-3). Data was acquired using a green laser (532 nm, Laserglow, Inc.) and prism-type total internal reflection microscopy on an inverted Olympus IX71 microscope equipped with an Andor IXON (897) CCD camera with 100 ms integration time.

Single molecule FRET analysis

Raw movie files were analyzed using in house written software available upon request (IDL and Matlab). FRET is defined as the efficiency of energy transfer between acceptor and donor dye, $I_A/(I_A+I_D)$. I_A is the acceptor intensity and I_D is the donor intensity. FRET histograms were compiled by combining data from ~ 100 individual single molecule trajectories. The Hidden Markov modeling was performed

using the HaMMy[27] software program. First, dye intensity traces used in the HaMMy analysis were normalized after photo-bleaching, then the period prior to photo-bleaching was cropped to be used in the analysis. Since individual traces did not typically possess a sufficient number of transitions to produce a reliable HMM fit, we opted to stitch all of the normalized and cropped trajectories into a single trace using in house written Matlab software. This compiled trace was then fit using the HaMMy software. The idealized FRET trace that was produced by the HaMMy fitting was then parsed back into individual traces using in house Matlab software to prevent the inclusion of artificial FRET transitions at the trace stitch points. The individual HaMMy fits were then used to generate a transition density plot using in house written Matlab software. To fit the smFRET distributions, data were binned (bin size was FRET = 0.03) and fit with multiple Gaussian functions using Origin software. For dwell time analysis, idealized traces produced by fitting the single molecule trajectories using the HaMMy software package were used to calculate the average lifetime of the folded or unfolded states. The smFRET histograms for the Na^+ and K^+ titrations, as well as in the presence of Mg^{2+}, were constructed by averaging the observed FRET for each molecule over two seconds (20 frames at 100 ms integration time) and binning the data with a bin size of FRET = 0.03.

Integrated FRET and Magnetic Tweezers measurements

The DNA tethers for FRET-Magnetic Tweezers measurements were attached to anti-dig coated magnetic beads. Anti-digoxigenin beads were made using EDC coupling chemistry (Pierce CAS# 22980) using carboxylated magnetic beads (Dynal, either

1μm (Invitrogen Cat.#650.11) or 2.8 μm (Invitrogen Cat.#143.05D) diameter. After 10 minutes incubation at 4^{o}C, the DNA-magnetic bead mixture was washed to remove the excess DNA fragments prior to being deposited onto the streptavidin coated coverglass (#1.5, Ted Pella) prepared using the same cleaning procedure described above. After 30 minutes incubation in room temperature to allow the DNA-magnetic bead complexes to settle to the surface, 100uL of 50mM Tris with 100mM Na^+ (for GQ experiments) or 200mM K^+ (for control DNA hairpin experiments) was flowed onto the slide by gravity to remove the unattached DNA linked magnetic beads.

To calibrate the stretching forces being applied to each DNA tether, forces were measured across a wide range of positions of the magnet assembly held above the sample chamber. Force measurements were determined using the expression $F = Lk_BT/\langle x^2 \rangle$, where F is the force in pico-Newtons (pN), L is the DNA tether length at a particular magnet height, k_BT is thermal energy (4.1pN*nm) at room temperature, and $\langle x^2 \rangle$ is the variance in the bead position in x-axis. The DNA tether length was measured by calibrating the diffraction ring pattern of the magnetic bead using a piezo-controlled objective positioning device (Mad City Labs) as previously described [28]. Separate calibrations were made for the 1μm and 2.8μm magnetic beads. The force was found to decrease exponentially as the distance of the magnets was increased from the sample chamber as described[28]. The magnet height vs. Force plots were fit with single exponential decay functions and the parameters from these fits were then used to calculate the forces applied to subsequent experimental setups

using the same size bead. The bead manufacturer estimates a deviation in the radius (r) of the beads on the order of 1-2%, and the force scales with the volume of the bead (r^3), thus we estimate the systematic error of our force measurements based upon an imperfect calibration to be on the order of 10%.

Data was collected in imaging buffer containing 100mM Na^+ or 200mM K^+ salt, 0.4% (w/v) D-glucose, 0.1 mg/ml glucose oxidase (Sigma Cat.# G2133-250KU), 0.02 mg/ml catalase (CalBiochem Cat.#219001) and saturated with Trolox (Aldrich Cat.#23,881-3). Data was acquired using objective-type total internal reflection microscopy with a green laser (532 nm, Laserglow, Inc.) on an Olympus IX71 microscope equipped with an Andor IXON (860) CCD camera. 100ms integration time was used for telomere DNA GQ and polyT measurements and 33 ms integration time was used for hairpin measurement. The stretching force applied to individual DNA tethers was altered by changing the linear position of a pair of rare earth magnets mounted on a computer controlled translation stage above the sample. The DNA extension and applied stretching force were determined by visualizing the magnetic bead with a blue LED (Thorlabs) mounted above the magnet assembly as described previously [28]. Prior to collecting smFRET data, the length of each DNA tether was monitored in real time as a function of applied twist to ensure that there was only a single DNA molecule attached between the bead and the surface. Molecules whose extension changed as a function of twist represented DNA braids and were not used in the measurements.

Raw movie files were first analyzed using in house written software (LabVIEW and Matlab), and then fit with the HaMMy[27] software to generate idealized FRET traces used for dwell time analysis. Single molecule FRET histograms generated from combined FRET-magnetic tweezers data were compiled from all molecules analyzed at a particular force set point. Traces were first normalized for photo-bleaching and then cropped prior to the photo-bleaching event for HaMMy analysis. We focused our analysis on the equilibrium between the predominant mid-FRET (folded) and the low-FRET (unfolded) state. To this end, for our analysis we used the idealized FRET traces generated by the HaMMy [27] fitting to remove the minority of time spent in the alternative high-FRET (folded) state. For dwell time analysis, idealized traces produced by fitting the smFRET trajectories using the HaMMy software package were used to calculate the average lifetime of the folded or unfolded states at each of the indicated forces.

RESULTS

Single molecule FRET analysis of telomere DNA GQ structure in the absence of force

The experimental geometry utilized in the present study of the force-dependent folding/unfolding of telomere DNA GQ structure includes attachment of duplex DNA handles to each end of a single GQ forming sequence (Figure 1A, right). A similar design has been employed in recent optical trapping experiments [24-26], but it has not been analyzed using smFRET; therefore, we first set out to characterize the structure and dynamics of our model human single stranded telomere DNA construct

(Tel24, $(TTAGGG)_4$) embedded within a duplex DNA molecule. For these experiments the Tel24 sequence was flanked by non-telomeric DNA extensions which were hybridized to their respective complementary DNA strands harboring either a FRET donor (Cy3) or acceptor (Cy5) dye (Figure. 1A, right). To minimize the possibility of the adjacent duplex altering the folding properties of the GQ forming sequence, abasic sites were introduced between the Tel24 and the non-telomere sequence. Experiments investigating the structure of Tel24 in the absence of force were conducted on a prism-type total internal reflection fluorescence (TIRF) microscope [29]. Throughout this study we define FRET = $I_A/(I_A + I_D)$, where I_A and I_D are the background corrected intensities of the acceptor and donor dyes, respectively.

Figure 1

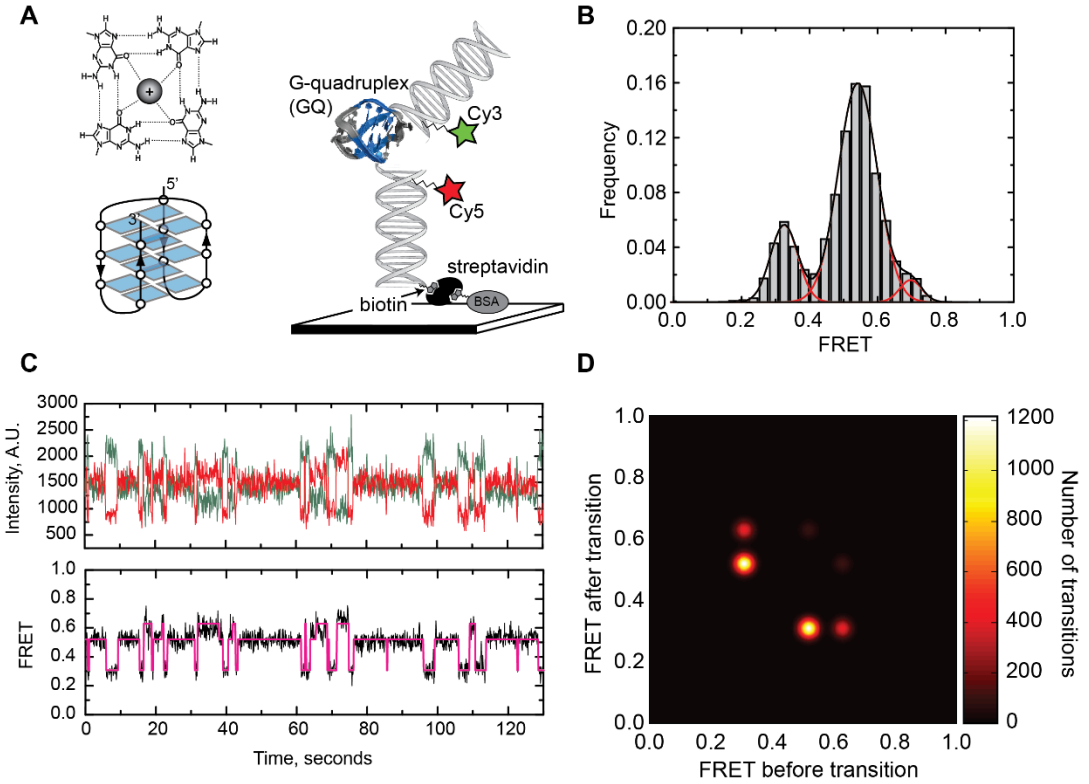


Figure 1. Single molecule FRET analysis of Na^+-induced telomere DNA G-quadruplex folding in the absence of force. (A) (top left) Diagram of H-bonding network within a single G-quartet with a monovalent metal ion coordinated at its center. (bottom left) Schematic illustration of the anti-parallel G-quadruplex folding topology with each planar G-quartet represented as a quartet of blue rectangles. (right) Schematic diagram of smFRET experimental setup. (A model telomere DNA sequence ($(TTAGGG)_4$, Tel24) with G-quartets highlighted in blue is flanked by two duplex DNA handles harboring a FRET donor (Cy3) and acceptor (Cy5) dye, respectively. Tel24 molecules were surface immobilized via a biotin-streptavidin linkage and imaged using prism-type total internal reflection microscopy. The antiparallel G-quadruplex structure solved by NMR is shown in the illustration (PDB code 143D). (B) Single molecule FRET histogram derived from data collected on >100 molecules is fit with three Gaussian functions centered at FRET = 0.33, 0.54, and 0.70. (C) A representative smFRET trajectory for Tel24 in the presence of 100 mM NaCl. (top) Individual donor (green) and acceptor (red) dye intensities, and (bottom) calculated FRET ratio are plotted as a function of time. FRET trajectories were fit with a Hidden Markov model yielding idealized FRET traces (magenta). (D) Idealized FRET traces from HMM fitting were used to generate a transition density plot indicating the frequency of each observed transition. Color bar on right indicates the number of transitions.

We first measured the folding properties of Tel24 over a wide range of NaCl concentrations (Supplementary Figure 1). As the NaCl concentration was gradually increased to 100 mM, we observed the emergence of a predominant FRET population centered at FRET = 0.54, a second population centered at FRET = 0.33, and a minor population centered at FRET = 0.70 (Figure 1B). Analysis of single molecule trajectories revealed the Tel24 construct is in a dynamic equilibrium between each of these three FRET states (Figure 1C). Conversion of the predominant mid-FRET state to the high-FRET state required formation of the transient low-FRET state, as has been previously reported [21]. Single molecule trajectories were fit to a Hidden Markov model which yielded idealized FRET trajectories (Figure 1C) [27]. The idealized traces were used to generate transition density plots which clearly indicated the majority of FRET transitions (70%, n = 2426/3475 transitions) occurred between the mid-FRET

and the low-FRET states (Figure 1D). The ratio of time spent in the unfolded folded (FRET = 0.33) and the predominant folded (FRET = 0.54) states yielded a $K_{eq} = 0.25$, corresponding to a $\Delta G_{unfold} = 0.81$ kcal mol^{-1}, a value that is in good agreement with previously reported calorimetric studies of Tel24 unfolding in NaCl [30]. Dwell time analysis yielded rate constants of 0.87 s^{-1} and 0.29 s^{-1} for the folding and unfolding reactions, respectively, giving a $K_{eq} = 0.33$ in reasonable agreement with the value obtained by fitting the smFRET histograms. NMR and CD spectroscopy experiments on telomere DNA substrates in the presence of NaCl have demonstrated the strong preference of single stranded telomere DNA to form the anti-parrallel GQ conformation under this folding condition [14,31]. Thus, while smFRET data alone are not sufficient to conclusively determine which GQ structure is present, it is likely the predominant FRET = 0.54 conformation observed in the presence of 100 mM Na^+ represents the anti-parallel conformation and the FRET = 0.33 state represents the unfolded conformation.

In contrast to the Na^+ folding condition, the presence of KCl produced a considerably broader distribution of FRET states, with increasing concentrations of K^+ stabilizing several distinct high FRET GQ conformations (Supplementary Figure 2). The smFRET distribution for Tel24 folded in 100 mM KCl is well fit by two Gaussian functions, with populations centered at FRET = 0.61 and FRET = 0.76. Consistent with this result, analysis of single molecule trajectories in the presence of 100 mM KCl revealed two distinct long-lived FRET states, which occasionally inter-converted via the obligatory low FRET intermediate (Supplementary Figure 3). In

addition, the presence of Mg^{2+} in the K^+ folding reaction appeared to have a slightly stabilizing effect on GQ structure, but did not significantly alter the smFRET histograms in the presence of 100 mM K^+ (Supplementary Figure 4). Taken together, our smFRET results are in good agreement with previous findings that demonstrated Na^+ ions promote homogeneous folding of telomere DNA GQs, while K^+ ions induced a more complex distribution of FRET states that are more thermodynamically stable than the Na^+-induced fold. Based upon these results, we conclude that the presence of duplex DNA flanking the Tel24 sequence does not significantly alter telomere DNA GQ folding and stability.

Integrated Fluorescence and Magnetic Tweezers Spectroscopy

Previously reported force spectroscopy experiments using optical traps achieved sub-nanometer spatial resolution by applying relatively large stretching forces (>20 pN) to the system of study [32,33]. Since many biologically important structural transitions in nucleic acids and proteins are induced by much smaller forces, we have developed a simple method for measuring sub-nanometer structural rearrangements within individual DNA molecules held under a wide range of stretching forces (0.1 – 50pN). In our integrated fluorescence and magnetic tweezers system, a single Tel24 sequence is embedded within a duplex DNA molecule specifically attached between a microscope slide and a magnetic bead (Figure 2A). The length of the DNA handle attached to the microscope slide was 29 base pairs (~10 nm), while the length of the DNA handle attached to the bead was made considerably longer (15.7 kb, 5.4 μm contour length) to prevent background signal

introduced by the large magnetic bead entering the evanescent field produced near the surface of the glass slide by total internal reflection of the excitation laser. The DNA construct is further modified with FRET donor (Cy3) and acceptor (Cy5) dyes on the DNA handles flanking the Tel24 sequence, as well as abasic sites between the duplex DNA handles and the Tel24 sequence. Stretching force was applied to the tethered DNA molecule with a pair of permanent rare-earth magnets mounted on a computer-controlled translation stage. Integrated fluorescence and magnetic tweezers experiments employed objective-type TIRF which is readily compatible with the presence of the magnet assembly held above the sample chamber. We note that due to slight differences in the optical components in the microscopes, smFRET values made by objective-type TIRF vary slightly from those measured using the prism-type TIRF setup described in Figure 1 (ie. mid-FRET (folded state) is 0.64 in low force objective-type TIRF experiments vs. 0.54 in prism-type TIRF microscope under zero force); however, this variation does not impact our assignment of the mid-FRET and low-FRET states to the folded and unfolded telomere DNA GQ conformations, respectively.

Figure 2

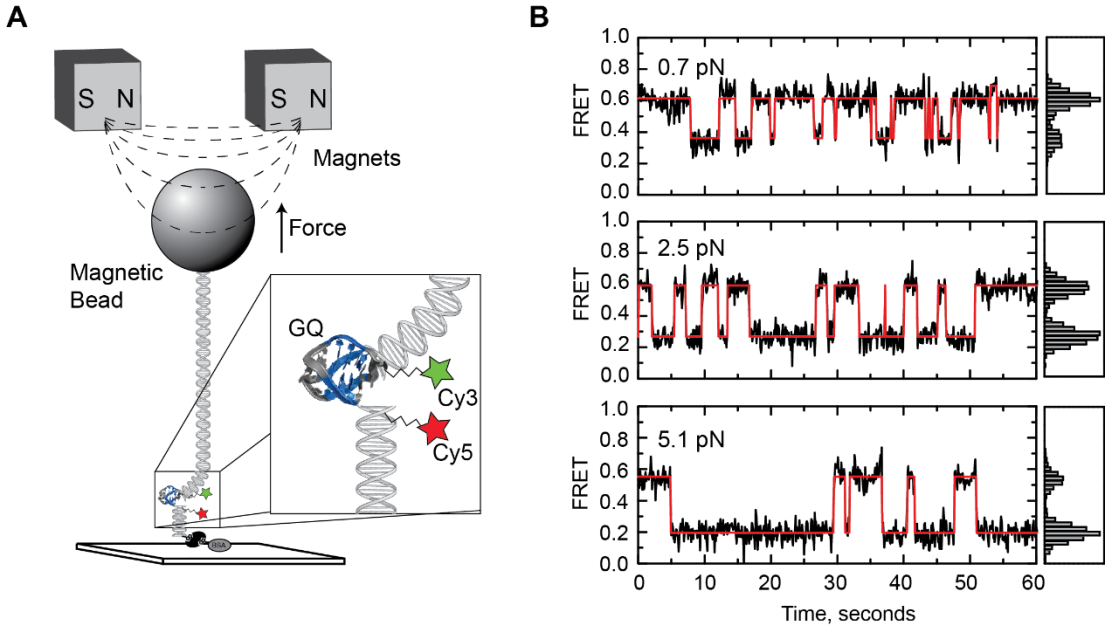


Figure 2. Force biases the telomere DNA G-quadruplex folding/unfolding equilibrium. (A) Schematic diagram of the experimental setup for integrated fluorescence and magnetic tweezers measurements. A FRET-labeled Tel24 molecule is flanked by a short biotinylated duplex DNA handle on one side, and a second longer duplex DNA handle with terminal digoxigenin modifications on the other side. Individual Tel24 constructs were attached between a streptavidin coated microscope slide and an anti-digoxigenin coated magnetic bead in a magnetic tweezers setup and imaged using objective-type total internal reflection microscopy. Variable stretching force was applied to individual Tel24 molecules by translating a pair of rare earth magnets held above the sample chamber. (B) Representative smFRET trajectories of a single Tel24 molecule held under three different forces (0.7, 2.5, and 5.1 pN). Histogram of the FRET values obtained at each force set point are shown on the right side of the FRET trajectories.

In Figure 2B, representative smFRET trajectories are plotted at three different stretching forces for a single telomere DNA GQ in the presence of 100 mM NaCl. It is evident that at low forces (0.7 pN) the molecule spends the majority of the time in the high FRET (folded) state consistent with our zero force measurements described above. In contrast, a higher force (5.1 pN) substantially shifts the telomere DNA GQ

folding equilibrium toward the low FRET (unfolded) state. Notably, the application of small degrees of tension appeared to substantially reduce the number of transitions into the minority high-FRET (FRET = 0.70) conformation when compared to experiments performed in the absence of force; thus, we focused our analysis on the structural equilibrium between the low-FRET (unfolded) and predominant mid-FRET (folded) conformation. These data qualitatively demonstrate the capacity of our technique to detect nanometer scale structural transitions within single telomere DNA GQs at low forces. However, to determine whether this approach can be used to extract quantitative information about force-induced structural transitions in DNA, we next performed control experiments on a model DNA hairpin construct (Supplementary Figure 5). Data taken on the model DNA hairpin was analyzed as a simple two state system separated by a single energetic barrier, yielding an unfolding force $F_{1/2}$ = 12.1 pN (defined as the force at which the hairpin folded and unfolded states are equally populated) and a distance between the folded and unfolded state, Δx = 16.2 nm. Furthermore, analysis of the force-dependent rate constants for folding and unfolding placed the transition state barrier 8.9 nm from the folded state and 6.1 nm from the unfolded state. These results are in good quantitative agreement with previously reported optical-trapping measurements on the same model DNA hairpin [2], demonstrating the ability of our integrated fluorescence and magnetic tweezers system to extract quantitative information about DNA folding processes.

Force dependence of telomere DNA GQ folding and unfolding

We next characterized in greater detail the force-dependence of the telomere DNA GQ folding/unfolding equilibrium. For these experiments, we performed force titrations for 75 individual telomere DNA GQ molecules (for each force a minimum of five different molecules were analyzed), and compiled the results into FRET histograms (Figure 3A, see Supplementary Figure 6 for complete data set). Analysis of the amount of time spent in the mid-FRET (folded) state vs. the low-FRET (unfolded) state as a function of force provided a direct measurement of the effect of stretching force on the telomere DNA GQ folding equilibrium. We fit each histogram with two Gaussian functions and calculated K_{eq} for the unfolding reaction as the area under the low-FRET peak divided by the area under the mid-FRET peak. From the slope of the plot of ln(Keq) as a function of stretching force, we determined the unfolding distance $\Delta x = 2.7$ nm and $F_{1/2} = 2.5$ pN (Figure 3B).

Figure 3

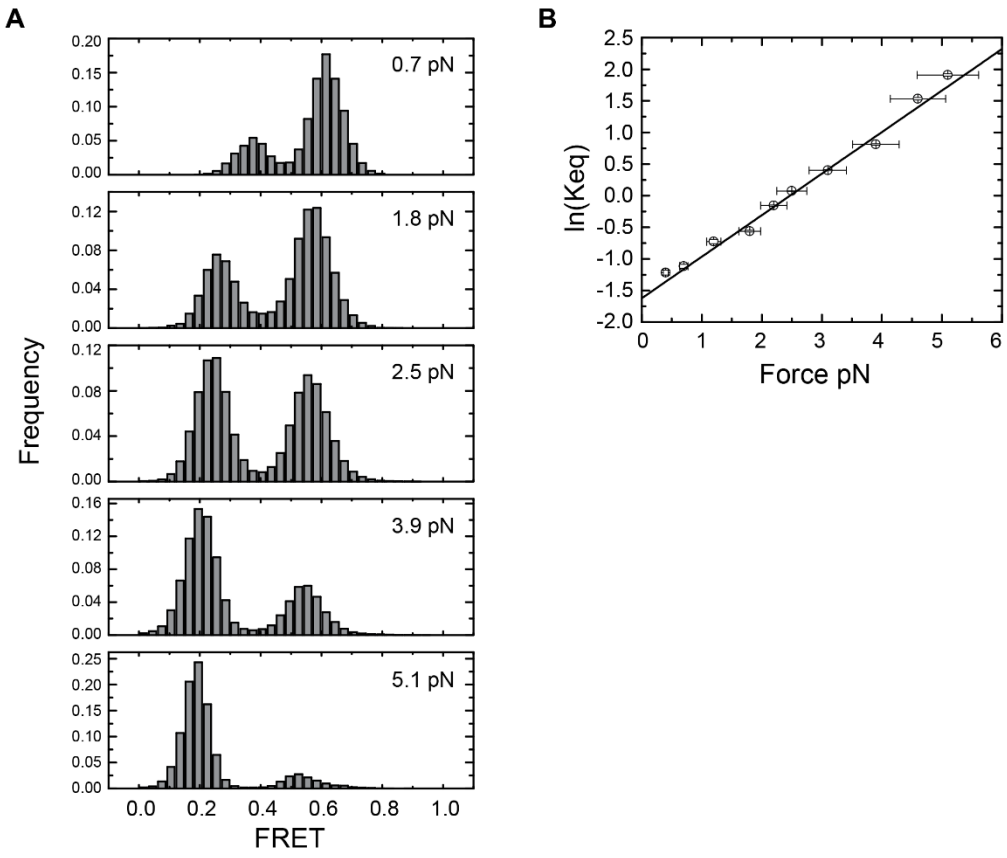


Figure 3. Telomere DNA G-quadruplex folding equilibrium as a function of applied force. (A) Compiled single molecule FRET histograms from data taken on 75 individual Tel24 molecules. For each force, a minimum of five different molecules were analyzed. Data for five representative forces ranging from 0.7 to 5.1 pN are shown. (see Supplementary Figure 6 for complete data set). (B) Plot of the $\ln(K_{eq})$ vs. Force. For each force set point at which folding/unfolding transitions were detected, smFRET histograms were fit with two Gaussian functions and Keq for unfolding was calculated as the area under the low FRET (unfolded) state divided by the area under the high FRET (folded) state. Plot was fit with the expression $\ln(K_{eq}) = F\Delta x/k_BT - \Delta G^o/k_BT$, yielding an unfolding distance, $\Delta x = 2.7$ nm and a $F_{1/2} = 2.5$ pN. Y-axis error bars are estimates of the standard error of the mean determined by bootstrapping analysis of smFRET data and x-axis error bars reflect an ~10% systematic error in determination of stretching force.

We observed that the centers of the Tel24 low-FRET (unfolded) and mid-FRET (folded) populations shifted toward lower values as the applied stretching force was increased (Figure 3A). The two different Tel24 conformations fall within significantly different regions of the FRET response curve; thus, the magnitude of the force-induced ΔFRET for the low-FRET population reflects a substantially larger change in distance between the FRET probes than the ΔFRET of the mid-FRET population (Supplementary Figure 7). The low-FRET Tel24 conformation, previously identified as an obligatory folding intermediate during inter-conversion between distinct topological isomers[21], was suggested to reflect an unfolded DNA conformation that would be expected to respond to stretching force by gradually increasing its end-to-end distance, as observed in our experiments. To further explore the structural properties of this obligatory unfolded intermediate, we compared the force response of the low-FRET (unfolded) telomere DNA GQ conformation to a 24 nucleotide polyT DNA strand, which has been utilized as a model for unstructured single stranded DNA in previous smFRET experiments [34]. We find that under low stretching force conditions, the unfolded conformation of the Tel24 construct is significantly more compact (higher FRET) than the polyT construct (Figure 4). This result demonstrates the G-rich Tel24 unfolded state experiences an attractive force, perhaps derived from increased stacking interactions between purine bases or transient H-bonding, which promotes telomere DNA GQ folding. Interestingly, the elastic properties of the unfolded Tel24 and polyT constructs converge within the force range (~8 pN) that completely inhibits telomere DNA GQ folding within the

100 msec time resolution of these measurements (Supplementary Figure 6). We note that the gradual increase in extension of the unfolded Tel24 state with increasing force is not consistent with the presence of any long-lived structured GQ folding intermediates such as a DNA hairpin or triplex structure. These structured intermediates have been previously suggested to occur during telomere DNA GQ folding[35-37], but may be too transient to detect in our experiments.

Figure 4

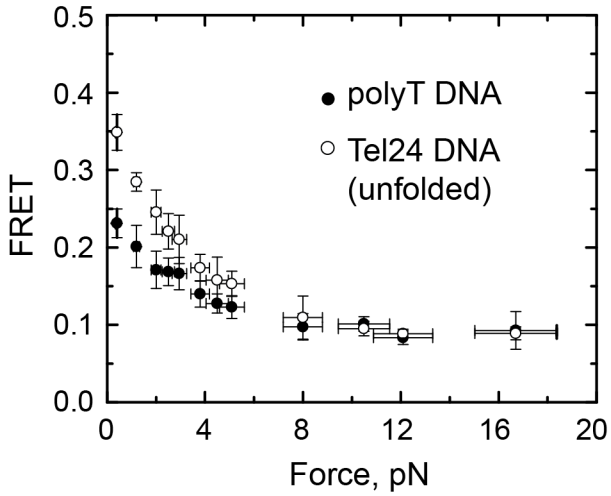


Figure 4. Force response of Tel24 low FRET (unfolded) state and polyT DNA. (open circles)The average FRET values of the Tel24 low FRET (unfolded) state (open circles) and a 24 nucleotide polyT DNA construct (closed circles) are plotted as a function of force. Y-axis error bars are the standard deviation of the average FRET value measured for at least 5 individual molecules at each force set point and x-axis error bars reflect ~10% systematic error in the calibration of the applied stretching force.

The force-dependence of the rate constants for telomere DNA GQ folding and unfolding can also be analyzed from our single molecule trajectories. Making the simplifying assumptions that along the reaction coordinate defined by the DNA

stretching axis, the position of the energy barrier is independent of the stretching force and the positions of the folded and unfolded energy wells are not substantially altered by the applied force, we modeled the lifetime of the folded state, τ_{folded}, as given by $\tau_{folded}(F) = \tau_{folded,0}*\exp(F\Delta x^{\ddagger}_{folded}/k_BT)$, where $\tau_{folded,0}$ is the folded lifetime at zero force, $\Delta x^{\ddagger}_{folded}$ is the distance from the folded to the transition state, k_B is Boltzman's constant, and T is temperature. Similarly, the lifetime of the unfolded state as a function of force can be expressed as $\tau_{unfolded}(F) = \tau_{unfolded,0}*\exp(F\Delta x^{\ddagger}_{unfolded}/k_BT)$. Fitting these expressions to a plot of the average lifetimes of the folded or unfolded states as a function of force yielded $\Delta x^{\ddagger}_{unfolded}$ = 2.1 nm, $\Delta x^{\ddagger}_{folded}$ = 0.6 nm, $\tau_{folded,0}$ =4.1 sec, and $\tau_{folded,0}$ = 0.76 sec (Figure 5A). The results of our kinetic analysis are in good agreement with the Δx value derived from the equilibrium data ($\Delta x = \Delta x^{\ddagger}_{folded} + \Delta x^{\ddagger}_{unfolded}$ = 2.7nm). In addition, the extrapolated zero force rate constants of k_{fold} = 1.3 s^{-1} and k_{unfold} = 0.24 s^{-1} are in good agreement with our smFRET measurements in the absence of force (see above). The short distance (0.6 nm) from the folded state to the transition state barrier suggests the telomere DNA GQ structure is stabilized by long-range contacts, and that disruption of just a few of the terminal base pairing contacts is sufficient to promote the unfolding of GQ structure (Figure 5B).

Figure 5

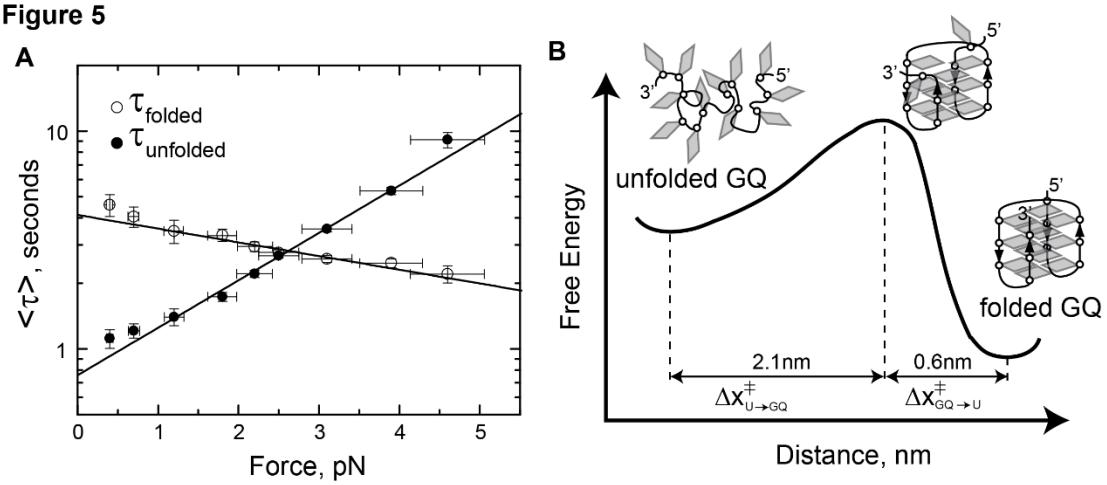


Figure 5. Force-dependent rate constants of telomere DNA G-quadruplex folding and unfolding. (A) Average lifetimes (τ) of the folded (high FRET) and unfolded (low FRET) states are plotted as a function of force. Plots are fit with exponential functions described in the main text, yielding $\Delta x^{\ddagger}_{unfolded}$ = 2.1 nm and $\Delta x^{\ddagger}_{folded}$ = 0.6 nm. (B) A schematic folding energy landscape for Tel24 in the presence of 100 mM NaCl. Cartoon model for the folded, transition state, and unfolded structures are superimposed on the folding energy landscape. The short distance between the transition state barrier and the folded conformation suggests just a few base pairs are disrupted in the transition state structure.

DISCUSSION

Here, we describe an integrated fluorescence and magnetic tweezers technique that combines objective-type total internal reflection microscopy with a 'classical' magnetic tweezers setup [28,38]. Although several groups have described combined fluorescence-force instruments using optical traps to apply stretching forces [6,9], these tools remain inaccessible to non-specialists, due to the technical challenges associated with aligning and calibrating these advanced optical trapping systems. Furthermore, the integration of a high-power trapping laser with single molecule fluorescence detection has the potential to negatively impact the photo-physical stability of the fluorescence probes, thereby limiting the duration of the FRET measurement. By

comparison, the technique described in the present study is a simple instrument to construct and maintain, and should therefore be a generally accessible tool for a wide range of biochemists and structural biologists. Moreover, the use of permanent magnets to apply stretching forces obviates concerns of photo-damage induced by a high power trapping laser and provides a means to stably apply low forces for arbitrarily long periods of time. Indeed, smFRET trajectories acquired using our approach routinely last for tens of minutes, facilitating measurement of a single molecule under a variety of experimental conditions. An additional advantage of hybrid force-fluorescence methods using magnetic tweezers is the ability to monitor torque-induced structural transitions by smFRET, as has recently been reported in studies of the torque-induced B-DNA to Z-DNA transition [7]. We expect the capability of integrated fluorescence and magnetic tweezers systems to detect nanometer scale structural rearrangements under conditions of low tensions and torques will be broadly applicable to a wide range of macromolecular folding studies.

In the present work, we have analyzed the force-dependent folding of a model human telomere DNA G-quadruplex (GQ) in the presence of Na^+. We find that the dynamic equilibrium between the Na^+-induced GQ folded and unfolded states is extremely sensitive to applied stretching forces between ~1-8 pN. By analyzing the effect of stretching force on the rate constants for telomere DNA GQ folding and unfolding, we have directly measured the position of the transition state barrier for GQ unfolding along the well defined DNA stretching reaction coordinate. Interestingly, we find the distance between the telomere DNA GQ folded state and

the transition state barrier for unfolding is quite short (~0.6 nm). Our finding is consistent with a recent report that used dynamic force spectroscopy (DFS) to reconstruct the distribution of rupture forces for a DNA GQ formed by a sequence upstream of the human insulin gene promoter [26]. Fitting the DFS results to theoretical models which predict the distribution of rupture forces assuming a single transition state barrier suggested the barrier for folding is within ~1nm of the GQ folded state, in close accord with our present results. However, as with previous optical trapping experiments which analyzed the mechanical stability of GQ DNA[24,25], the DFS measurements did not directly observe a reversible folding/unfolding equilibrium in the low force regime as reported here. At present, it is unclear why the low force 'hopping' behavior we report for the telomere DNA GQ in the present study is not observed in the optical trapping experiments, but the discrepancy may be due to differences in the sequences being used or challenges associated with detecting small extension changes at low stretching forces in an optical tweezers instrument.

Force spectroscopy experiments with RNA and DNA hairpins revealed the transition state barrier for unfolding lies closer to the unfolded state than the folded state[1,2]. This long distance to the transition state barrier for unfolding is expected for compliant structures whose folding stability is derived primarily from local interactions (ie. zipping of adjacent base pairs within a DNA hairpin). However, a short distance to the transition state barrier for unfolding, as measured here for the telomere DNA GQ, is indicative of a more brittle structure whose stability is reliant upon critical contributions of long-range interactions. This result can be explained

within the framework of the NMR structure for the Na^+- induced anti-parallel GQ fold[14], wherein guanine bases separated by as many as 20 nucleotides must be brought together within a planar G-quartet configuration. Thus, disruption of just a few H-bonding interactions by mechanically pulling on the terminal bases of the telomere DNA GQ structure is sufficient to effectively destabilize the entire fold (Figure 5B), a finding that has direct implications on the molecular mechanisms of telomere-associated proteins and enzymes which must resolve DNA GQs during telomere maintenance. Telomere DNA GQs are typically classified as DNA secondary structure; however, our studies reveal these unique DNA folds share important structural properties with RNA and protein tertiary structures, which have also been characterized by sub-nanometer distances to the transition state barrier for unfolding[1,39-41]. Finally, understanding the critical contribution of long-range interactions in promoting telomere GQ folding stability should aid efforts in designing small molecule drugs that target and stabilize GQ structure in order to disrupt telomere homeostasis.

REFERENCES

1 Liphardt, J., Onoa, B., Smith, S. B., Tinoco, I., Jr. & Bustamante, C. Reversible unfolding of single RNA molecules by mechanical force. *Science* **292**, 733-737 (2001).

2 Woodside, M. T. *et al.* Nanomechanical measurements of the sequence-dependent folding landscapes of single nucleic acid hairpins. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 6190-6195 (2006).

3 Harlepp, S. *et al.* Probing complex RNA structures by mechanical force. *The European physical journal. E, Soft matter* **12**, 605-615 (2003).

4 Li, P. T., Collin, D., Smith, S. B., Bustamante, C. & Tinoco, I., Jr. Probing the mechanical folding kinetics of TAR RNA by hopping, force-jump, and force-ramp methods. *Biophysical journal* **90**, 250-260 (2006).

5 Onoa, B. *et al.* Identifying kinetic barriers to mechanical unfolding of the T. thermophila ribozyme. *Science* **299**, 1892-1895 (2003).

6 Hohng, S. *et al.* Fluorescence-force spectroscopy maps two-dimensional reaction landscape of the holliday junction. *Science* **318**, 279-283 (2007).

7 Lee, M., Kim, S. H. & Hong, S. C. Minute negative superhelicity is sufficient to induce the B-Z transition in the presence of low tension. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 4985-4990 (2010).

8 Shroff, H. *et al.* Biocompatible force sensor with optical readout and dimensions of 6 nm3. *Nano letters* **5**, 1509-1514 (2005).

9 Brau, R. R., Tarsa, P. B., Ferrer, J. M., Lee, P. & Lang, M. J. Interlaced optical force-fluorescence measurements for single molecule biophysics. *Biophysical journal* **91**, 1069-1077 (2006).

10 Tarsa, P. B. *et al.* Detecting force-induced molecular transitions with fluorescence resonant energy transfer. *Angewandte Chemie* **46**, 1999-2001 (2007).

11 Palm, W. & de Lange, T. How shelterin protects mammalian telomeres. *Annual review of genetics* **42**, 301-334 (2008).

12 Oganesian, L. & Bryan, T. M. Physiological relevance of telomeric G-quadruplex formation: a potential drug target. *BioEssays : news and reviews in molecular, cellular and developmental biology* **29**, 155-165 (2007).

13 Rezler, E. M., Bearss, D. J. & Hurley, L. H. Telomeres and telomerases as drug targets. *Current opinion in pharmacology* **2**, 415-423 (2002).

14 Wang, Y. & Patel, D. J. Solution structure of the human telomeric repeat d[AG3(T2AG3)3] G-tetraplex. *Structure* **1**, 263-282 (1993).

15 Parkinson, G. N., Lee, M. P. & Neidle, S. Crystal structure of parallel quadruplexes from human telomeric DNA. *Nature* **417**, 876-880 (2002).

16 Ambrus, A. *et al.* Human telomeric sequence forms a hybrid-type intramolecular G-quadruplex structure with mixed parallel/antiparallel strands in potassium solution. *Nucleic acids research* **34**, 2723-2735 (2006).

17 Dai, J., Carver, M. & Yang, D. Polymorphism of human telomeric quadruplex structures. *Biochimie* **90**, 1172-1183 (2008).

18 Li, J., Correia, J. J., Wang, L., Trent, J. O. & Chaires, J. B. Not so crystal clear: the structure of the human telomere G-quadruplex in solution differs from that present in a crystal. *Nucleic acids research* **33**, 4649-4659 (2005).

19 Luu, K. N., Phan, A. T., Kuryavyi, V., Lacroix, L. & Patel, D. J. Structure of the human telomere in K+ solution: an intramolecular (3 + 1) G-quadruplex scaffold. *Journal of the American Chemical Society* **128**, 9963-9970 (2006).

20 Jena, P. V. *et al.* G-quadruplex DNA bound by a synthetic ligand is highly dynamic. *Journal of the American Chemical Society* **131**, 12522-12523 (2009).

21 Lee, J. Y., Okumus, B., Kim, D. S. & Ha, T. Extreme conformational diversity in human telomeric DNA. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 18938-18943 (2005).

22 Ying, L., Green, J. J., Li, H., Klenerman, D. & Balasubramanian, S. Studies on the structure and dynamics of the human telomeric G quadruplex by single-molecule fluorescence resonance energy transfer. *Proceedings of the National Academy of Sciences of the United States of America* **100**, 14629-14634 (2003).

23 Lynch, S., Baker, H., Byker, S. G., Zhou, D. & Sinniah, K. Single molecule force spectroscopy on G-quadruplex DNA. *Chemistry* **15**, 8113-8116 (2009).

24 Koirala, D. *et al.* A single-molecule platform for investigation of interactions between G-quadruplexes and small-molecule ligands. *Nature chemistry* **3**, 782-787 (2011).

25 Yu, Z. *et al.* Click chemistry assisted single-molecule fingerprinting reveals a 3D biomolecular folding funnel. *Journal of the American Chemical Society* **134**, 12338-12341 (2012).

26 de Messieres, M., Chang, J. C., Brawn-Cinani, B. & La Porta, A. Single-molecule study of g-quadruplex disruption using dynamic force spectroscopy. *Physical review letters* **109**, 058101 (2012).

27 McKinney, S. A., Joo, C. & Ha, T. Analysis of single-molecule FRET trajectories using hidden Markov modeling. *Biophysical journal* **91**, 1941-1951 (2006).

28 Lipfert, J., Hao, X. & Dekker, N. H. Quantitative modeling and optimization of magnetic tweezers. *Biophysical journal* **96**, 5040-5049 (2009).

29 Roy, R., Hohng, S. & Ha, T. A practical guide to single-molecule FRET. *Nature methods* **5**, 507-516 (2008).

30 Lane, A. N., Chaires, J. B., Gray, R. D. & Trent, J. O. Stability and kinetics of G-quadruplex structures. *Nucleic acids research* **36**, 5482-5515 (2008).

31 Rezler, E. M. *et al.* Telomestatin and diseleno sapphyrin bind selectively to two different forms of the human telomeric G-quadruplex structure. *Journal of the American Chemical Society* **127**, 9439-9447 (2005).

32 Abbondanzieri, E. A., Greenleaf, W. J., Shaevitz, J. W., Landick, R. & Block, S. M. Direct observation of base-pair stepping by RNA polymerase. *Nature* **438**, 460-465 (2005).

33 Moffitt, J. R., Chemla, Y. R., Smith, S. B. & Bustamante, C. Recent advances in optical tweezers. *Annual review of biochemistry* **77**, 205-228 (2008).

34 Murphy, M. C., Rasnik, I., Cheng, W., Lohman, T. M. & Ha, T. Probing single-stranded DNA conformational flexibility using fluorescence spectroscopy. *Biophysical journal* **86**, 2530-2537 (2004).

35 Gray, R. D., Li, J. & Chaires, J. B. Energetics and kinetics of a conformational switch in G-quadruplex DNA. *The journal of physical chemistry. B* **113**, 2676-2683 (2009).

36 Mashimo, T., Yagi, H., Sannohe, Y., Rajendran, A. & Sugiyama, H. Folding pathways of human telomeric type-1 and type-2 G-quadruplex structures. *Journal of the American Chemical Society* **132**, 14910-14918 (2010).

37 Zhang, Z., Dai, J., Veliath, E., Jones, R. A. & 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 research* **38**, 1009-1021 (2010).

38 Strick, T. R., Allemand, J. F., Bensimon, D. & Croquette, V. Behavior of supercoiled DNA. *Biophysical journal* **74**, 2016-2028 (1998).

39 Carrion-Vazquez, M. *et al.* The mechanical stability of ubiquitin is linkage dependent. *Nature structural biology* **10**, 738-743 (2003).

40 Li, P. T., Bustamante, C. & Tinoco, I., Jr. Unusual mechanical stability of a minimal RNA kissing complex. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 15847-15852 (2006).

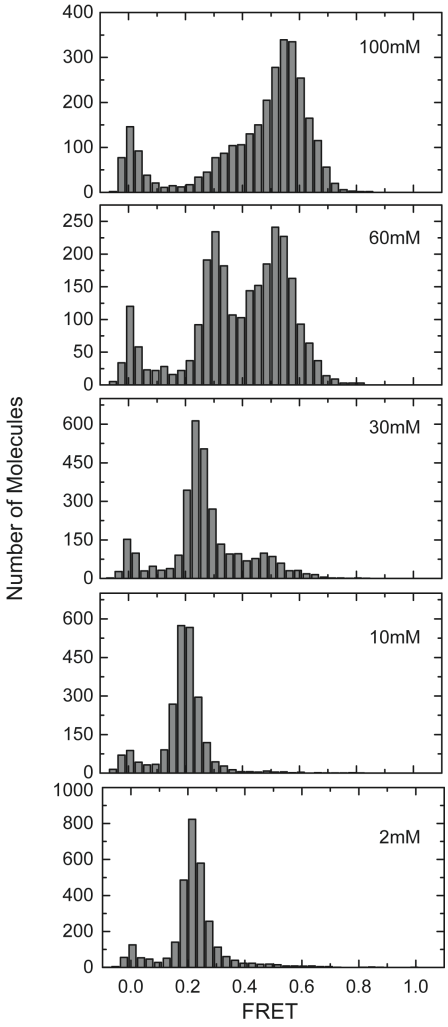
41 Williams, P. M. *et al.* Hidden complexity in the mechanical properties of titin. *Nature* **422**, 446-449 (2003).

SUPPLEMENTARY INFORMATION

Supplementary Table 1. DNA Oligos

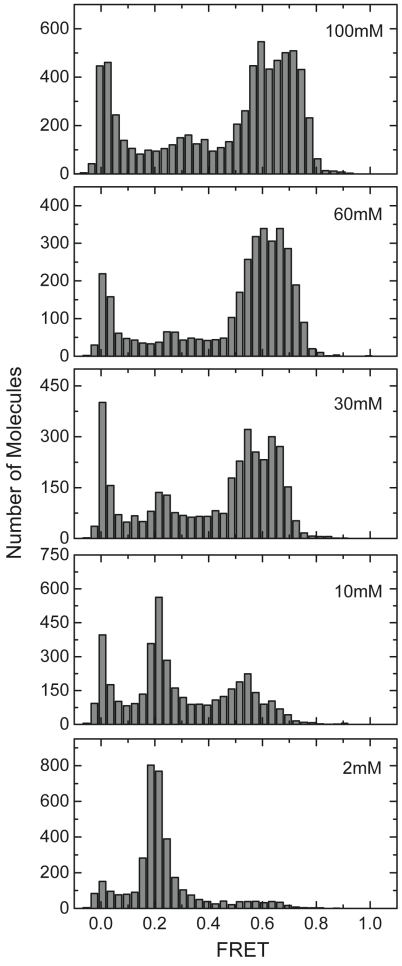
Tel24 Abasic	5' GCGTGGCACCGGTAATAGGAAAATGGAGA(abasic spacer)$(TTAGGG)_4$(abasic spacer)CAGCTACGTGCAC-3'
polyT Abasic	5' GCGTGGCACCGGTAATAGGAAAATGGAGA(abasic spacer) T_{24} (abasic spacer)CAGCTACGTGCAC-3'
15R60/T8 Hairpin Abasic	5' GCGTGGCACCGGTAATAGGAAAATGGAGA(abasic spacer)GAGTCCTGGATCCTGTTTTTTTTCAGGATCCAGGACTC(abasic spacer)CAGCTACGTGCAC-3'
Biotin DNA Handle T3 Amino	5' TCT(AminoC6)CCATTTTCCTATTACCGGTGCCACGC-Biotin 3'
DNA Handle T16 Amino with EcoRI overhang	5' Phosphate-AATTGTGCACGTAGCT(Amino C6)G 3'
PCR Primer I for Synthesized Digoxigenin labeled DNA Linker	5' ACATTTCCCCGAAAAGTGCCA 3'
PCR Primer II for Synthesized Digoxigenin labeled DNA Linker	5' GTTTCGCCACCTCTGACTTGA 3'

Supplementary Figure 1



Supplementary Figure 1. Single molecule FRET distributions for Tel24 molecules in the absence of force as a function of NaCl concentration. Increasing NaCl concentrations stabilizes the fraction of molecules in the folded FRET = 0.54 state. At low ionic strength Tel24 is unfolded as indicated by the FRET ~ 0.25 population. Histograms were compiled by averaging the FRET value for each molecule over a two second period (20 frames at 100 msec integration time). The peak centered on FRET = 0.0 corresponds to molecules that posses an inactive Cy5 dye.

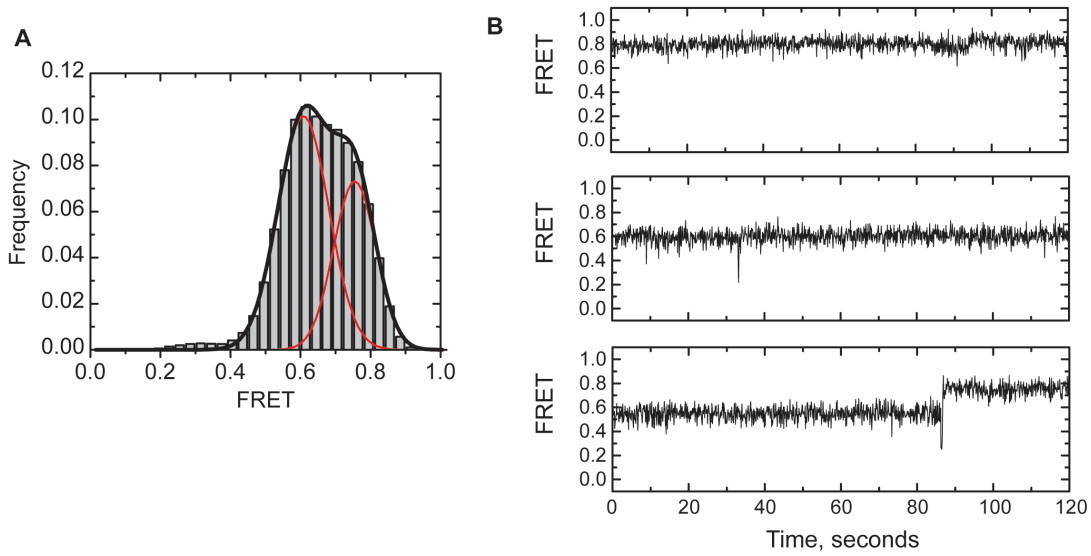
Supplementary Figure 2



Supplementary Figure 2. Single molecule FRET distributions for Tel24 molecules in the absence of force as a function of KCl concentration. Increasing KCl concentrations stabilizes the fraction of molecules in one of the higher FRET folded conformations. At low ionic strength Tel24 is unfolded as indicated by the FRET ~0.2 population. Histograms were compiled by averaging the FRET value for each molecule over a two second period (20 frames at 100 msec integration time). The

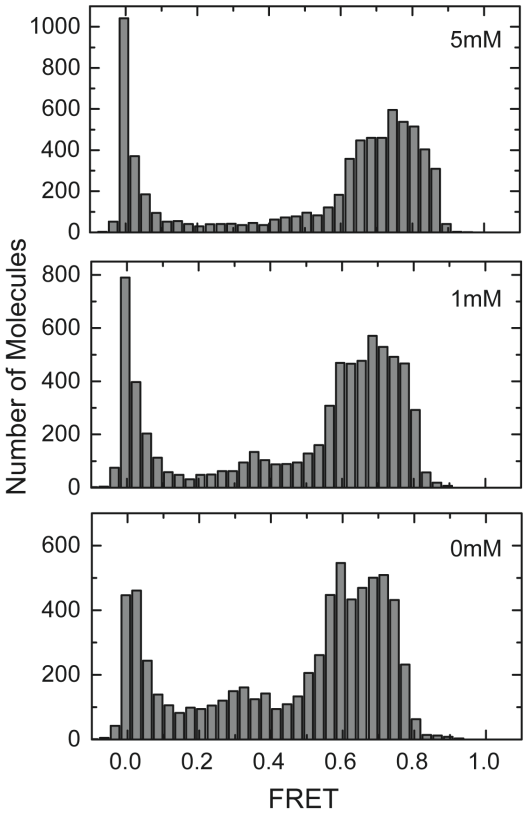
peak centered on FRET = 0.0 corresponds to molecules that posses an inactive Cy5 dye.

Supplementary Figure 3



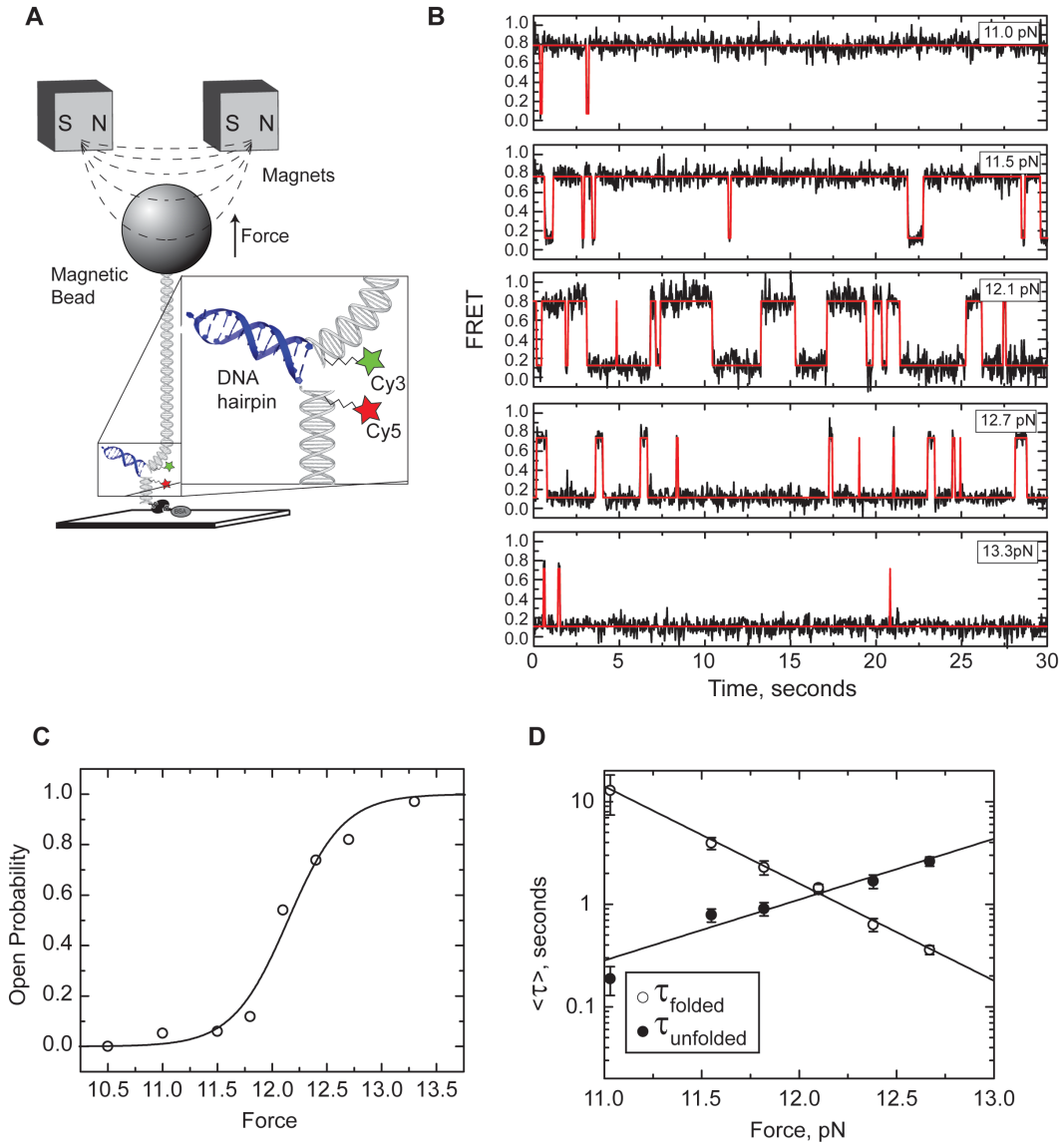
Supplementary Figure 3. Increased complexity of Tel24 folding in the presence of 100 mM KCl. (A) Single molecule FRET histogram comprised from data taken on >100 molecules. Histogram was fit with two guassian functions centered at FRET = 0.61 and FRET= 0.76. (B) Representative smFRET trajectories demonstrating stable high FRET (top), mid FRET (middle), and occasional interconversion between the two FRET states (bottom).

Supplementary Figure 4



Supplementary Figure 4. The effect of Mg^{2+} on the smFRET distributions of Tel24 in the presence of 100 mM KCl. To determine whether adding Mg^{2+} to the KCl folding buffer had any measurable impact on the FRET conformations observed in our assay, smFRET distributions were constructed for molecules folded in the presence of 100 mM KCl and the indicated amount of $MgCl_2$. The presence of physiological levels of Mg^{2+} appeared to have a slightly stabilizing effect on the Tel24 GQ fold, with fewer molecules occupying the unfolded low FRET state. However, the general behaviour of the Tel24 molecules was not substantially altered by the presence of Mg^{2+} in the reaction buffer.

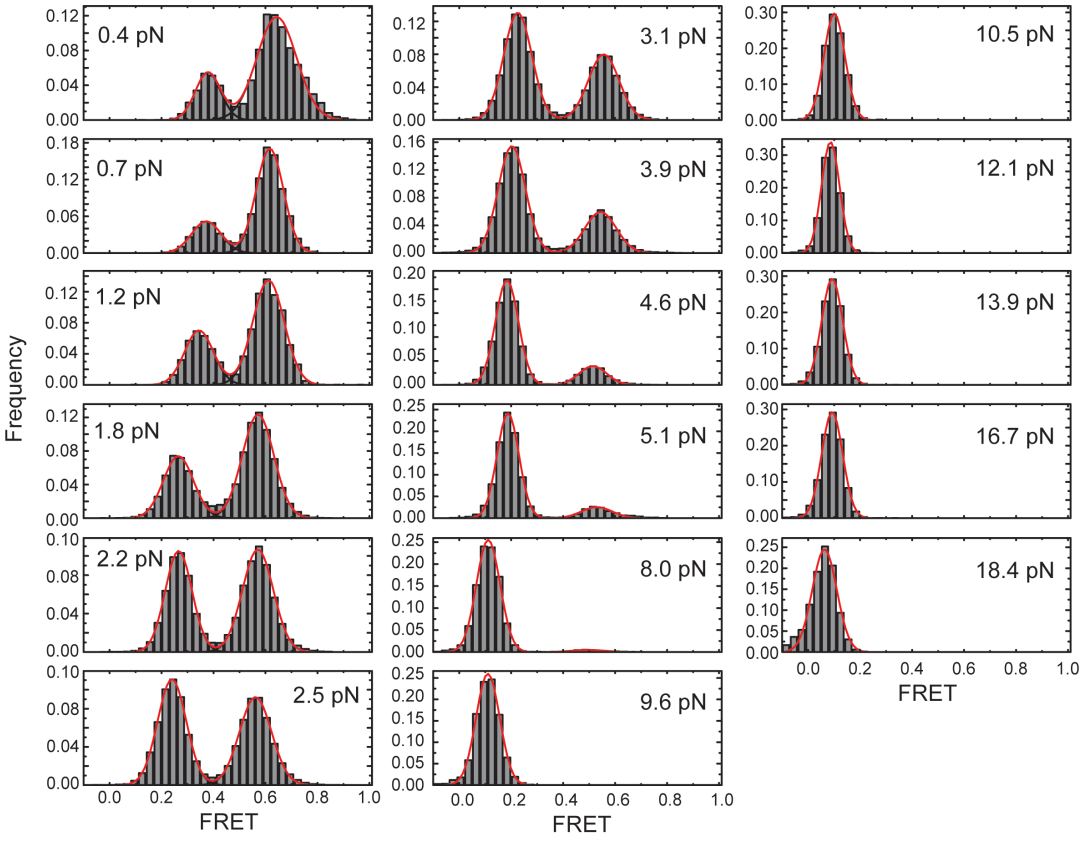
Supplementary Figure 5



Supplementary Figure 5. Integrated fluorescence and magnetic tweezers analysis of a model DNA hairpin. (A) Schematic diagram of the experimental setup for measurements. A FRET-labeled DNA hairpin molecule (see Supplementary Table 1 for sequence) is flanked by a short biotinylated duplex DNA handle on one side, and a second longer duplex DNA handle with terminal digoxigenin modifications on the

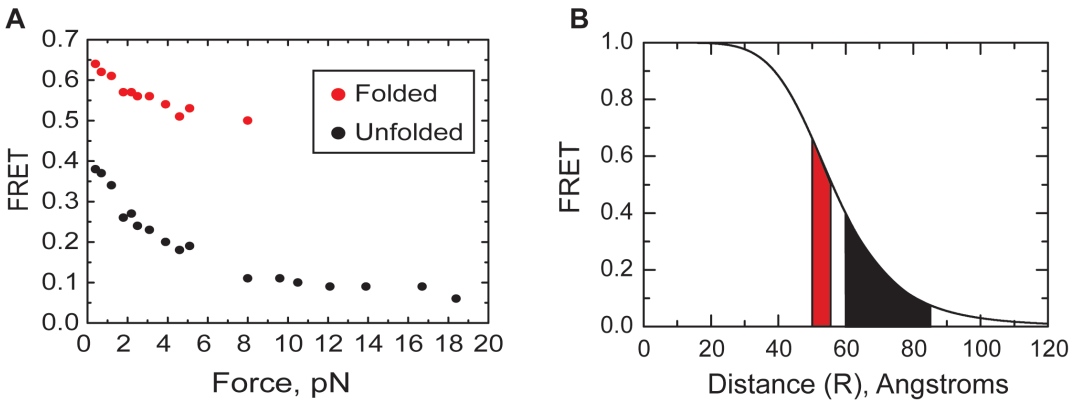
other side. Individual DNA hairpin constructs were attached between a streptavidin coated microscope slide and an anti-digoxigenin coated magnetic bead in a magnetic tweezers setup and imaged using objective-type total internal reflection microscopy. Variable stretching force was applied to individual DNA molecules by translating a pair of rare earth magnets held above the sample chamber. (B) Representative smFRET trajectories of a single DNA hairpin molecule held under five different forces (11,11.5, 12.1, 12.7, and 13.3 pN). The red line represents the model produced by fitting the smFRET trajectory to a Hidden Markov model using the HaMMy software package. (C) Probability of unfolded state as a function of force fit with the two-state Boltzmann relation, $P_0(F)=\{1+\exp[((F_{1/2} - F)\bullet\Delta x)/k_BT]\}^{-1}$, yielding $F_{1/2}$ = 12.1pN and Δx = 16.2 nm. (D) Average lifetimes (τ) of the folded (high FRET) and unfolded (low FRET) states are plotted as a function of force. Plots are fit with exponential functions described in the main text for Figure 5A, yielding $\Delta x^{\ddagger}_{unfolded}$ = 6.1 nm and $\Delta x^{\ddagger}_{folded}$ = 8.9 nm.

Supplementary Figure 6



Supplementary Figure 6. Complete data set for all Tel24 molecules analyzed in study. A total of 75 different molecules were analyzed. At each force, a minimum of five different molecules were studied. Each smFRET trajectory was fit with a Hidden Markov model using the HaMMy software, yielding an idealized FRET trace. Data corresponding to the rare transitions to the high FRET ~0.7 FRET state (see Figure 1) were removed prior to compiling the smFRET histograms, and each histogram was fit with either one or two Gaussian functions using the Origin software package.

Supplementary Figure 7



Supplementary Figure 7. Analysis of the change in FRET observed for the mid FRET (folded, red) and low FRET (unfolded, black) states as a function of force. (A) Plot of the average FRET value derived from fits shown in Supplementary Figure 6 as a function of force. (B) Plot of the FRET response curve described by the equation, $FRET = 1/[1+(R/R_o)^6]$, where R is the distance between the FRET dyes and R_o is estimated to be 56 angstroms. Note, since we have not determined the exact value of R_o under our specific experimental conditions, this plot is not intended to be interpreted as a precise conversion of our measured FRET values to absolute distances. Rather, we provide the plot to illustrate an estimate of the relative distance change that would correspond to the ΔFRET measured for the folded (red) and unfolded (black) states as a function of applied stretching force.

CHAPTER V

The telomerase essential N-terminal domain promotes DNA synthesis by stabilizing short RNA-DNA hybrids.

*Originally published in Nucleic Acids Research 2015

ABSTRACT

Telomerase is an enzyme that adds repetitive DNA sequences to the ends of chromosomes and consists of two main subunits: the telomerase reverse transcriptase (TERT) protein and an associated telomerase RNA (TER). The telomerase essential N-terminal (TEN) domain is a conserved region of TERT proposed to mediate DNA substrate interactions. Here, we have employed single molecule telomerase binding assays to investigate the function of the TEN domain. Our results reveal telomeric DNA substrates bound to telomerase exhibit a dynamic equilibrium between two states: a docked conformation and an alternative conformation. The relative stabilities of the docked and alternative states correlate with the number of basepairs that can be formed between the DNA substrate and the RNA template, with more basepairing favoring the docked state. The docked state is further buttressed by the TEN domain and mutations within the TEN domain substantially alter the DNA substrate structural equilibrium. We propose a model in which the TEN domain stabilizes short RNA-DNA duplexes in the active site of the enzyme, promoting the docked state to augment telomerase processivity.

INTRODUCTION

Telomerase is a ribonucleoprotein enzyme that maintains the ends of eukaryotic chromosomes by synthesizing repetitive DNA sequences that serve as the foundation for protective nucleoprotein structures called telomeres [1]. Telomerase counteracts the loss of telomeric DNA that arises due to the inability of the conventional DNA replication machinery to completely replicate DNA ends. Thus, telomerase solves the 'end replication problem' and helps to avoid cell growth arrest triggered by the presence of critically short telomeres [2]. Mutations within subunits of the telomerase holoenzyme give rise to genetic disorders characterized by deterioration of proliferative tissues, such as the heritable diseases dyskeratosis congenita and aplastic anemia [3]. On the other hand, inappropriate telomerase activation helps to confer the ability for cells to divide indefinitely and is associated with ~90% of human cancers, making telomerase a promising target for potential cancer therapies [4].

Telomerase consists of two main components, a protein telomerase reverse transcriptase (TERT) and a telomerase RNA (TER) (Figure 1A). TERT is tightly associated with TER, and functions by repetitively reverse transcribing a short template region of TER into telomeric DNA [5]. The template region basepairs with the DNA primer to form an RNA-DNA hybrid that is recognized by the TERT active site (Figure 1B). The telomerase catalytic cycle can be sub-divided into two distinct activities: nucleotide addition processivity (NAP) and repeat addition processivity (RAP). During NAP, the telomere DNA substrate is progressively extended to the strictly defined template boundary. Next, during RAP the nascent DNA must

dissociate from the RNA template, re-anneal downstream, and enter the TERT active site for the subsequent round of NAP [5] (Figure 1B).

TERT has several conserved domains, including the telomerase N-terminal (TEN) domain, the RNA binding domain (RBD), the reverse transcriptase domain, and a C-terminal extension (Figure 1A) [6]. Previous studies in *Oxytricha aediculatus* have shown that TERT cross-links to a region of the telomeric DNA ~20 nucleotides upstream of the nascent telomeric DNA 3' end [7]. Subsequently, yeast studies identified the site of a similar cross-link in the TEN domain and determined that mutations that disrupted this cross-link also had an effect on telomerase extension activity, suggesting a functional interaction [8]. Studies in human telomerase also revealed an interaction site between human TERT and single-stranded DNA that occurs independently of telomerase RNA and further mapped this contact to the N-terminal region of the protein [9].

The crystal structure of the TEN domain from *Tetrahymena thermophila* revealed a protein domain with a novel fold [10]. Though a co-crystal structure with DNA was not obtained, a series of surface-exposed residues were implicated in both DNA cross-linking activity and telomerase extension activity using mutagenesis experiments. A particularly interesting mutant in this study was to residue Q168, which is a highly conserved amino acid in the TEN domain found within the T2 (also known as the GQ) motif of TERT [11]. A mutation to this amino acid had a severe effect on both the cross-linking activity of the enzyme and the catalytic activity of reconstituted telomerase [10]. In a separate cross-linking study, a second cross-link

between the DNA and the *Tetrahymena* TEN domain was identified, mapping a contact between residue W187 and the telomeric DNA primer in a region directly adjacent to the 3' end of the DNA in the active site of the enzyme [12].

Taken together, these experiments suggested the TEN domain mediates an "anchor site" interaction between the telomeric DNA substrate and the enzyme [7,8,10,12,13]. The TEN anchor site model posits that the 5' end of the DNA is bound by the TEN domain, such that when the 3' end of the DNA dissociates from the template RNA during RAP, the anchor site interaction with the 5' end of the DNA is maintained, preventing dissociation of the primer. Following the formulation of this model, yet another TEN domain mutant (L14A) was characterized in the *Tetrahymena* telomerase system which had a severe effect on RAP without affecting the anchor site interaction in the protein [14]. Telomerase harboring a mutation to residue L14 was competent to complete a single round of telomere repeat synthesis, but could not undergo productive translocation to generate RAP products [14]. Interestingly, when the L14A mutant was tested in the context of endogenous *Tetrahymena thermophila* telomerase processivity factors, the enzyme retained the ability to undergo RAP, albeit at a substantially reduced rate [15]. More recently, studies on human telomerase demonstrated that the TEN domain is required for RAP and that certain TEN domain constructs could complement a TEN domain deletion mutant in *trans* to restore RAP [16].

Further studies utilizing a sensitized telomerase enzyme lacking an internal RNA template showed that human telomerase could elongate a short RNA-DNA

hybrid provided in *trans* [17], and do so in a TEN domain-dependent manner [18], raising the possibility that the TEN domain may possess activities beyond providing a distal 5' anchor site. Indeed, another study using human telomerase demonstrated that TEN domain mutants exhibit kinetic defects independent of their binding defects, suggesting that the TEN domain may have an additional function to its role as an anchor site [19].

Here, we have employed a single molecule telomerase binding assay [20,21], together with telomerase direct primer extension assays, to directly interrogate the impact of TEN domain mutants on telomere DNA dynamics and telomerase activity. We demonstrate that DNA substrates bound within wild-type telomerase enzymes exhibit dynamic rearrangements between two clearly resolved conformations. The observed internal DNA structural equilibrium correlates with the extent of potential hybrid formation between the DNA primer and the RNA template. Furthermore, under our experimental conditions, mutations to TEN domain residues L14, Q168, or F178 significantly alter the DNA structural equilibrium in primers with the capacity to form short RNA-DNA hybrids (< 5 bp) but not for primers that may form longer heteroduplexes in the telomerase active site. Taken together, our experiments provide evidence that a DNA primer bound within the *Tetrahymena* telomerase complex may exist in one of several possible conformations: a docked conformation wherein the DNA is hybridized to the template RNA in the active site, or a second alternative conformation in which the DNA remains bound to the enzyme but is positioned away from the active site. These results provide support for a model in which a critical

function of the conserved TEN domain is to stabilize the docked conformation of the enzyme for short primers where basepairing stability is expected to be minimal.

MATERIAL AND METHODS

Dye-labeling of synthetic oligonucleotides

Dye-labeling of synthetic DNA and RNA fragments was performed as previously described [21,22]. Synthetic DNA primers (IDT) were ordered containing an amine modification at the desired labeling site and incubated with amine-reactive Cy3 dye (GE Lifesciences) in 0.1M sodium bicarbonate solution. Synthetic RNA fragments were also ordered containing site-specific amine modifications (Dharmacon) and labeled in the same fashion. Dye-labeled oligonucleotides were purified by reverse-phase HPLC. Synthetic RNAs were then splint-ligated to generate full-length telomerase RNA, and the desired RNA was PAGE purified.

Telomerase reconstitution and purification

Telomerase was reconstituted in rabbit reticulocyte lysate (RRL) as previously described [21]. Briefly, 6 pmol of dye-labeled TER was incubated with 25 pmol of recombinant purified p65 in a final volume of 12.5 µl for 10 minutes at room temperature. This was added to a mixture containing 200 µl T7-coupled transcription/translation RRL (Promega), 4.13 µg FLAG-TERT expression plasmid, 5 µl PCR enhancer, and 5 µl 1 mM methionine in a final volume of 250 µl. This was incubated at 30°C for 2 hours. Assembled telomerase was purified by immunoprecipitation using anti-FLAG conjugated beads (Sigma). Telomerase-containing RRL was incubated overnight with anti-FLAG beads. The beads were then

washed in a wash buffer containing 300 mM potassium glutamate. Telomerase was eluted in a buffer containing 1 mg/ml FLAG peptide (Sigma), 50 mM Tris pH 8.0, 1.25 mM $MgCl_2$, and 10% glycerol. Aliquots of purified telomerase were flash-frozen in liquid nitrogen for future use.

TERT mutagenesis

The FLAG-TERT expression plasmid was mutagenized using PCR mutagenesis and custom PCR primers (IDT). Linear PCR amplicons were ligated using DNA ligase (NEB) and used to transform DH5α competent cells and isolated by mini-prep (Qiagen). Each plasmid was then sequenced to determine whether it had the correct modification.

Single-molecule FRET telomerase activity assay

Single-molecule FRET slides were thoroughly cleaned and PEGylated as described [23]. Prepared slides were then incubated in 10 mg/ml BSA for 10 minutes, and rinsed with T50 buffer (10 mM Tris pH 8.0, 50 mM NaCl). Next, 200 µL of 10 pM purified Cy3-labeled DNA was flowed over the slide. Eluted telomerase containing a Cy5-modification in the TER subunit was added in a buffer containing 10 µl eluted telomerase, 18 µl telomerase imaging buffer (50 mM Tris pH 8.0, 1.25 mM $MgCl_2$, 0.5% glucose, 10% glycerol, 1 mg/mL trolox), 1.5 µl 10 mg/ml BSA, and 0.5 µl glucose-oxidase catalase solution (100 mg/ml glucose oxidase, 0.4 mg/mL catalase in T50). FRET was observed using a prism-type total internal reflection microscope on an Andor CCD camera with an integration time of 100 ms. FRET traces were analyzed using custom Matlab software (Matlab). FRET was measured

over the course of the binding event using the formula $E = 1/(1 + \gamma(I_D/I_A))$, where E is FRET efficiency, I_D is donor intensity and I_A is acceptor intensity. The factor γ adjusts for differences in dye quantum yields and can be useful in correcting FRET efficiency when there is a protein-induced Cy3 enhancement, as was observed in a subset of our traces. Because we cannot distinguish between acceptor bleaching events and enzyme dissociation from the primer, we could not determine γ by the previously established method [24]. Instead, we approximated γ as $(I_{D1} + I_{A1})/(I_{D2} + I_{A2})$, where $I_{D1} + I_{A1}$ represents the sum of the donor and acceptor intensity before protein binding and $I_{D2} + I_{A2}$ represents the sum of the donor and acceptor intensities after binding [21]. The factor γ was determined individually for each trace and was consistent with previously reported values of protein-induced Cy3 enhancement.

Telomerase extension assays

Telomerase for *in vitro* telomerase extension assays was prepared in RRL as described above, however instead of dye-labeled telomerase RNA, *in vitro* transcribed telomerase RNA was used instead. RRL reactions were not immunopurified, but were used directly in telomerase extension assays. 5 µL RRL reaction was added to1 µM DNA primer, 100 µM dTTP, 9 µM dGTP, 1 µM ^{32}P α-dGTP, in a final volume of 15 µL in a buffer containing 50 mM Tris pH 8.0, 1.25 mM $MgCl_2$, and 10% glycerol. Reactions were then phenol:chloroform extacted and ethanol precipitated. Prior to phenol:chloroform extraction a radiolabeled recovery control was added, consisting of 5'-end-labeled ^{32}P TER. Extension products were resolved on a 12% PAGE DNA sequencing gel and imaged using a Typhoon scanner

(GE Lifesciences) with a phosphor screen (GE Lifesciences).Telomerase activity assays were performed in triplicate, and gels were quantified using the program SAFA [25]. The intensity of each band was then corrected for the number of radio-labeled guanines incorporated at that band. Nucleotide addition processivity was measured as the sum of the corrected quantity of the first repeat addition band and all bands above the first repeat band. This value was then further corrected using the recovery control. For each enzyme, the extension activity on the primer that can form the most RNA-DNA basepairs—the $(GGGGTT)_3$ primer—was set to 1.0. Therefore the quantification measured the amount of primers extended to, or beyond, the first repeat addition band relative to the maximum activity observed on that enzyme.

RNA dot blot quantification

Telomerase was prepared in RRL and immunoprecipitated as previously described for single-molecule FRET assays, however instead of dye-labeled telomerase RNA, *in vitro* transcribed telomerase RNA was used. 2.5 µl and 5µl aliquots of immunopurified telomerase was diluted to 10 µl in formamide loading buffer (90% deionized formamide, 0.1% bromphenol blue, 0.1% xylene cyanole, and 1X TBE) and heated at 70°C for 5 minutes and placed on ice. The solution was dotted onto a piece of Hybond N+ membrane (GE Lifesciences) and cross-linked to the surface using a UV transilluminator for 1 minute. The membrane was blocked in 10 ml Church buffer (1% BSA, 1 mM EDTA, 500 mM sodium phosphate pH 7.2, 7% SDS) at 55°C for 30 minutes. Approximately 3 x 10^6 cpm of a 5'-^{32}P-end-labeled DNA probe was added to the solution (sequence: 5'-

TATCAGCACTAGATTTTTGGGGTTGAATG-3') and incubated at 55°C overnight. The membrane was washed three times in 0.1X saline-sodium-citrate buffer (15 mM NaCl, 1.5 mM trisodium citrate, pH 7.0) containing 0.1% SDS at room temperature. The membrane was imaged using a phosphor screren (GE Lifesciences) and a typhoon scanner (GE lifesciences). Quantification of the blot was performed with ImageJ. To determine concentrations, samples were compared against *in vitro* transcribed telomerase RNA standards dotted onto the same blot.

HaMMy analysis

Individual single-molecule traces were analyzed by HaMMy [26]. HaMMy was instructed to identify 3 states for each trace: the 0.0 FRET unbound state, the 0.75 FRET active state, and the 0.25 FRET alternative state for U36-labeled enzyme and the 0.0 FRET state, the 0.50 FRET state, and the 0.90 FRET state for the U63-labeled enzyme. Individual dwell times for each trace were compiled together in a single table and plotted as a histogram using Origin (Originlab). The histograms were fit to an exponential decay function $y = A_0e^{-x/\tau} + y_0$, where A_0 represents the amplitude, τ represents the average dwell time, and y_0 represents the y offset.

RESULTS

Direct observation of DNA primer dynamics within single telomerase enzymes

To determine how the TEN domain influences conformational rearrangements within the telomerase holoeznyme, we required a method that permits direct observation of structural dynamics in telomerase bound to its DNA substrate. Traditional methods for measuring telomerase-DNA interactions cannot directly

detect such dynamic structural rearrangements in a DNA primer bound within a telomerase enzyme. To overcome this challenge, we turned to a single molecule telomerase binding assay that monitors DNA dynamics within individual telomerase-primer complexes via Förster resonance energy transfer (FRET) [20,21]. The single molecule FRET (smFRET) assay provides a unique opportunity to analyze how previously characterized TEN domain mutants might alter the movement of DNA within the telomerase holoenyzme, and correlate these measurements with the effects of the mutations on telomerase activity.

In a typical experiment we analyze conformational properties of telomerase-DNA complexes using a prism-type total internal reflection fluorescence (TIRF) microscope and measure the distance-dependent energy transfer efficiency between a donor and an acceptor dye incorporated into a telomeric DNA primer and telomerase RNA (Figure 1C) [27]. Unless otherwise indicated, TER was labeled with a Cy5 acceptor dye at residue U36 (Figure 1A), reconstituted into an active telomerase RNP complex, and purified using a FLAG-tag engineered onto the N-terminus of TERT. [21,22,28]. Telomeric primers used in this study were labeled at the 5' most alignment residue with a Cy3 dye modification. Each primer possessed a 5'-$(TG)_8$ dinucleotide repeat sequence followed by varying amounts of native *Tetrahymena* telomere DNA sequence (Figure 1C). These dye modifications in either TER or the DNA primer have no detectable effect on telomerase assembly or catalysis [20,21]. Furthermore, the $(TG)_8$ repeat primers support wild-type telomerase activity and simplify the present

experiments by ensuring the 3'-end of the DNA primer can only bind to the RNA template in a single alignment register [29].

Figure 1

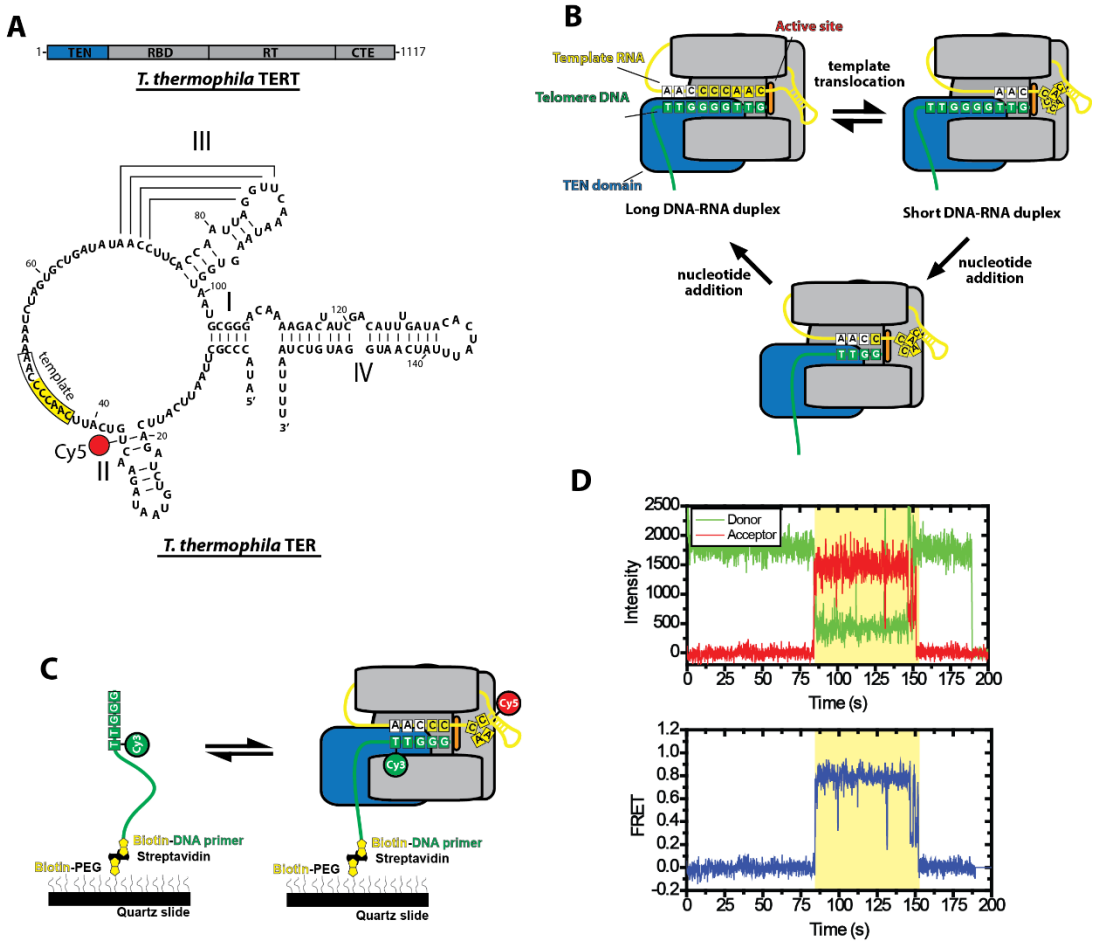


Figure 1. Overview of telomerase smFRET binding assay. (**A**) Domain organization of *Tetrahymena thermophila* TERT and secondary structure of *Tetrahymena thermophila* TER. TERT is divided into the telomerase essential N-terminal domain (TEN, blue), the RNA binding domain (RBD), the reverse transcriptase domain (RT), and the C-terminal extension (CTE). TER contains stems I, II, III, and IV as well as a conserved RNA template (boxed region). The position of the Cy5 modification used for smFRET studies at U36 is indicated. (**B**) Diagram of telomerase catalytic cycle. TERT is represented in grey with the TEN domain highlighted in blue and the active site in orange. The telomeric DNA substrate is represented in green and the telomerase RNA is represented in yellow. The template RNA and telomere DNA form basepairing interactions and this heteroduplex is

positioned in a central channel of the enzyme adjacent to the active site [33]. When the end of the template is reached, the RNA-DNA duplex is denatured and the RNA template re-anneals downstream to position the template for another round of synthesis (template translocation). The post-translocation state of the enzyme contains a short RNA-DNA duplex which must be stabilized in the active site in order to become extended by the enzyme's reverse transcriptase activity to complete the catalytic cycle. (**C**) Schematic diagram of smFRET telomerase binding assay. DNA primers containing telomeric DNA sequence are labeled with a donor Cy3 dye at their 5' most alignment residue and immobilized on a quartz microscope slide by a biotin-streptavidin linkage. Telomerase labeled with Cy5 in its TER subunit is flowed onto the slide and FRET is measured on individual molecules for the duration of the binding events. (**D**) Example smFRET trace for a $(TG)_8T_2G_3$ primer incubated with telomerase labeled at the U36 position of the TER subunit. Donor (Cy3) and acceptor (Cy5) intensity are plotted over time (Top panel). The binding event (shaded region) is marked by the onset of a FRET signal, characterized by an anti-correlated drop in donor fluorescence and rise in acceptor fluorescence. Loss of FRET signal at ~150s occurs either due to Cy5 photobleaching or diffusion of telomerase off of the primer. Loss of Cy3 signal at ~190s is due to a normal process of Cy3 photobleaching. Donor and acceptor intensity values from the top panel are used to calculate a FRET trace in the bottom panel. The FRET values from each point during the binding event are combined with multiple other binding events to generate smFRET histograms.

To measure the interaction between telomerase and DNA substrates, Cy3-labeled DNA primers were surface-immobilized on a microscope slide followed by the addition of purified telomerase harboring Cy5-labeled TER. Binding of a Cy5-labeled telomerase enzyme to the Cy3-labeled DNA primer on the surface was observed as a sudden onset of FRET, characterized by a drop in the donor (Cy3) intensity and an increase in the acceptor (Cy5) intensity (Figure 1D, top). Raw dye intensity values were used to calculate the observed FRET efficiency (Figure 1D, bottom), defined as $FRET = I_A/(I_A+\gamma I_D)$, where I_A and I_D are the intensities of the acceptor and donor dyes, respectively, while γ is a correction factor used to account for effects of the local environment on the photophysical properties of the FRET dyes [24]. smFRET measurements were conducted in the absence of dNTPs; thus, each

experiment represents a telomerase complex bound at a different stage of the telomere repeat synthesis reaction, depending on the telomere sequence present at the 3'end of the DNA primer. For our initial measurements, we used a $(TG)_8T_2G_3$ primer sequence which has the capacity to form up to five basepairs of RNA-DNA hybrid when bound to telomerase (Figure 1C). We note that attempts to measure binding of primers with less telomeric sequence, $(TG)_8T_2G$ or $(TG)_8T_2G_2$, yielded very few binding events, prohibiting accurate measurements. Incubation of wild-type telomerase with $(TG)_8T_2G_3$ primers yielded FRET trajectories that displayed a high FRET ~0.75 state and transient excursions to a lower FRET ~ 0.25 conformation (Figure 2A, top). When many of these binding events are compiled into a smFRET histogram the same two predominant FRET populations are observed (Figure 2A, bottom) consistent with previously reported results using the same primer and enzyme [21]. These experiments demonstrate the ability of the smFRET assay to directly report on the internal structural equilibrium of a telomeric DNA primer bound to a telomerase enzyme.

Mutations in the TEN domain alter telomere DNA dynamics

Next, we analyzed the binding properties of telomerase complexes harboring one of several single amino acid substitutions (L14A, Q168A, F178A, or D94A) in the TEN domain of TERT [10,14,15]. We initially focused on the L14A substitution due to the severe RAP defect that was reported previously for this mutant [14]. FRET trajectories collected on L14A telomerase-DNA complexes differed markedly from wild-type, with the low FRET ~ 0.25 conformation becoming more populated and the high FRET ~ 0.75 state less populated (Figure 2B). In addition, the L14A mutation

gave rise to a general increase in the overall heterogeneity of the FRET behavior, as evidenced by the appearance of transient mid-FRET states in both the single molecule FRET trajectories and histograms (Figure 2B).

Figure 2

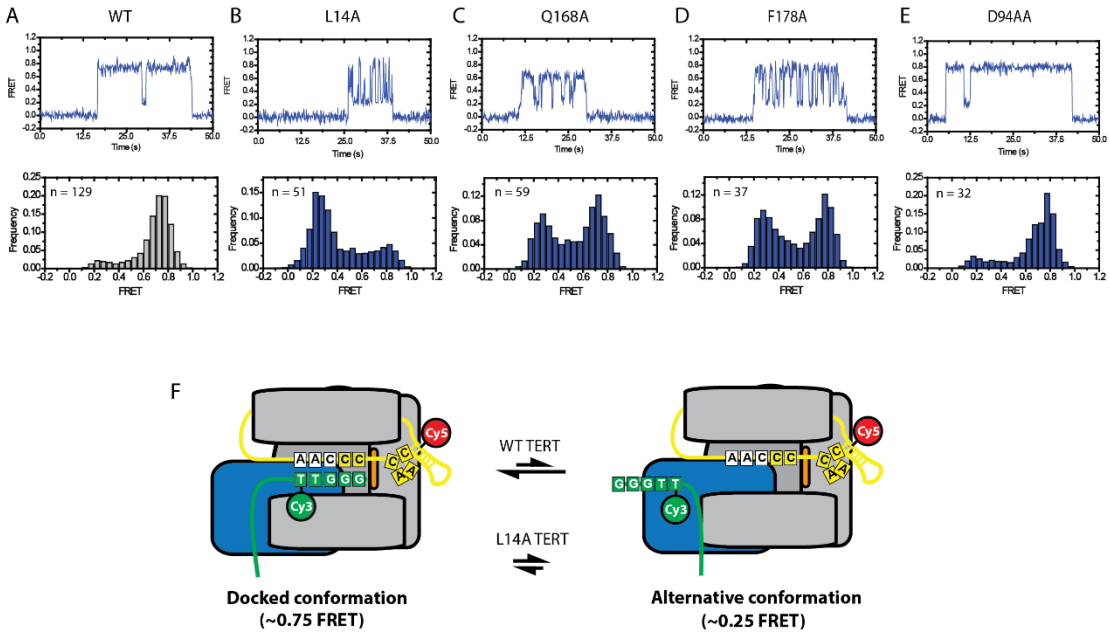


Figure 2: Representative smFRET traces and histograms for wild-type and mutant telomerase. (**A**) Representative smFRET trace (top) and smFRET histogram (bottom) for Cy3-labeled $(TG)_8T_2G_3$ primer incubated with wild-type telomerase labeled with Cy5 at the U36 position of TER. Wild-type enzyme demonstrates a stable ~0.75 FRET state with transient excursions to a ~0.25 FRET state (top panel). This is also reflected in a smFRET histogram of FRET values compiled from 129 separate binding events (bottom panel) demonstrating a predominant ~0.75 FRET distribution with a small shoulder at ~0.25 FRET. (**B,C,D,E**) Representative smFRET traces (top) and smFRET histograms (bottom) for L14A, Q168A, and F178A mutant telomerase respectively. (**F**) Model of telomerase DNA binding dynamics. smFRET data indicates that DNA associated with telomerase can exist in one of at least two conformations. In the docked state, represented by the ~0.75 FRET population, the RNA-DNA duplex is positioned in the enzyme active site. The ~0.25 FRET population represents an alternative state that exists in an equilibrium with the docked state. In this conformation the 3' end of the DNA is positioned away from the enzyme active site. TEN domain residues L14, Q168, and F178 bias the internal equilibrium towards the docked conformation. Importantly, smFRET alone does not provide sufficient information to fully map the contacts present in the alternative state.

Therefore, although we can confidently assert that an alternative state exists, the schematic layout presented in this figure represents only one of many possible organizations that could comprise the alternative state of the enzyme.

We next examined the binding properties of the $(TG)_8T_2G_3$ primer to telomerase enzymes with either a Q168A or F178A mutation in the TEN domain of TERT. These two mutations were shown to reduce the efficiency of cross-linking to the 5'-end of the DNA primer and to reduce the rate of RAP in telomerase activity assays, albeit to a lesser extent than was observed with L14A mutants [10,14]. When the Q168A and F178A TERT mutants were tested in our smFRET assay, a destabilization of the high FRET ~ 0.75 state was once again observed (Figure 2C and 2D). However, the effect of these mutations was less pronounced than was observed with the L14A TERT mutant. Thus, it appears the degree of destabilization of the high FRET ~ 0.75 state conferred by the L14A, Q168A, and F178A mutations correlates well with the extent of the activity defects observed in telomerase direct primer extension assays. It should be noted that the mutation to Q168A slightly changed the center of the ~0.75 FRET state distribution. This could represent a slight rearrangement of this region of the DNA upon mutation of Q168, or this could be due to statistical drift in our measurements.

Finally, we also investigated the binding of $(TG)_8T_2G_3$ primers to enzymes containing a D94A mutation to the TEN domain. The crystal structure of the TEN domain indicated that D94 is positioned on the surface of the TEN domain, distal from the surface containing L14, Q168, and F178 [10]. The D94A mutation was previously shown to have no effect on DNA cross-linking activity and to have a

modest effect on telomerase extension activity [10]. While RAP wasn't explicitly quantified in this study, D94A mutants display a clear banding pattern in telomerase extension assays suggesting they can perform RAP [10]. In our smFRET binding assay, D94A mutants highly resemble the wild-type FRET distribution (Figure 2E and Figure 2A), suggesting that D94 has no effect on the equilibrium between the ~0.75 and ~0.25 FRET states.

The dynamic FRET behavior observed in our experiments could, in principle, arise due to motions in the telomere DNA substrate, the region of TER labeled with Cy5, or both. Therefore, to investigate the physical basis for the different FRET states, we next prepared a telomerase complex reconstituted with TER labeled with Cy5 at residue U63, located within the template recognition element on the opposite side of the RNA template from residue U36 [30] (Supplementary Figure S1A). This labeling position was strategically chosen based upon previous experiments that demonstrated a FRET dye at residue U63 is well-tolerated by the enzyme [20,21]. Interestingly, smFRET traces and histograms for U63-labeled wild-type telomerase bound to Cy3-labeled $(TG)_8T_2G_3$ primer showed a reciprocal FRET signal compared to that observed for the U36-labeled telomerase enzyme, with a predominant mid FRET ~ 0.5 state and a transient high FRET ~ 0.9 state (compare Figure 2A and Supplementary Figure S1B). Moreover, when U63-labeled telomerase harboring the L14A mutation was incubated with $(TG)_8T_2G_3$, the observed smFRET distribution shifted considerably, showing an increase of the FRET ~ 0.9 state and decrease of the FRET ~ 0.5 state (Supplementary Figure S1C). The reciprocal nature of the smFRET

results from experiments with the U36- and U63-labeled telomerase is consistent with these two sites being distant from each other in three-dimensional space as suggested previously [20], and supports the notion that dynamics observed in our smFRET experiments are primarily due to movements of the DNA primer between distinct conformations, rather than RNA conformational changes.

In prior work, the ~ 0.75 FRET state observed with U36-labeled enzyme was interpreted to represent a telomerase enzyme in which the DNA primer is hybridized to the RNA template and positioned in the active site poised for nucleotide extension [21], a conformation we will refer to as the docked state (Figure 2F). The assignment of the ~ 0.75 FRET state to the docked conformation is further supported by our mutagenesis results, which demonstrate that mutations with known activity defects disrupt this state, and the degree of disruption correlates with the known severity of the mutation. The ~ 0.25 FRET state observed in U36-labeled enzyme represents a substantial deviation in FRET from the docked state, indicating the Cy3 label within the DNA primer has been repositioned across a length scale of several nanometers. Furthermore, the increased probability of adopting the ~ 0.25 FRET state observed for processivity-defective mutants strongly suggests this state is not competent for telomere DNA primer extension. We therefore assign the ~ 0.25 FRET state to a conformation in which the 3' end of the DNA primer is displaced from the enzyme active site while the DNA remains bound to the enzyme via other contacts. However, since the smFRET experiments do not provide sufficient structural constraints to know precisely where the DNA is while in the ~ 0.25 FRET state, we refer to this

conformation as the alternative state to differentiate it from the docked state (Figure 2F). Comparing the FRET distributions from the U63-labeled enzyme (Supplementary. Figure. S2B) and the U36-labeled enzyme (Figure 2) strongly suggests that the ~ 0.5 FRET state in the U63 enzyme corresponds to the docked conformation and the ~ 0.9 FRET distribution corresponds to the alternative conformation. Finally, observation of dynamics within the smFRET traces (Figure 2) demonstrates that the two conformations are in dynamic equilibrium and the role of the TEN domain (and in particular residues L14, Q168, and F178) is to bias the equilibrium toward the docked conformation for nucleotide extension (Figure 2F arrows).

The TEN domain stabilizes the docked conformation of the enzyme

Importantly, the analysis of smFRET histograms alone cannot determine whether mutations to the TEN domain destabilize the docked state of the enzyme, or if these mutations bias the internal DNA equilibrium by stabilizing the alternative state of the enzyme. Therefore, to distinguish between these two possibilities we used a hidden-Markov modelling software program that generates idealized FRET trajectories for each of the telomerase-DNA binding events (Figure 3A and 3B) [26]. To simplify the analysis, we elected to treat the internal structural equilibrium of the DNA primer as a two-state system between the low and high FRET states observed in the data collected on the wild-type and L14A enzymes (Figure 2A and 2B). The idealized FRET trajectories were then used to generate dwell time distributions of the time spent in either the high FRET docked or low FRET alternative state (Figure 3C

and 3D). Each dwell time distribution contained at least 100 individual dwell time measurements and was well fit by a single exponential decay function. For wild-type telomerase bound to the $(TG)_8T_2G_3$ DNA primer, the average time spent in the high FRET docked conformation (τ_{docked}) was 5 seconds while the alternative conformation dwell time (τ_{alt}) was 0.8 seconds (Figure 3C). By comparison, when L14A mutant telomerase was incubated with the $(TG)_8T_2G_3$ primer, the average dwell time for the high FRET docked conformation dropped by an order of magnitude to 0.5 seconds, whereas the dwell time for the low FRET alternative state remained essentially unchanged at 0.9 s (Figure 3D). As a control, we performed the same kinetic analysis using data collected on U63-labeled wild-type and L14A mutant telomerase enzymes bound to the $(TG)_8T_2G_3$ primer. In this case, the L14A mutation exerted the largest effect on the dwell time distribution of the predominant FRET ~ 0.5 state, corresponding to the docked conformation (Supplementary Figure S2A-D). Since a mutation to L14 destabilizes the docked conformation but has a negligible impact on the stability of the alternative conformation, we conclude the TEN domain stabilizes the docked conformation of the DNA primer.

We note that in addition to affecting the binding lifetimes of the high and low FRET states, these TEN domain mutations could also have an effect on the overall binding lifetimes of the DNA primer to telomerase. Unfortunately, under the conditions of our assay we cannot accurately measure the off-rate of the enzyme as we cannot distinguish between enzyme dissociation from the DNA primer and photobleaching of the acceptor dye. However, although there could be subtle

differences in the off-rates of the various TEN domain mutants, observation of our single-molecule traces indicate that all of the enzymes tested have smFRET binding events on similar timescales (Figure 2).

Figure 3

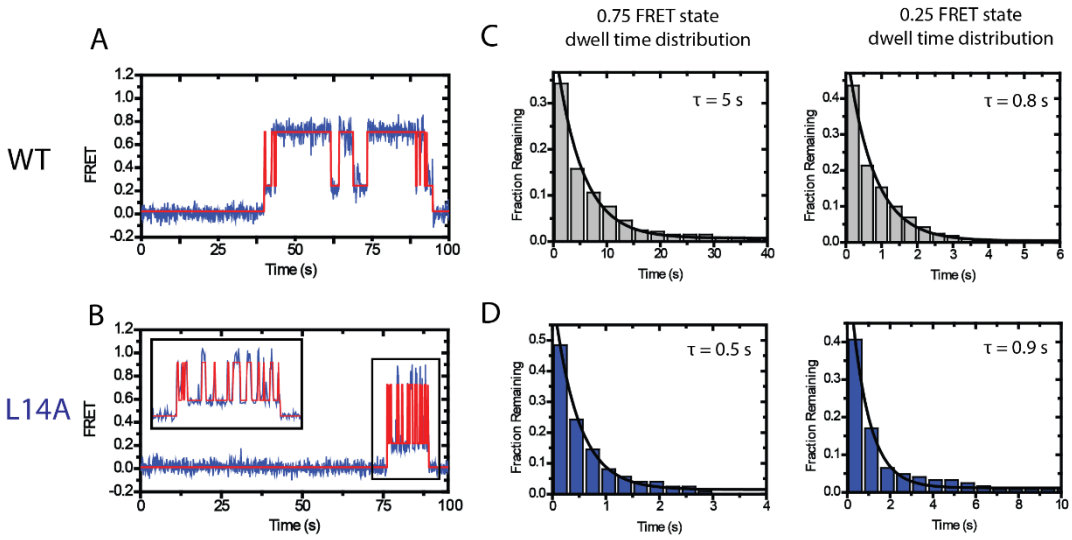


Figure 3. Dwell time analysis of TEN domain mutants demonstrate that the TEN domain stabilizes the docked state. (**A**) smFRET traces (blue) were analyzed by HaMMy [26] to generate idealized traces (red). These were used to determine the dwell time of the enzyme in each state. (**B**) The dwell times for the WT enzyme in the 0.75 FRET state and the 0.25 FRET state incubated with primer $(TG)_8T_2G_3$ were compiled into histograms. The histograms were fit to an exponential function to identify the average dwell time. Wild-type TERT demonstrated a dwell time of τ_{docked} = 5 s for the 0.75 FRET state and a dwell time of τ_{alt} = 0.8 s for the 0.25 FRET state. (**C**) Representative smFRET trace and idealized HaMMy trace for L14A TERT telomerase incubated with the $(TG)_8T_2G_3$ primer. (**D**) Compiled histograms for L14A enzyme. L14A TERT demonstrated a dwell time of τ_{docked} = 0.5 s for the 0.75 FRET state and a dwell time of τ_{alt} = 0.9 s for the 0.25 FRET state.

The TEN domain stabilizes short RNA-DNA duplexes

Having identified a critical role for the TEN domain in stabilizing the docked conformation for the $(TG)_8T_2G_3$ primer, we next set out to analyze the effect of varying the amount of telomeric DNA sequence in the primer. We repeated the

smFRET telomerase binding experiments with a set of DNA primers, each having one additional nucleotide of telomeric sequence added at the 3'-end. Therefore, these primers can in principle make increasing numbers of basepairing contacts with the template region of TER, with the $(TG)_8T_2G_3$ primer having the potential to make a five basepair RNA-DNA duplex and the $(TG)_8T_2G_4T_2G$ primer having the capacity to form up to nine basepairs of RNA-DNA duplex (Figure 4A).

Figure 4

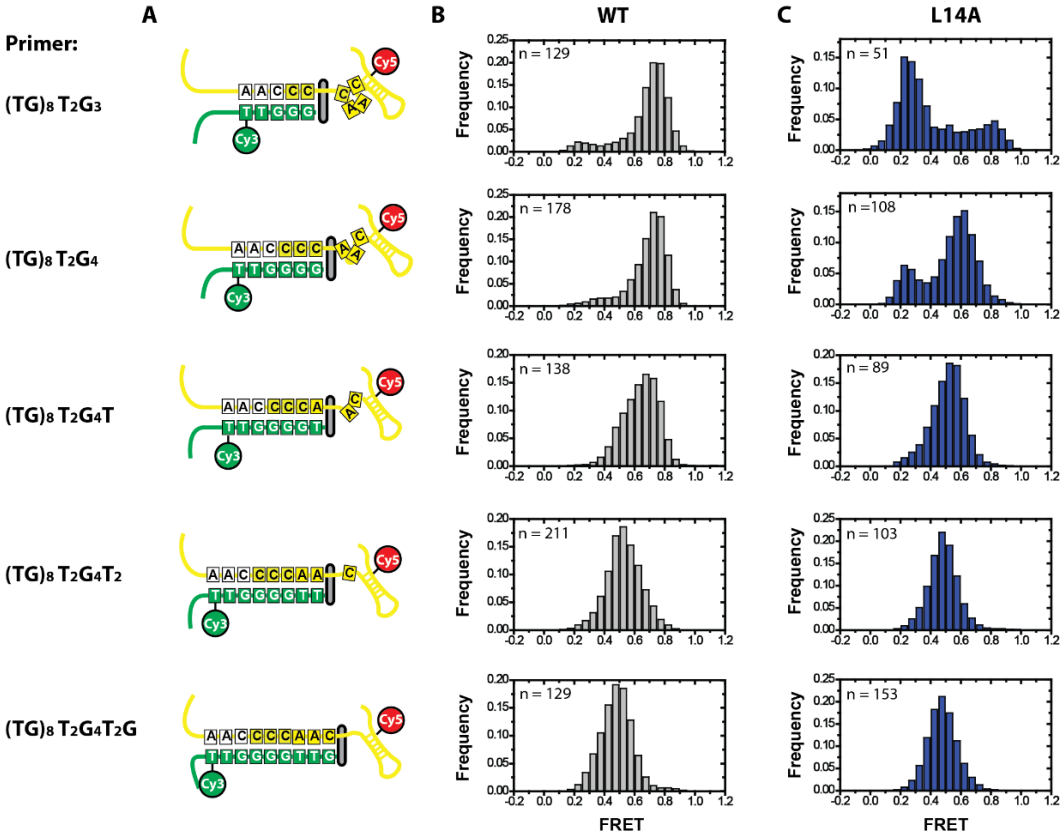


Figure 4. Effect of primer-template hybrid formation on FRET distributions. Primers capable of forming 5-9 basepairs with template RNA were tested in smFRET telomerase binding assays with U36-labeled telomerase. (**A**) Schematic diagram of the docked state for all six primers used in smFRET experiments, demonstrating the number of basepairs formed and the expansion of the template RNA as the RNA-DNA duplex becomes progressively longer [21]. (**B**) smFRET histograms for wild-type enzyme. As primers contain progressively more telomeric DNA sequence, the

predominant FRET distribution of the docked state shifts from ~0.75 FRET to ~0.5 FRET. In addition, the ~0.25 FRET alternative state disappears. (**C**) smFRET histograms for L14A mutant enzyme.

As primers with increasing amounts of telomeric sequence are used, the wild-type FRET distribution undergoes two main changes. First, the predominant FRET state, which we have assigned to the docked conformation, undergoes a gradual shift from a distribution centered at ~0.75 FRET to a distribution centered at ~0.5 FRET (Figure 4B). This drop in FRET was previously demonstrated to represent an expansion in the flexible region of RNA 5' of the template as the RNA-DNA duplex is extruded from the active site [21], consistent with the notion that the high FRET state is reporting on the docked conformation of the DNA primer. The second trend observed is that the ~0.25 FRET distribution disappears in smFRET histograms for primers with increasing telomeric DNA sequence. A likely explanation for this observation is that later catalytic intermediates possess greater numbers of RNA-DNA basepairs which would be expected to stabilize the docked conformation of the enzyme at the expense of the alternative conformation.

Next, we tested L14A mutant telomerase with primers possessing increasing amounts of telomeric sequence. Interestingly, L14A mutant telomerase showed only a modest defect in formation of the high FRET docked conformation with the $(TG)_8T_2G_4$ primer, which has the capacity to form one additional basepair with the TER template region when compared with the $(TG)_8T_2G_3$ primer (compare Figure 4B and 4C). As the amount of telomeric sequence was further increased, the DNA binding properties of the L14A mutant enzyme resembled the wild-type distributions,

with no detectable difference in smFRET distributions observed for the $(TG)_8T_2G_4T_2G$ primer, which can form up to nine basepairs of RNA-DNA duplex (Figure 4B & C). A similar result was obtained when the same set of experiments was performed U63-labeled telomerase, only with the expected inversion of the high and low FRET states as described earlier (Supplementary Figure S3A-C). In addition, the same correlation between increased occupancy of the transient alternative DNA conformation with DNA primers possessing less telomeric sequence was evident for the Q168A and F178A mutants (Supplementary Figure S4A-D). We note that in addition to impacting the dynamic equilibrium of the docked and alternative states, the L14A mutant also has a detectable impact on the centers of the FRET distributions of the docked conformation with the $(TG)_8T_2G_4$ and $(TG)_8T_2G_4T$ primers, which may be due to a slight difference in the structure of the docked state when compared with the wild type enzyme. Taken together, our results demonstrate the L14A, Q168A, and F178A mutants exhibit detectable defects in early catalytic intermediates that contain short RNA-DNA duplexes; however, this DNA binding defect is suppressed for primers corresponding to late catalytic intermediates with the potential to form long RNA-DNA duplexes.

TEN domain mutants fail to extend primers with low RNA-DNA hybrid potential

The telomerase catalytic cycle is often sub-divided into two separate stages: nucleotide addition processivity (NAP) and repeat addition processivity (RAP). NAP is typically used to describe the synthesis of one telomeric DNA repeat, while RAP

refers to the series of molecular rearrangements required to realign the telomerase RNA and telomere DNA subunits in order to add additional telomeric repeats. RAP involves several steps including: pausing nucleotide extension correctly at the end of one telomeric repeat, melting of the existing RNA-DNA duplex, reannealing of a short 3 basepair RNA-DNA duplex in the next alignment register, and extension of the newly-formed short RNA-DNA duplex (Summarized in Figure 1B). Our data suggests that mutations to L14, Q168, and F178 impact this final step in RAP, by preventing the stable association of short RNA-DNA duplexes in the enzyme active site. If this interpretation is true, we would anticipate that L14A, Q168A, and F178A mutants would not only manifest themselves as RAP defective mutants, but should also exhibit NAP defects for primers that form short RNA-DNA duplexes.

To test this prediction, we performed direct primer extension assays with either wild-type TERT or TERT bearing mutations in the TEN domain (L14A, Q168A, F178A, or D94A) using a set of DNA primers that were all 18 nucleotides in length, but had staggered sequences that permitted formation of varying amounts of RNA-DNA hybrid in the telomerase active site (Figure 5A), similar to an assay originally performed on L14A mutants upon discovery of the mutation [14]. As expected, wild-type telomerase efficiently extended all 6 primers to the first complete telomeric DNA repeat band (Figure 5B, red asterisks) and exhibited RAP as evidenced by the accumulation of products extended by multiple telomere repeats. In contrast, the activity of telomerase enzyme harboring the L14A mutation was severely perturbed in NAP for primers with low RNA-DNA duplex potential, but as RNA-DNA hybrid

potential increased NAP activity was restored (Figure 5B, compare lanes 7-9 with 10-12). This is in agreement with previous activity assays on the L14A enzyme [14]. Moreover, primers that were successfully extended by the L14A mutant telomerase to the end of the first nascent telomeric DNA repeat failed to extend beyond this point. This result demonstrates that the RAP defect observed in the L14A mutant enzyme is due to the inability of this enzyme to position short RNA-DNA hybrids in the active site, consistent with our smFRET observations. The Q168A and F178A mutants also displayed a bias in their extension activity with respect to RNA-DNA duplex potential; however the defect was not as severe when compared to the L14A mutant (Figure 5B, compare lanes 7-9, 13-15, and 19-21). The D94A mutant, which demonstrated no detectable defect in our smFRET assays (Figure 2E), similarly did not display a bias against short primers in primer extension assays (Figure 5B, compare lanes 1-3 and 25-27).

We quantified each enzyme's primer-dependent NAP activity by measuring the relative number of primers extended to their first repeat addition band (Figure 5B red asterisks) as a function of the initial RNA-DNA duplex length (Figure 5C). The most populated band, obtained using a primer with an 8 basepair initial hybrid length, was correct to 1.0 for each enzyme. Therefore our quantification consisted of comparing the relative efficiency of extension for primers with short initial RNA-DNA duplexes against primers with longer RNA-DNA duplexes within a single enzyme. In this manner, telomerase activity assays were internally controlled against

variations in enzyme preparation by only comparing activity assays performed with the same enzyme.

Figure 5

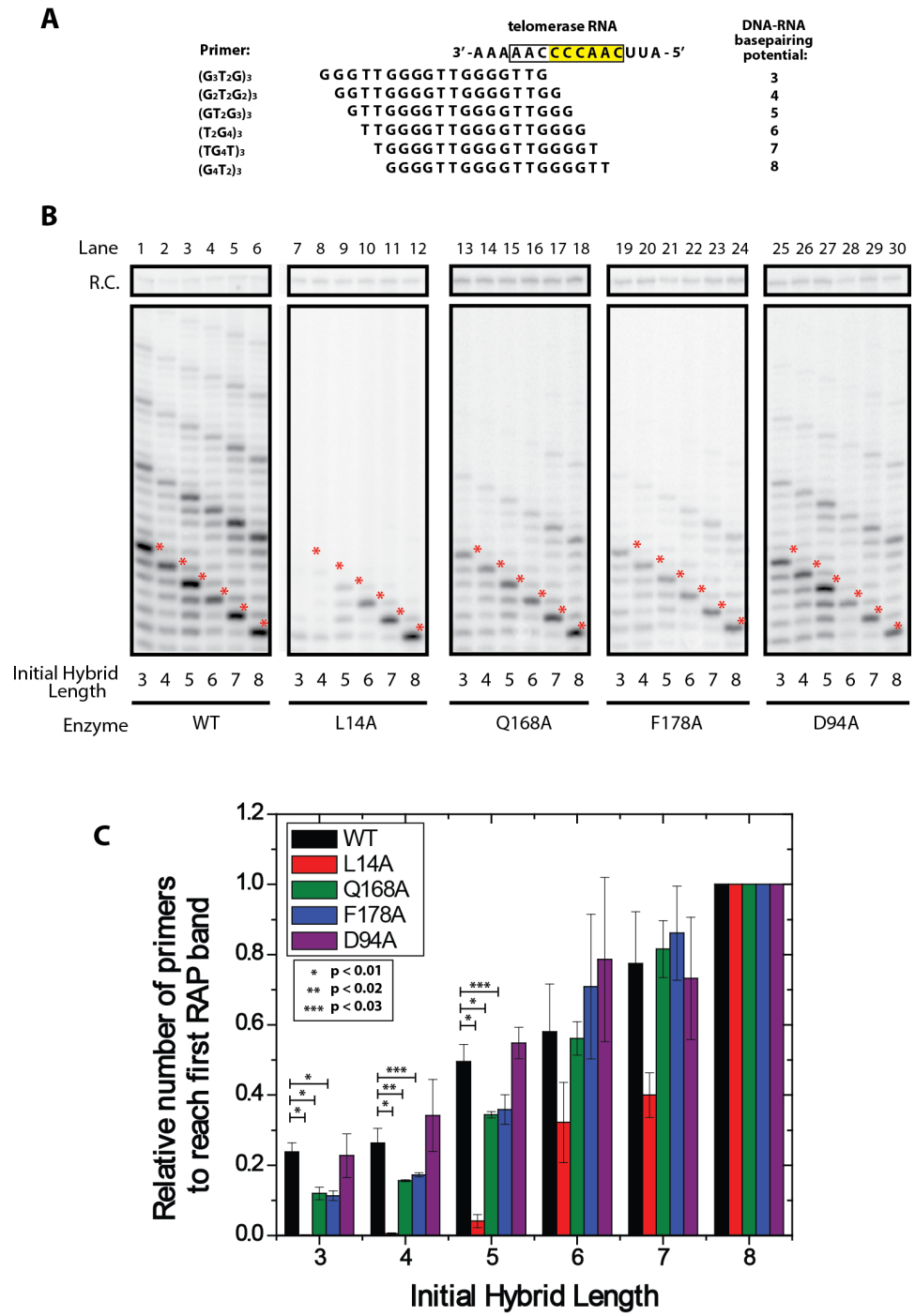


Figure 5. Telomerase activity assays demonstrate TEN domain mutations affect nucleotide addition processivity of primers with short RNA-DNA duplexes. (**A**) Primer permutants used in *in vitro* extension assays. Primers were length-matched at 3 telomeric repeats (18 nts), but staggered such that they formed different initial potential RNA-DNA duplex lengths with template RNA. (**B**) Telomerase was reconstituted in rabbit reticulocyte lysate and telomerase activity was assayed on the six DNA primers corresponding to six potential RNA-DNA hybrid lengths. WT enzyme was compared against enzyme harboring L14A, Q168A, F178A, and D94A mutations. Mutants were assayed for primer-specific NAP defects by comparing the accumulation of the first repeat addition band (red asterisks) between primers tested with the same enzyme. (**C**) Quantification of relative NAP as a function of initial primer duplex length. Telomerse activity assay gels (Figure 5B) were performed in triplicate and quantified using the program SAFA [25]. The quantification of each band was then corrected for specific activity. NAP was quantified as the amount of product that reached the first repeat addition band (Figure 5B, red asterisks) or passed through an intermediate containing the first repeat addition band (all material above the red asterisks). For each enzyme, the primer containing an 8 basepair initial hybrid was corrected to 1.0, and the telomerase activity on all other primers was measured as a fraction of the NAP of this primer. The results demonstrate that L14A, Q168A, and F178A exhibit a strong bias against primers containing short RNA-DNA duplexes. P-values indicating statistical significance are as marked on the graph, error bars indicate one standard deviation based on triplicate measurements.

When corrected for specific activity, our results demonstrate that even wild-type enzyme extends significantly fewer primers to their first repeat addition band when primers contain short initial RNA-DNA duplex lengths versus when primers contain long RNA-DNA duplexes (Figure 5C, black bars) and that the primers gradually recover activity as the initial RNA-DNA duplex length is extended (Figure 5C, black bars). This defect for short initial duplexes is shown to be significantly exaggerated in many of the TEN domain mutants (L14A, Q168A, and F178A) (Figure 5C). For a primer initially containing a 3 basepair RNA-DNA duplex, wild-type enzyme demonstrated 24% of maximal activity, however L14A enzyme demonstrated less than 1% maximal activity and Q178A and F178A mutants each demonstrated 9% activity. As the initial RNA-DNA duplex length of the primers was

extended, the TEN domain mutants increasingly resembled wild-type levels of activity. For a 6 basepair initial hybrid length, wild-type enzyme demonstrated 62% activity, whereas L14A, Q168A, and F178A mutants demonstrated 33%, 55%, and 69% activity respectively. D94A mutants, which demonstrated no detectible defect in smFRET assays, highly resembled wild-type enzyme in their extension activity with all of the primers in the assay (Figure 5C).

Our results demonstrate that the TEN domain mutants which demonstrate a primer-specific defect in our smFRET assays also fail to extend primers with short initial RNA-DNA duplexes in direct primer extension assays and demonstrate a clear trend across primers. Notably, conditions that favor the alternative DNA conformation as measured in the smFRET assays (Figure 4) (ie. primers with low RNA-DNA hybrid potential or TEN domain mutations) also manifest as NAP defects in the direct primer extension assays (Figure 5).

One concern in evaluating how TEN domain mutants affect telomerase activity is whether or not the TEN domain mutants have a secondary effect on protein expression, stability, or telomerase RNA association. Our telomerase activity assays were internally controlled for this possibility by only comparing the activity of telomerase on different primers within the same enzyme preparation. However, to further confirm that TEN domain mutants did not affect TERT expression or RNA association, we performed filter binding assays on telomerase complexes immunopurified from rabbit reticulocyte lysate to measure the amount of RNA assembled with TERT in the context of TEN domain mutants (Supplementary Figure

S5A). The results demonstrated that most TEN domain mutants assemble a similar amount of TERT-TER RNP complexes in the lysate, with only the Q168A mutant demonstrating a slightly reduced amount of protein-RNA complex. These results suggest that these mutations do not act at the level of protein stability or protein-RNA assembly. Furthermore, the material used in the filter binding assays continued to display RAP defects in the L14A, Q168A, and F178A mutants (Supplementary Figure S5B).

DISCUSSION

Previous experiments investigating the role of the TEN domain established the TEN domain as an important site of DNA interaction and identified TEN domain mutants that severely affect the rate of RAP [7,8,10,14]. Here, we conducted smFRET assays to investigate in real-time the effect of TEN domain mutants on DNA dynamics within the telomerase holoenzyme. These assays revealed that telomerase bound to a DNA primer exists in two distinct conformations that are in dynamic equilibrium (Figure 2). We assigned these two conformations to a docked state of the enzyme and an alternative state in which the 3' end of the DNA is displaced from the enzyme active site (Figure 2F). Several lines of evidence support this model. First, there is a strong correlation between the relative occupancies of the alternative and docked states and the strength of processivity defects observed in TEN domain mutants in our telomerase extension assays (Figure 2 and Figure 5). This is highly consistent with the interpretation that the docked state contains DNA positioned in the active site and the alternative state contains DNA positioned away from the active

site. The correlation between activity and the docked state occupancy extends not only to several different mutations but also extends across several DNA primers, such that long RNA-DNA duplex primers that favor the docked state demonstrate a reduced sensitivity to TEN mutations (Figure 4 and Figure 5). Finally, the model that the docked state is stabilized by template-product basepairing is further supported by our smFRET results, as primers with low RNA-DNA hybrid potential demonstrate increased primer dynamics, increased occupancy of the alternative state, and increased susceptibility to TEN domain mutations (Figure 4 and Figure 5). We conclude that an essential role of the TEN domain in *Tetrahymena* TERT is to stabilize the short RNA-DNA duplex in the active site of the enzyme at the start of each telomere repeat synthesis cycle (Figure 6).

Figure 6

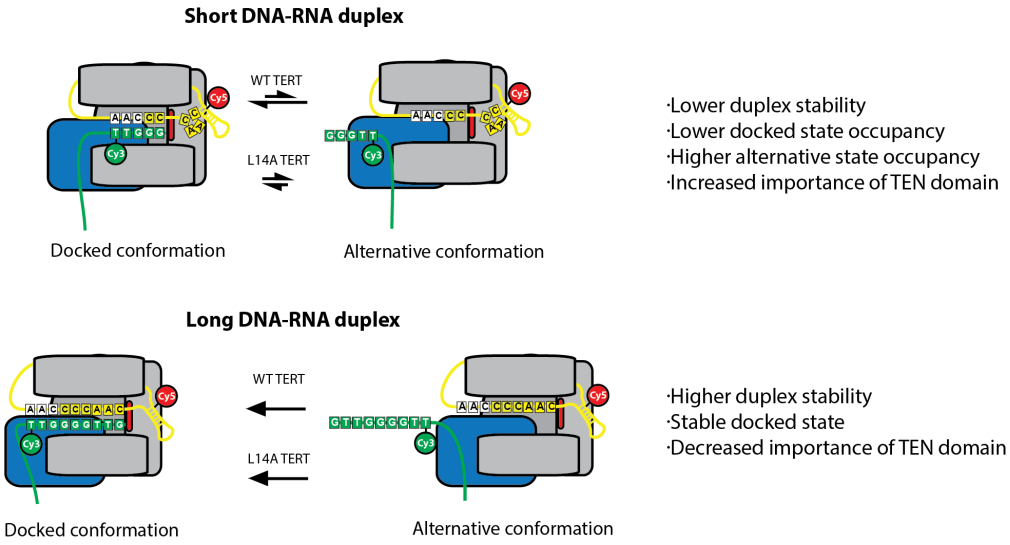


Figure 6. Model demonstrating the role of the TEN domain in stabilizing short RNA-DNA duplexes. Primers corresponding to early catalytic intermediates that contain fewer RNA-DNA basepairs are in a conformational equilibrium between a docked state and an alternative state (top panel). TEN domain mutants destabilize the docked state such that the alternative state is favored, disrupting the catalytic activity

of the enzyme. In primers corresponding to late catalytic intermediates, the docked state is heavily favored due to the presence of additional RNA-DNA basepairs (bottom panel). As a result, the alternative state is not observed, even in the presence of TEN domain mutants.

Our results also explain how telomerase accessory factors can partially rescue TEN domain mutants as previously observed [15]. In our smFRET observations, TEN domain mutants retain the ability to enter the docked state, but do not remain in the docked state stably enough for nucleotide extension. Accessory factors which tether the DNA to the enzyme and lower the substrate off-rate would permit the DNA to remain bound for a sufficiently long time such that the enzyme has an increased probability of stably entering the docked state by chance, permitting multiple rounds of RAP to occur in these complexes, albeit at a reduced rate.

While our smFRET results demonstrate an unambiguous equilibrium between the docked and alternative states, the mechanistic and structural details of these two states remain unclear. For this reason, we do not yet know the exact mechanism by which L14, Q168, and F178 stabilize short RNA-DNA duplexes. Interestingly, residues Q168 and F178 were previously implicated in TEN domain DNA binding by mutagenesis and binding studies suggesting they may interact directly with telomeric DNA [10]. Glutamine and phenylalanine also contain functional groups that can form hydrogen bonding and base stacking interactions with DNA, respectively. This suggests a possible mechanism of DNA interaction.

On the other hand, L14 was not implicated in direct DNA interactions by mutagenesis studies [14]. In the crystal structure, L14 is surface exposed and makes interactions with several other hydrophobic side chains near the surface of the domain

[10]. L14 therefore may be important in protein-protein interactions, which may either help organize an adjacent region of TEN within the domain, or alternatively interact with another domain of TERT to aid in the positioning and/or dynamics of the TEN domain within the context of full-length TERT as suggested previously [14]. Future experiments designed to directly interrogate movements of the TEN domain during telomerase catalysis will be necessary to support or refute these models.

The conservation of the TEN domain across species—including *Tetrahymena*, *S. cerevisae*, and human telomerase—suggests that the role of the TEN domain in stabilizing short RNA-DNA hybrids may be evolutionarily conserved. A conserved role for the TEN domain is further supported by a recent report that demonstrated a sensitized human telomerase enzyme, lacking both an internal RNA template and the TEN domain, can only extend a short RNA-DNA hybrid *in trans* if the TEN domain is added as a separately folded polypeptide [18]. Another argument for the conserved role of the TEN domain is the incredible degree of conservation observed between species for the residue Q168. This residue is found in a region of high conservation in both yeast and human telomerases and recently it was demonstrated that a mutation to the equivalent Q169 residue in human TERT had a similar defect in RAP in human telomerase [31]. Since this glutamine is conserved between *T. thermophila* and humans, and mutations to this glutamine have analogous defects in *T. thermophila* and humans, it is likely that the glutamine acts by a similar mechanism in the two systems, namely by stabilizing short RNA-DNA duplexes in the telomerase active site.

The conservation of L14 between *T. thermophila* and human telomerase is less clear, however previous experiments demonstrated that a double mutation to leucines 13 and 14 in human telomerase has a severe activity defect, including at least a modest defect in RAP[14]. Residue F178 in *T. thermophila* telomerase does not appear to be strongly conserved, and it is less clear if it has an analogue in the human system. Nevertheless, when one considers the conservation of the other two TEN residues involved in the stabilization of short RNA-DNA duplexes and the similar activity defects observed between TEN mutants in *T. thermophila* and human telomerases, it appears that the mechanism of the TEN domain is likely conserved between species.

A recent smFRET study on human telomerase revealed a DNA dynamic equilibrium between two separate template annealing registers—pre- and post-translocation—for DNA primers corresponding to late catalytic intermediates [32]. However, DNA conformational changes in early catalytic intermediates corresponding to an analogous alternative state to the one described in the present work on *Tetrahymena* telomerase were not observed. It is possible that the distinct number of alignment residues in the *Tetrahymena* and human telomerase RNA templates may confer different levels of stability to the realigned RNA-DNA hybrid at the start of each NAP cycle. Therefore, the human enzyme may require TEN domain mutants to sufficiently destabilize the docked state in order to reveal an equilibrium between a docked state and an alternative state. Future experiments comparing FRET distributions between wild-type and TEN mutant enzymes in the

human enzyme will be valuable to determine if the TEN domain plays a conserved role in stabilization of short RNA-DNA duplexes in other telomerase systems.

REFERENCES

1 Palm, W. & de Lange, T. How shelterin protects mammalian telomeres. *Annual review of genetics* **42**, 301-334 (2008).

2 Bodnar, A. G. *et al.* Extension of life-span by introduction of telomerase into normal human cells. *Science* **279**, 349-352 (1998).

3 Vulliamy, T. J. & Dokal, I. Dyskeratosis congenita: the diverse clinical presentation of mutations in the telomerase complex. *Biochimie* **90**, 122-130 (2008).

4 Kim, N. W. *et al.* Specific association of human telomerase activity with immortal cells and cancer. *Science* **266**, 2011-2015 (1994).

5 Greider, C. W. Telomerase is processive. *Molecular and cellular biology* **11**, 4572-4580 (1991).

6 Blackburn, E. H. & Collins, K. Telomerase: an RNP enzyme synthesizes DNA. *Cold Spring Harbor perspectives in biology* **3** (2011).

7 Hammond, P. W., Lively, T. N. & Cech, T. R. The anchor site of telomerase from Euplotes aediculatus revealed by photo-cross-linking to single- and double-stranded DNA primers. *Molecular and cellular biology* **17**, 296-308 (1997).

8 Lue, N. F. A physical and functional constituent of telomerase anchor site. *The Journal of biological chemistry* **280**, 26586-26591 (2005).

9 Wyatt, H. D., Lobb, D. A. & Beattie, T. L. Characterization of physical and functional anchor site interactions in human telomerase. *Molecular and cellular biology* **27**, 3226-3240 (2007).

10 Jacobs, S. A., Podell, E. R. & Cech, T. R. Crystal structure of the essential N-terminal domain of telomerase reverse transcriptase. *Nature structural & molecular biology* **13**, 218-225 (2006).

11 Moriarty, T. J., Huard, S., Dupuis, S. & Autexier, C. Functional multimerization of human telomerase requires an RNA interaction domain in the N terminus of the catalytic subunit. *Molecular and cellular biology* **22**, 1253-1265 (2002).

12 Romi, E. *et al.* High-resolution physical and functional mapping of the template adjacent DNA binding site in catalytically active telomerase. *Proceedings of the National Academy of Sciences of the United States of America* **104**, 8791-8796 (2007).

13 Collins, K. & Greider, C. W. Tetrahymena telomerase catalyzes nucleolytic cleavage and nonprocessive elongation. *Genes & development* **7**, 1364-1376 (1993).

14 Zaug, A. J., Podell, E. R. & Cech, T. R. Mutation in TERT separates processivity from anchor-site function. *Nature structural & molecular biology* **15**, 870-872 (2008).

15 Eckert, B. & Collins, K. Roles of telomerase reverse transcriptase N-terminal domain in assembly and activity of Tetrahymena telomerase holoenzyme. *The Journal of biological chemistry* **287**, 12805-12814 (2012).

16 Robart, A. R. & Collins, K. Human telomerase domain interactions capture DNA for TEN domain-dependent processive elongation. *Molecular cell* **42**, 308-318 (2011).

17 Qi, X. *et al.* RNA/DNA hybrid binding affinity determines telomerase template-translocation efficiency. *The EMBO journal* **31**, 150-161 (2012).

18 Wu, R. A. & Collins, K. Human telomerase specialization for repeat synthesis by unique handling of primer-template duplex. *The EMBO journal* **33**, 921-935 (2014).

19 Jurczyluk, J. *et al.* Direct involvement of the TEN domain at the active site of human telomerase. *Nucleic acids research* **39**, 1774-1788 (2011).

20 Wu, J. Y., Stone, M. D. & Zhuang, X. A single-molecule assay for telomerase structure-function analysis. *Nucleic acids research* **38**, e16 (2010).

21 Berman, A. J., Akiyama, B. M., Stone, M. D. & Cech, T. R. The RNA accordion model for template positioning by telomerase RNA during telomeric DNA synthesis. *Nature structural & molecular biology* **18**, 1371-1375 (2011).

22 Akiyama, B. M. & Stone, M. D. Assembly of complex RNAs by splinted ligation. *Methods in enzymology* **469**, 27-46 (2009).

23 Selvin, P. R. H. T. *Single-molecule techniques : a laboratory manual.*, (2008).

24 Ha, T. *et al.* Single-molecule fluorescence spectroscopy of enzyme conformational dynamics and cleavage mechanism. *Proceedings of the National Academy of Sciences of the United States of America* **96**, 893-898 (1999).

25 Das, R., Laederach, A., Pearlman, S. M., Herschlag, D. & Altman, R. B. SAFA: semi-automated footprinting analysis software for high-throughput quantification of nucleic acid footprinting experiments. *Rna* **11**, 344-354 (2005).

26 McKinney, S. A., Joo, C. & Ha, T. Analysis of single-molecule FRET trajectories using hidden Markov modeling. *Biophysical journal* **91**, 1941-1951 (2006).

27 Roy, R., Hohng, S. & Ha, T. A practical guide to single-molecule FRET. *Nature methods* **5**, 507-516 (2008).

28 Bryan, T. M., Goodrich, K. J. & Cech, T. R. Tetrahymena telomerase is active as a monomer. *Molecular biology of the cell* **14**, 4794-4804 (2003).

29 Collins, K. & Gandhi, L. The reverse transcriptase component of the Tetrahymena telomerase ribonucleoprotein complex. *Proceedings of the National Academy of Sciences of the United States of America* **95**, 8485-8490 (1998).

30 Miller, M. C. & Collins, K. Telomerase recognizes its template by using an adjacent RNA motif. *Proceedings of the National Academy of Sciences of the United States of America* **99**, 6585-6590 (2002).

31 Wyatt, H. D., Tsang, A. R., Lobb, D. A. & Beattie, T. L. Human telomerase reverse transcriptase (hTERT) Q169 is essential for telomerase function in vitro and in vivo. *PloS one* **4**, e7176 (2009).

32 Parks, J. W. & Stone, M. D. Coordinated DNA dynamics during the human telomerase catalytic cycle. *Nature communications* **5**, 4146 (2014).

33 Mitchell, M., Gillis, A., Futahashi, M., Fujiwara, H. & Skordalakes, E. Structural basis for telomerase catalytic subunit TERT binding to RNA template and telomeric DNA. *Nature structural & molecular biology* **17**, 513-518 (2010).

SUPPLEMENTARY INFORMATION

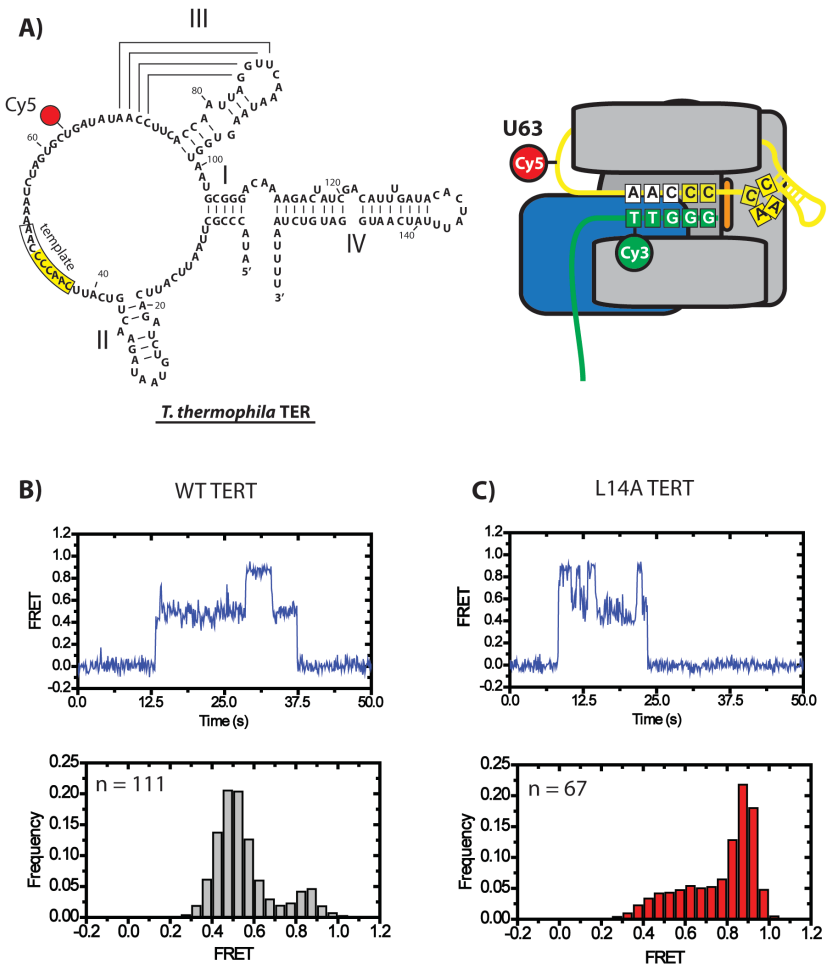


Figure S1. Telomerase labeled with an alternative labeling site demonstrates a similar distribution of states. A) (Left panel) Secondary structure of *Tetrahymena thermophila* TER with the position of the U63 Cy5 modification indicated. (Right panel) Cartoon model of the relative positions of the U63 Cy5 label and primer Cy3 label within the telomerase holoenzyme. **B)** Wild-type TERT was reconstituted with U63-labeled TER and tested in smFRET binding assays with $(TG)_8T_2G_3$ primers as in Figure 2. The top panel demonstrates a representative single molecule trace and the bottom panel represents a histogram compiled from 111 separate binding events. Both single molecule traces and histograms demonstrate a major population at ~ 0.50 FRET and a minor population at ~ 0.90 FRET. **C)** Representative single molecule trace (top) and smFRET histogram (bottom) for L14A TERT reconstituted with U63-labeled TER and Cy3-labeled $(TG)_8T_2G_3$ primers.

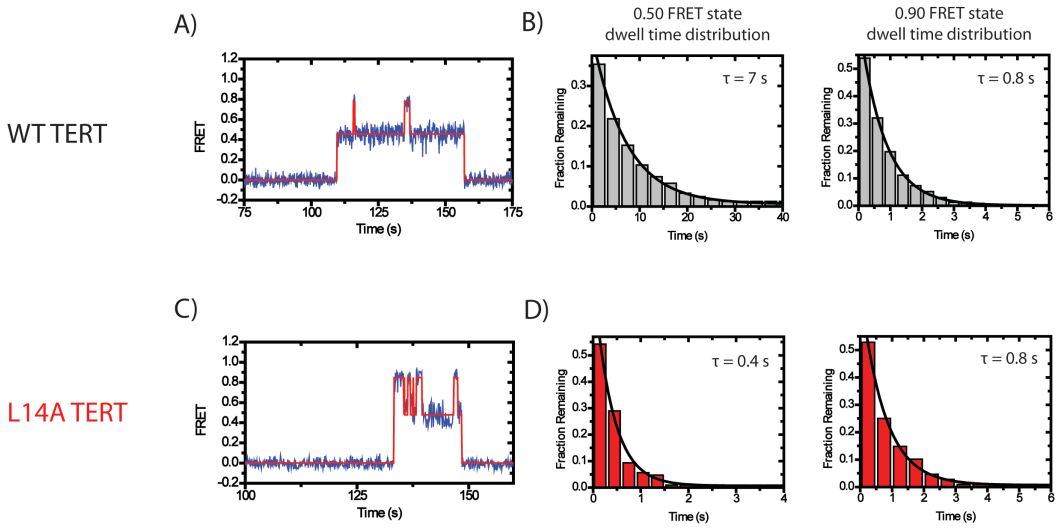


Figure S2. Dwell-time analysis using U63-labeled telomerase. A) U63-labeled TER was reconstituted with wild-type TERT and tested in the smFRET binding assay using $(TG)_8T_2G_3$ primers. smFRET traces (blue) were analyzed by HaMMy [1] to generate idealized traces (red). These were used to determine the dwell time of the enzyme in each state. **B)** The dwell times for the wild-type enzyme in the 0.50 FRET state and the 0.90 FRET state incubated with primer $(TG)_8T_2G_3$ were compiled into histograms. The histograms were fit to an exponential function to identify the average dwell time. Wild-type TERT demonstrated a dwell time of τ_{docked} = 7 s for the 0.50 FRET state and a dwell time of τ_{alt} = 0.8 s for the 0.90 FRET state. **C)** Representative smFRET trace and idealized HaMMy trace for L14A TERT telomerase incubated with the $(TG)_8T_2G_3$ primer. **D)** Compiled histograms for L14A enzyme. L14A TERT demonstrated a dwell time of τ_{docked} = 0.4 s for the 0.50 FRET state and a dwell time of τ_{alt} =0.8 s for the 0.90 FRET state. As observed for U36-labeled enzyme, the L14A mutant demonstrates a strong effect on the docked state, but has a minimal effect on the alternative state.

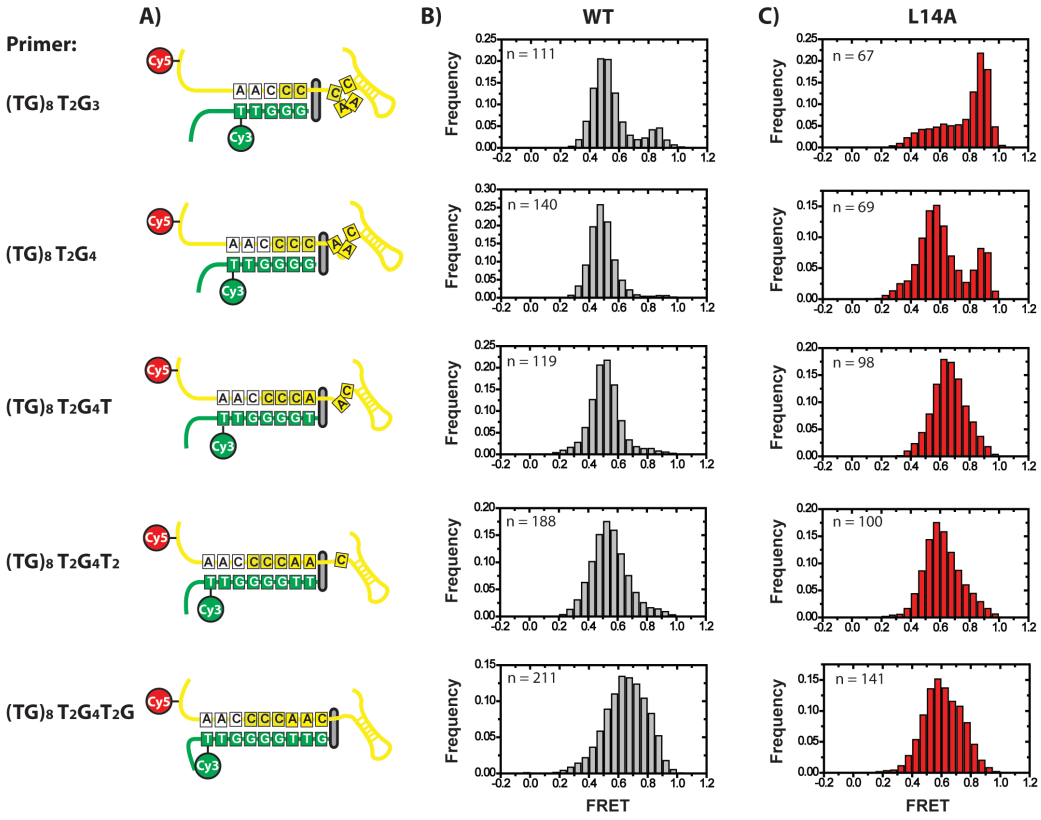


Figure S3. Effect of primer-template hybrid length on U63-labeled telomerase. Primers capable of forming 5-9 basepairs with template RNA were tested in smFRET telomerase binding assays with U63-labeled telomerase. **A)** Schematic diagram of the docked state for all six primers used in smFRET experiments, demonstrating the number of basepairs formed and positions of the Cy3 and Cy5 labels on the DNA primer and telomerase RNA, respectively. **B)** smFRET histograms for wild-type enzyme. As primers contain progressively more telomeric DNA sequence, the predominant FRET distribution of the docked state shifts from ~0.50 FRET to ~0.65 FRET. In addition, the ~0.90 FRET alternative state disappears. **C)** smFRET histograms for L14A mutant enzyme. The $(TG)_8T_2G_3$ primer is highly enriched for the alternative ~0.90 FRET state when compared against the wild-type enzyme. As the primer-template duplex is progressively extended, the L14A mutant FRET distribution increasingly resembles the wild-type distribution. The relative occupancies of the docked and alternative states in U63-labeled enzyme are highly consistent with the FRET distributions obtained with U36-labeled enzyme.

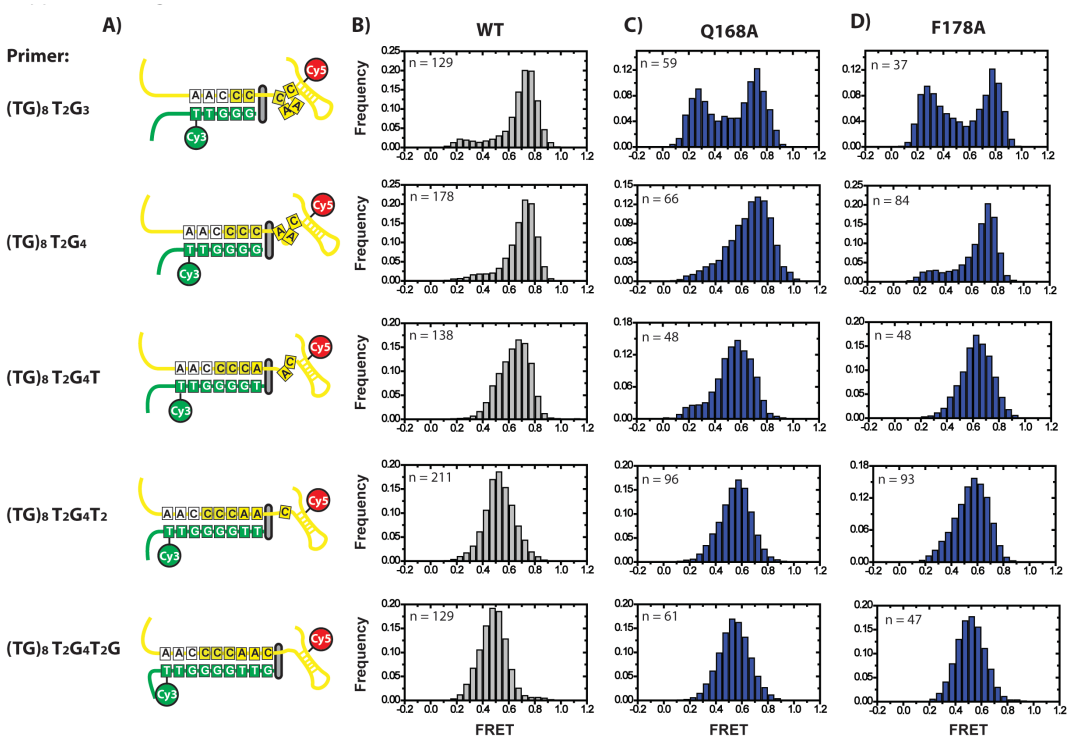


Figure S4. Effect of primer-template hybrid length on Q168A and F178A mutants. Primers capable of forming 5-9 basepairs with template RNA were tested in smFRET telomerase binding assays with U36-labeled telomerase. **A)** Schematic diagram of the docked state for all six primers used in smFRET experiments, demonstrating the number of basepairs formed and the expansion of the template RNA as the DNA-RNA duplex becomes progressively longer [2]. **B)** smFRET histograms for wild-type enzyme. As primers contain progressively more telomeric DNA sequence, the predominant FRET distribution of the docked state shifts from ~ 0.75 FRET to ~ 0.50 FRET. In addition, the ~ 0.25 FRET alternative state disappears. **C)** smFRET histograms for Q168A mutant enzyme. As shown in Figure 2, Q168A mutants demonstrate an enrichment for the alternative state with the $(TG)_8T_2G_3$ primer. As the primer-template hybrid length becomes longer, the Q168A mutant increasingly resembles the wild-type distribution. **D)** smFRET histograms for F178A mutant enzyme. F178A mutants demonstrate similar trends to Q168A mutants in their FRET distributions.

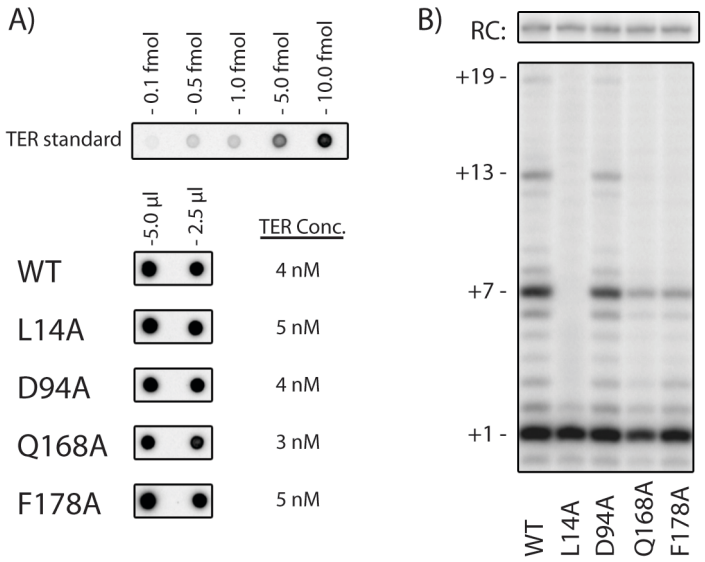


Figure S5. Quantification of immunopurified telomerase protein-RNA complexes. A) Wild-type and TEN domain mutant TERT was expressed in rabbit reticulocyte lysate and immunopurified by means of a FLAG tag on the N-terminus of TERT. Eluate from IP reactions was spotted on a blot and probed with a ^{32}P-labeled probe against TER to determine the concentrations of RNA that co-purified with TERT. 5 µl and 2.5 µl aliquots of the IP elution were compared against a TER standard. L14A, D94A, and F178A mutants demonstrated no detectible defect in the amount of TER that co-purified with TERT, whereas Q168A exhibited a modest defect in TER-TERT complexes. **B)** Telomerase primer extension assays on immunopurified telomerase. The material used in Figure S5A was used in telomerase primer extension assays using a $(GGGGTT)_3$ primer. As observed previously, L14A TERT exhibited a severe RAP defect and Q168A and F178A TERT also demonstrated RAP defects, whereas D94A telomerase activity largely resembled wild-type [3,4].

Supplementary References

1 McKinney, S. A., Joo, C. & Ha, T. Analysis of single-molecule FRET trajectories using hidden Markov modeling. *Biophys J* **91**, 1941-1951 (2006).

2 Berman, A. J., Akiyama, B. M., Stone, M. D. & Cech, T. R. The RNA accordion model for template positioning by telomerase RNA during telomeric DNA synthesis. *Nature structural & molecular biology* **18**, 1371-1375 (2011).

3 Zaug, A. J., Podell, E. R. & Cech, T. R. Mutation in TERT separates processivity from anchor-site function. *Nat Struct Mol Biol* **15**, 870-872 (2008).

4 Jacobs, S. A., Podell, E. R. & Cech, T. R. Crystal structure of the essential N-terminal domain of telomerase reverse transcriptase. *Nat Struct Mol Biol* **13**, 218-225 (2006).

CHAPTER VI

Integrated magnetic tweezers and single-molecule FRET for investigating the mechanical properties of nucleic acid.

*Submitted to Methods 2016

ABSTRACT

Many enzymes promote structural changes in their nucleic acid substrates via application of piconewton scale forces over nanometer length scales. Magnetic tweezers (MT) is a single molecule force spectroscopy method widely used for studying the energetics of such mechanical processes. MT permits stable application of a wide range of forces over long time scales, but suffers from relatively low spacial resolution in comparison to other single molecule manipulation methods. To overcome this limitation, researchers have integrated a spacially sensitive fluorescence spectroscopy method, single molecule-FRET, with MT to allow simulataneous observation and manipulation of nanoscale structural transitions. Here, we describe a method for using this hybrid instrument to analyze the mechanical properties of nucleic acids. We expect that this method for analysis of nucleic acid structure will be easily adapted for experiments aiming to interrogate the mechanical responses of other biological macromolecules.

INTRODUCTION

Recent developments in biological research have revealed that mechanical forces are critical modulators of molecular, cellular, and organismal function [1-3]. Initial biophysical characterization of DNA revealed that cellular processes such as

replication, transcription, and chromatin compaction rely on mechanical forces for proper biological function [4-10]. These studies revealed that enzymes convert chemical energy into pico-newton scale forces to modulate the twist, stretch, and bend properties of DNA.

Single molecule force spectroscopy has revolutionized how these force-induced structural transitions are characterized, permitting direct manipulation and analysis of individual DNA molecules [1,11-13]. Optical trapping (OT) and magnetic tweezers (MT) have emerged as the preferred methods for characterizing biologically relevant forces by allowing precise calibration of force and torque[14,15]. Although these techniques access similar force regimes, each possess certain advantages and limitations. For example, a well-aligned and vibration-isolated OT system can achieve a remarkable degree of spatial resolution, capable of detecting molecular displacements with sub-nanometer resolution [16-18]. By comparison, the lower trap stiffness of MT typically provides lower spatial resolution, but allows application of very low stretching forces in comparison to OT [12].

This difference in accessible stretching forces leads to a dichotomy for ideal MT and OT experiments. In both methodologies, typically long, flexible DNA handles are used as mechanical transducers for stretching and torqueing forces [19]. Increased handle length provides magnification of biological scale motions, but also introduces noise into spacial positioning for short time-scale measurements. For this reason, shorter timescale experiments are typically performed with OT, where high stretching forces stiffen the system and improve the signal-to-noise ratio [12,16,17]. When

lower forces are necessary, MT is the method of choice and provides accurate spacial positioning through time averaging of the data. This requirement limits low force MT experiments to analysis of long timescale transitions to achieve maximal spatial resolution.

In order to access high spacial and temporal resolution at relatively low forces, several research groups have employed hybrid force-fluorescence spectroscopy methods, which decouple the application of stretching force from the measurement of conformational rearrangements within a structure of interest (SOI). Typically the nanometer-scale rearrangements within the SOI are detected by single-molecule Förster Resonance Energy Transfer (smFRET), a time-resolved fluorescence method that is highly sensitive to distance changes on the nanometer scale[20,21]. In a smFRET experiment, a donor fluorophore is directly excited with a laser (i.e. 532 nm laser for cy3 excitation). If an acceptor dye is in close spacial proximity to the donor (2-8 nm), dipolar coupling of the dyes allows non-radiative energy transfer from the excited donor to the acceptor, where closer proximity results in a higher degree of energy transfer [22]. Experimentally, FRET is defined by the relationship $I_A/(I_A + I_D)$, where I_A and I_D are the measured acceptor and donor intensities, respectively. In this way, simultaneous measurement of donor and acceptor fluorescences can provide a sensitive metric for conformational changes within individual molecules.

Hybridization of smFRET with both OT and MT systems has been reported [23-28]; however, the integration of MT with smFRET has certain practical advantages. OT

experiments are achieved through a high-power trapping laser which has the potential to induce photo-damage of the smFRET probes. Although this challenge can be overcome by inter-lacing the trapping laser with the smFRET excitation source using high frequency shutters or acousto-optical modulators (AOM)[29], implementing such a system is expensive and technically challenging. In contrast, magnetic force application in MT is ideally-suited for integration with standard objective-type total internal reflection fluorescence (TIRF) microscopy[30,31], without concerns of induced photo-damage by a trapping laser. The relatively simple integration of an existing MT system with a TIRF microscope should enable biochemistry laboratories to establish a force-FRET assay 'in house', rather than exclusively relying on collaborative efforts.

Here, we will review the MT-smFRET system for study of nucleic acid structures. First, we will describe, in detail, the design of the MT-smFRET instrument. Next, we will address experimental protocols involved in MT-FRET analysis. Finally, we will discuss successful applications of MT-smFRET in studying the mechanical properties of DNA structure.

1. Magnetic Tweezers- Single molecule FRET microscope

2.1 Design and construction of a hybrid MT-smFRET microscope

The basic layout of a MT-smFRET instrument is illustrated in Figure 1. The hybrid device is built on a standard objective-type total internal reflection fluorescence (TIRF) microscope equipped with an electron multiplying charged coupled device (CCD) camera with single-molecule detection sensitivity. The excitation lasers are coupled through the back port of the microscope and guided to

the sample through an objective. This objective is mounted onto a piezo-controlled stepping device, which is later used for force calibrations and bead tracking. Emitted light from the fluorescent probes is collected by the same objective, spectrally separated by dichroic mirrors to generate donor and acceptor channels, and subsequently focused onto the CCD. In general, this objective-based TIRF setup closely resembles the systems described in detail in other reviews[30,31].

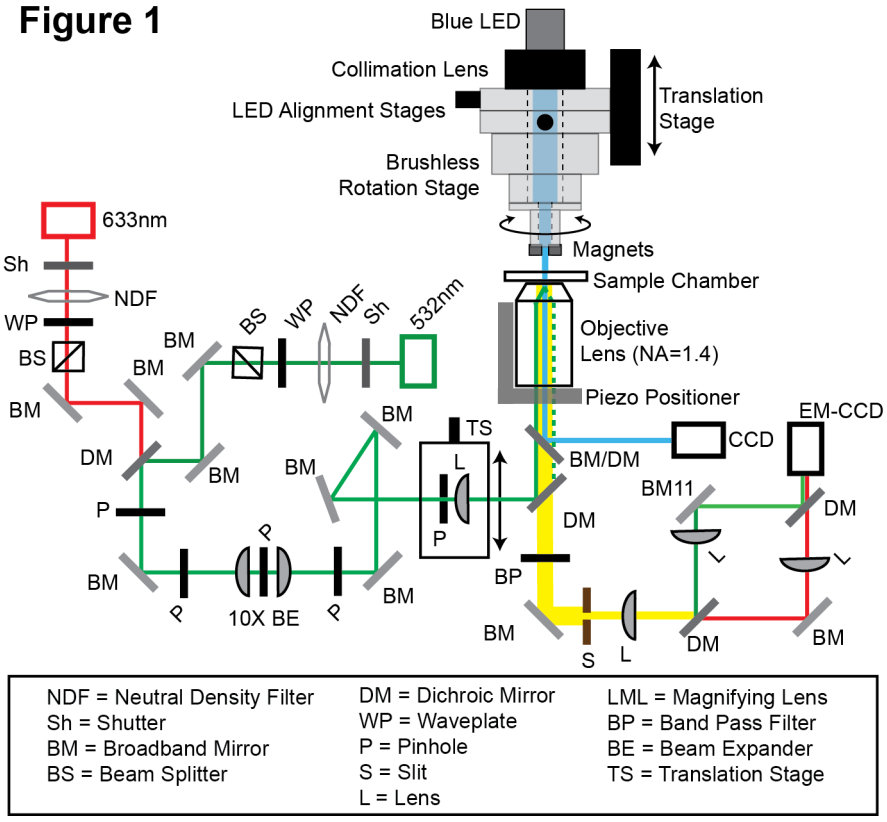


Figure 1. **Hybrid MT-smFRET microscope setup.** Schematic diagram of the MT-smFRET instrument. The MT-FRET instrument consists of an objective-type TIRF microscope, smFRET detection optics, and an EM-CCD camera. The excitation lasers are coupled through the back port of the microscope and guided to the sample through an oil-immersion objective. The magnetic tweezers apparatus, which controls the linear and angular positions of the magnets, is mounted above the sample chamber. The collimated blue LED light source is mounted directly onto the magnet assembly and is used for direct imaging of the magnetic beads.

Simultaneous measurement of molecular extension and smFRET is possible by integrating additional imaging components. A second, camera is mounted to a sideport of the microscope to detect the diffraction pattern, and thereby spacial positioning, of a molecule-bound magnetic bead. This pattern is generated by overhead illumination from a collimated LED (both blue and near IR LEDs have been used), which can be separated from fluorescence light by a dichroic mirror in the microscope. Spectral separation of the bead image from the fluorescence light provides simultaneous measurement of molecular extension and smFRET, respectively.

Objective-type TIRF microscopy provides an ideal geometry for integrating smFRET with a MT system. The magnetic tweezer apparatus resides above the sample chamber and consists of a pair of permanent magnets mounted onto computer controlled translation and rotation stages. The motorized stage assembly is used to change the linear position and angular displacement of the magnets to apply specific degrees of tension and twist, respectively. A more detailed description of the magnet control system is discussed in a variety of excellent reports [6,32].

2.2 Calibration of applied force

In a MT experiment, the force being applied to a DNA molecule can be directly extracted by measuring the length of the DNA tether and the amplitude of beads lateral fluctuations in the x-y plane (Figure 2a). The methods used for tracking the height of the tethered bead from the surface (z-position), as well as tracking the x-

and y-positions of the bead in real-time have been described[33,34], and therefore will not be described in detail in this review.

In the context of an integrated MT-smFRET experiment, it is desirable to determine a systematic calibration of the applied stretching force as a function of a magnet height (Figure 2b). This generalized calibration allows force estimations for any DNA tether without performing individual force calibrations on each molecule. One critical assumption made by this approach is that the magnetic beads used in the experiments are highly uniform with respect to their iron content. We have found the one micron (MyOne) beads from Dynal to be sufficiently uniform so as to provide force estimates with an error of maximally 10% among different beads. This level of accuracy is sufficient for most Force-smFRET experiments. If more precise calibration of the force is required, bead fluctuation and tether length data can be collected for each individual molecular setup allowing direct calibration of each molecule. Accuracy of the force calibration is typically assessed by fitting the force-extension curve to the well-established worm like chain model for DNA elasticity (Figure 2c)[35,36]. When larger forces (> 15 pN) are required, 2.8 micron magnetic beads must be employed. For these experiments, we find the size distribution of the beads to be too large to use a general calibration, and therefore forces are recalibrated for each DNA molecule following each MT-smFRET measurement. A detailed description of the required measurements and methods for force calibration has been discussed elsewhere[15].

Figure 2

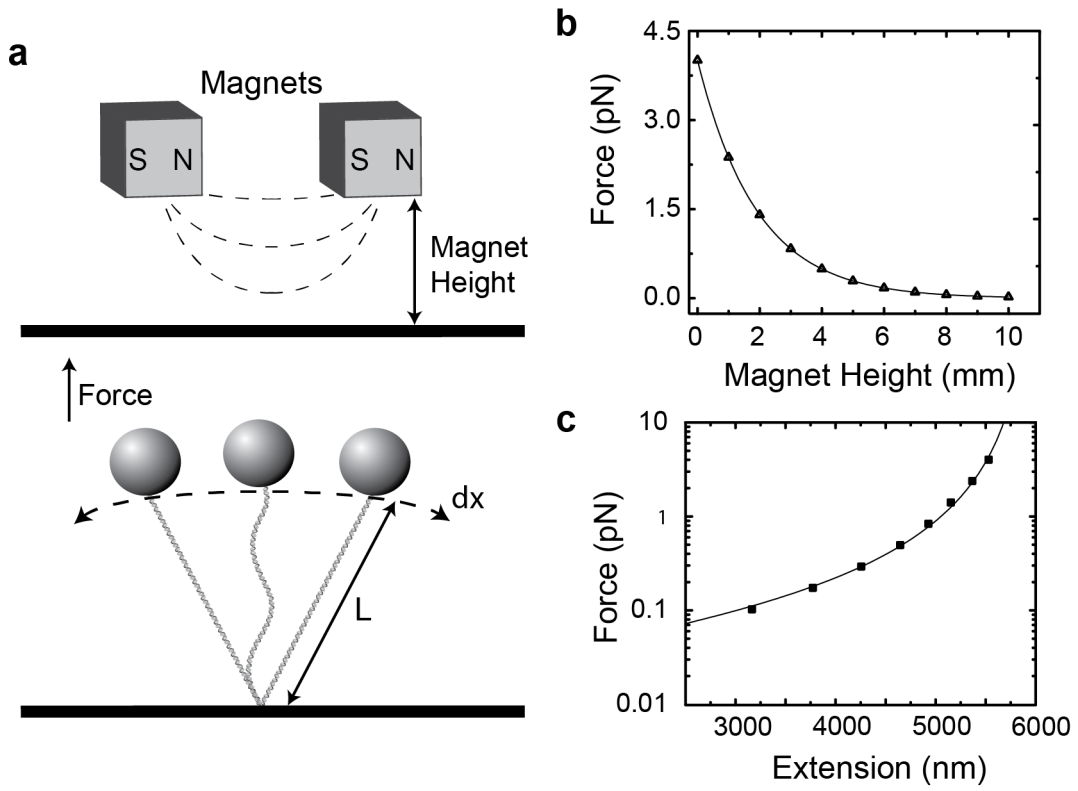


Figure 2. Magnetic tweezers force calibration. (a) Schematic illustration of a DNA molecule tethered between a coverglass and a magnetic beads. Forces are measured according to the expression $F = kTL/\langle \delta x^2 \rangle$, where F is the stretching force, k is Boltzman's constant, T is temperature, and δx is the amplitude of the bead fluctuation in the x-plane. The stretching force can be changed by varying the position of the magnet assembly held above the sample chamber. (b) Measured stretching forces in pico-Newtons (pN) as a function of magnet position in millimeters (mm) are plotted and fit by a single exponential decay function. The parameters of the exponential fit are then used to convert magnet position to force values across different experiments. (c) Experimental verification of force calibration by fitting a force-extension plot of a DNA molecule with the worm-like chain model, yielding the expected persistence length of 45 nm for DNA under our experimental conditions.

METHODS

In the following sections, we describe the preparation of sample chambers and magnetic beads, the construction of FRET dye-labeled molecule, the immobilization of the molecule, and MT-smFRET measurement.

3.1 Preparation of sample chambers and magnetic beads

As with all MT experiments, the DNA molecules to be studied must be specifically attached between a microscope coverglass and a magnetic bead. Sample chambers are prepared by sandwiching two pieces of parafilm between a standard glass microscope slide and a coverglass (Figure 3a). We note that if higher forces are required, the 1 mm thick microscope slide may be replaced with a standard coverglass, permitting a closer approach of the magnet assembly. Prior to chamber assembly, the microscope slide is cleaned using a plasma-cleaner to prevent non-specific sticking of DNA and beads. Since this surface is not imaged during the experiment, additional cleaning steps on the microscope slide are not necessary. In contrast, the coverglass is extensively cleaned by a series of alternating sonication treatments in a strong base and acetone to reduce fluorescence background. After assembly, the chambers can be stored for up to three weeks before use. Directly before the experiment, channels are treated with biotinylated bovine serum albumin (BSA), washed with buffer, and then incubated with streptavidin (Figure 3b). Following the streptavidin incubation, flow channels are washed again in a suitable buffer, and stored in a humidity chamber to prevent evaporation. An example of a typical channel assembly is described as follows:

Figure 3

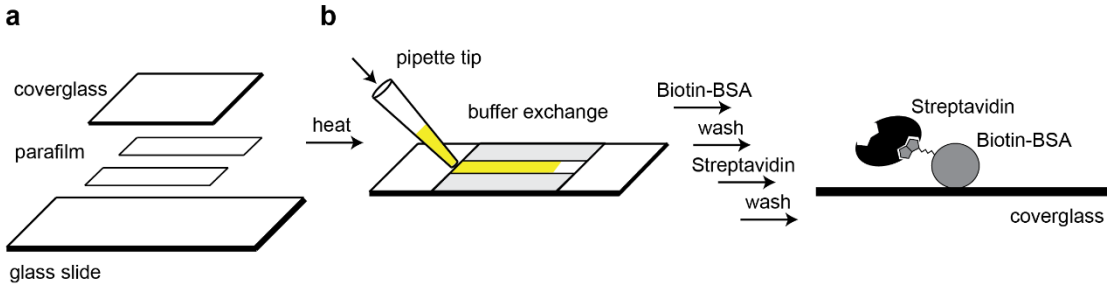


Figure 3. Preparation of sample chambers. (a) Sample chambers are assembled by sandwiching two strips of parafilm between a coverglass and a glass slide. To adhere the parafilm to the glass, the sample chamber is briefly heated and then allowed to cool to room temperature. (b) To prepare the cleaned and assembled sample chamber for an experiment, a series of buffer exchange steps are performed using a pipette. First a solution of biotinylated-BSA is introduced which adheres non-specifically onto the coverglass. After washing away the unbound biotin-BSA, a solution of streptavidin is flowed over the surface. Following a final wash step to remove excess streptavidin, sample chambers are stored in a humidity chamber to prevent evaporation.

a. Sonicate the coverglasses to be imaged in a glass staining jar for 20 min in each solution: 10% w/v Alconox, acetone, 1M KOH, and Milli-Q H_2O.
b. To begin sample chamber assembly, start by drying one of the cleaned coverglasses with nitrogen gas.
c. Place coverglass on top of a clean quartz slide.
d. Quickly flame the imaging surface of the coverglass using a handheld propane torch to burn away any fluorescent debris.
e. After plasma-cleaning the microscope slide for 5 min, assemble sample chambers by sandwiching strips of parafilm between the microscope slide and the coverglass.

f. Place the assembly beneath a heat block (95^{o}C) for ~ 10 seconds to adhere the parafilm to the glass (parafilm will become partially clear).

g. Once the assembled slide cools to room temperature, pipette 35µL of 1mg/mL biotinylated BSA into the sample chamber and incubate for 5 min.

h. Wash the chamber with ~200 µL T50 (10 mM Tris pH 8, 50 mM NaCl) buffer, followed by incubation with 50µL of 0.2 mg/mL streptavidin (made in T50 buffer).

i. Wash the chambers with ~200 µL T50 buffer to remove excess streptavidin.

j. Store the equilibrated slides in a humidity box to avoid evaporation. A humidity box can be made by filling the bottom of an empty pipette tip box with some water and covering tightly with lid.

The most common magnetic beads used in MT experiments are 1 micron or 2.8 micron diameter Dynal beads. As mentioned above, the 1 micron bead is designed for experiments using low stretching forces (<15pN). In contrast, the 2.8 micron bead can exhert forces >15pN. These magnetic beads come with a wide variety of functionalized coatings, which give the flexibility to tether DNA molecules with different antibodies. For our smFRET-MT experiments, we typically use anti-digoxigenin-coated magnetic beads, derived from carboxlated beads as described [27].

3.2 Preparation of smFRET-labeled DNA molecules

The molecules employed in a hybrid MT-smFRET experiment are minimally comprised of three separate fragments: (i) a segment that harbors the SOI, together with appropriately placed smFRET probes and a biotin modification for surface

immobilization; (ii) a relatively long DNA handle to space the large magnetic bead away from the imaging surface; and (iii) a linker DNA fragment that contains digoxigenin-modified dUTP for association with anti-digoxigenin coated magnetic bead (Figure 4). An example of a molecule preparation strategy employed in our laboratory is outlined below:

Figure 4

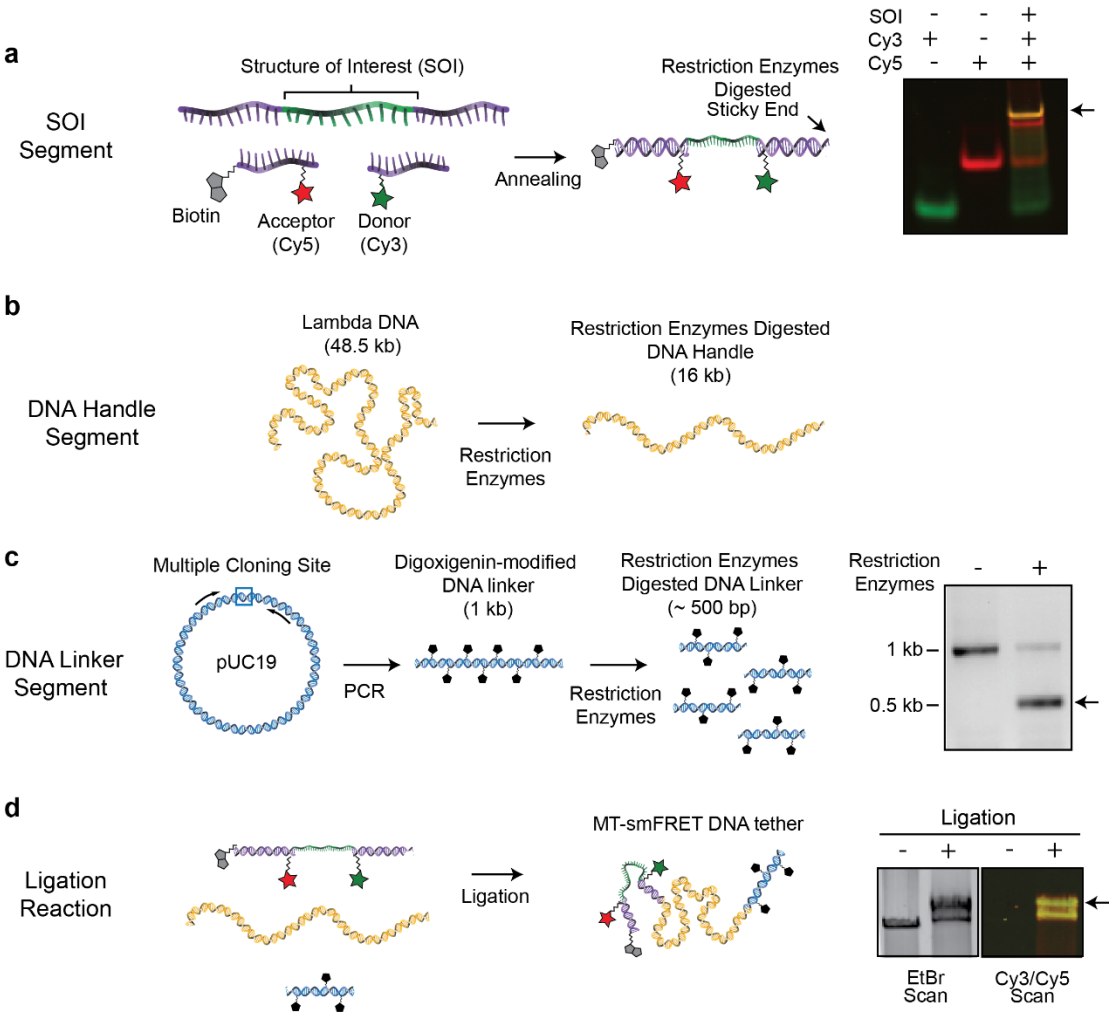


Figure 4. Preparation of DNA molecules for MT-smFRET measurement. The DNA molecule for MT-smFRET measurement is constructed from three different DNA segments. (**a**) The segment harboring the structure of interest (SOI) labeled

with smFRET dyes. The efficiency of the annealing reaction is determined by native polyacrylamide gel and taking a fluorescence scan for the donor (Cy3) and acceptor (Cy5) dyes. The yellow band (small black arrow) represents the target product of the annealing reaction possessing all three oligonucleotides. (**b**) The DNA handle segment is prepared by restriction enzymes digestion of lambda DNA. The desired basepair segment is purified on an agarose gel. (**c**) The DNA linker segment is prepared by PCR reaction with a plasmid containing multiple cloning site template using primers flanking the multiple cloning site in the presence of digoxigenin-11-dUTP. (right) The digoxigenin modified 1 kb PCR reaction product is digested with restriction enzymes to obtain ~500 DNA linker segment (small black arrow) and analyzed on a DNA agarose gel. (**d**) To assemble the final MT-smFRET DNA molecule, a three segment DNA ligation reaction including the SOI, DNA linker, and DNA handle segments is performed. Ligation efficiency is determined by comparing the ethidium bromide (EtBr) stained gel image of the ligation products with a fluorescence scan for Cy3 and Cy5. The largest band that is yellow (small black arrow) represents the target product possessing a DNA linker, DNA handle, and SOI segment.

(i) SOI construction: The DNA oligonucleotides used to construct the SOI segment are generated synthetically with site-specific amino modifications. These modifications allow site-specific dye coupling through standard succinimide chemistry, followed by reverse phase HPLC purification of labeled and unlabeled DNA fragments.

a. Dissolve amino-modified DNA oligonucleotides to be dye labeled (50 nmol) in 30μL Milli-Q H_2O and transfer to a 1.5 ml of microcentrifuge tubes. Ethanol precipitate to remove residual chemicals from DNA synthesis. Take care to use sodium acetate rather than ammonium acetate during precipitation as residual ammonium salt will quench the amino coupling dye labeling reaction.

b. Dissolve the mono-reactive dye packs (GE Cy3 or Cy5) with 50 μL of 0.2M $NaHCO_3$, combine with 50 μL of the appropriate resuspended DNA oligonucleotide, and incubate at 37 °C for 2 h.

c. Precipitate the dye labeled DNAs and dry pellets.

d. Purify labeled from unlabeled DNA fragments by reverse-phase HPLC with a C8 column as described[37].

e. Mix the dye-labeled and SOI DNA oligonucleotide in T50 Buffer in 1.5 μL microcentrifuge tube in a final volume of 30 μL, yielding a final concentration of 0.1 μM of each DNA oligonucleotide.

f. Anneal the dye labeled SOI segments by heating the solution to 95°C for 4 min, followed by slow cooling to room temperature on the benchtop.

(ii) Handle construction: The large magnetic bead introduces a substantial amount of background fluorescence if it is not sufficiently far away from the imaging surface in the TIRF microscope. Therefore, one can use a large DNA fragment generated by digesting commercially available lambda DNA with restriction endonucleases (Figure 4b). Using a 0.6 % w/v agarose gel, this lambda fragment can be purified and used for MT-smFRET DNA molecule construction.

(iii) Linker construction: A DNA fragment containing digoxigenin modifications is also created in order to specifically immobilize the final molecule to the

magnetic bead. This fragment is generated by performing a polymerase chain reaction (PCR) across a plasmid DNA containing a multiple cloning site (MCS) in the presence of digoxigenin-11-dUTP Figure 4c). Primers are designed to amplify a DNA fragment with a selected endonuclease restriction site positioned approximately in the middle of the fragment. The digoxigenin-modified PCR product is then cut with a restriction enzyme to yield a sticky-end linker fragment for subsequent ligation to the DNA handle. For experiments aiming to study torque-induced structural changes in DNA, it is necessary to torsionally constrain the DNA molecule between the coverglass and the magnetic bead. One can create an additional linker fragment containing biotin-modified dUTP for immobilization at the slide surface.

(iv) Ligation Assembly: The final step of the molecule preparation is to perform a multi-component ligation reaction of the (i) SOI segment, (ii) the purified lambda DNA handle, (iii) and the digoxigenin-modified DNA linker fragment (1:1:1 molar ratio of the three fragments at a final concentration of 0.3 μM). Ligations are carried out using manufacturer protocols (NEB). The efficiency of the ligation reaction can be determined by running a diagnostic agarose gel, and comparing an ethidium bromide stained image with a fluorescence scan for the smFRET dyes (Figure 4d).

3.3 Sample Immobilization

As ligation reactions are not 100% efficient, optimization is necessary to minimize the presence of unwanted side products or unreacted material in the sample chamber (Figure 5). First, an empirically determined dilution of the ligation reaction is incubated with anti-digoxigenin coated magnetic beads, pre-blocked with BSA to avoid nonspecific DNA binding (Figure. 5a). The beads are next washed during a magnetic bead pull-down step to remove any DNA fragments that were not ligated to an anti-digoxigenin DNA linker fragment (Figure. 5b). The washed beads bound with DNA are next introduced onto the streptavidin coated coverglass (see Figure. 3) and allowed to settle to the surface. Following this incubation period, non-immobilized molecules are washed out of the sample chamber using gravity-controlled buffer flow. This immobilization strategy greatly enriches for those beads that are specifically bound to the streptavidin coverglass and the anti-digoxigenin coated magnetic bead (Figure. 5c). An example of our immobilization protocol is included below:

Figure 5

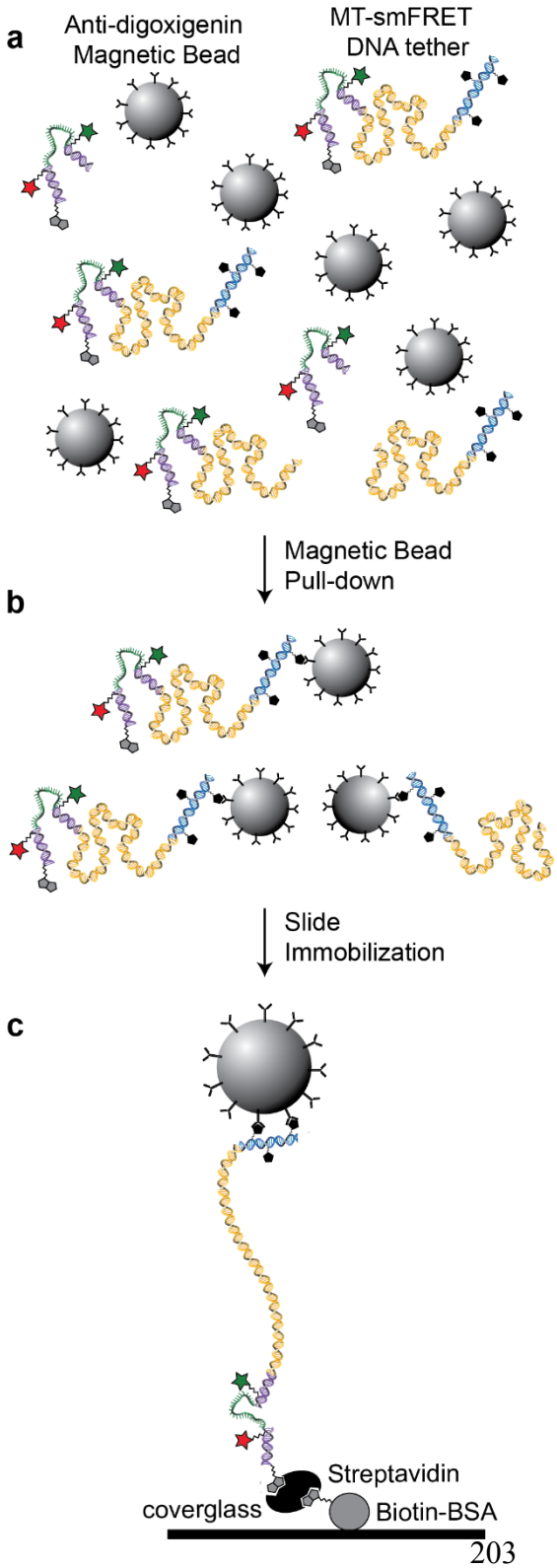

a
Anti-digoxigenin
Magnetic Bead
MT-smFRET
DNA tether
Magnetic Bead
Pull-down
b
Slide
Immobilization
c
Streptavidin
coverglass
Biotin-BSA

Figure 5. **Surface immobilization of DNA molecules for MT-smFRET measurement.** (**a**) Products of DNA ligation reactions are diluted and incubated with pre-blocked anti-digoxigenin coupled magnetic beads. (**b**) The DNA-magnetic bead mixture is placed on a magnetic separation rack to pull-down only those DNA tethers attached specifically to the antibody coupled magnetic beads via a digoxigenin moiety, while non-digoxigenin modified DNA is washed away at this step. (**c**) The DNA-magnetic bead mixture is introduced into a sample chamber and immobilized on the coverglass via a biotin-streptavidin linkage. The sample chamber is then gently washed with imaging buffer using gravity controlled buffer exchange to remove any unbound beads. This process greatly enriches for DNA tethers that are site-specifically oriented between the coverglass and magnetic bead.

a. Wash 50 μL anti-digoxigenin-coated magnetic beads from the stock with 100 mM NaCl and 20 mM Tris (pH 8.0) buffer.
b. Block the anti-digoxigenin-coated magnetic beads with 500 μL of 5 mg/mL BSA, 50 mM NaCl and 10 mM Tris (pH 8.0) buffer and at 4°C for at least 1 hour.
c. Dilute the DNA ligation 1:5 with 1X T4 DNA ligase buffer.
d. Incubate 6 μL of diluted DNA ligation with 3 μL of blocked anti-digoxigenin-coated magnetic beads at 4°C for 10 min. Gently agitate the DNA-magnetic bead mixture every 2 min.
e. Place the DNA-magnetic bead mixtures on a magnetic separation rack. Carefully pipette out the supernatant. Do not allow beads to dry completely.
f. Resuspend DNA-bound magnetic beads with 200 μL blocking buffer
g. Store the DNA-bound magnetic beads in 15 μL blocking buffer at 4°C for up to one day.
h. Mix 6 μL of DNA-bound magnetic beads with 194 μL of blocking buffer.

i. Deposit 100 μL of DNA-magnetic beads in blocking buffer to the streptavidin-coated chamber.
j. Incubate the DNA-magnetic bead mixture at room temperature for 30 min in a humidity chamber with the coverglass side faced down to allow the DNA-bound magnetic beads to settle to the surface by gravity.
k. Gently wash the chamber with 100 μL of equilibration buffer by gravity to remove the unattached DNA-linked magnetic beads.

3.4 MT-smFRET measurement

An MT-smFRET measurement starts by first identifying a DNA molecule of the appropriate length. This process can be achieved using the eye piece and manual focus knob of the microscope. Once a candidate molecule is found, quality control checks are performed on the molecule before data collection can begin. The first quality control utilizes the rotation stage to check whether the molecule is torsionally constrained. If the molecule is not torsionally constrained, the length of the DNA tether is not affected by rotating the magnets. The length change for a torsionally unconstrained molecule indicates the bead is tethered to the surface by more than one DNA molecule. In the case of experiments utilizing a torsionally constrained DNA tether, the length of the molecule should change as turns are introduced into the molecule. In this instance, one can use the differential response of positively and negatively supercoiled DNA to stretching force as a check for the presence of a single DNA tether[38]. The second quality control is to determine whether the molecule has an active FRET pair by directly exciting both fluorophores. Once the presence of an

active smFRET pair has been verified, data is then acquired at the user-determined frame rate and desired stretching force until irreversible photo-bleaching of one of the smFRET dyes is observed. The execution of a typical single molecule experiment is described in detail below:

a. Prepare oxygen scavenging system (Gloxy) and imaging buffer. Gloxy solution is prepared by mixing 100 mg/mL of glucose oxidase, 4 mg/mL of catalase, 50 mM NaCl and 10 mM Tris (pH 8.0) buffer. Imaging buffer includes 50 mM Tris pH 8.0, 100 mM NaCl, 0.4% w/v D-glucose, and is then saturated with Trolox and filtered twice through a 0.2 micron syringe filter
b. Prepare imaging buffer with oxygen scavenging system by mixing 2 μL of Gloxy solution with 200 μL of imaging buffer.
c. Gently flow 100 μL of imaging buffer plus oxygen scavenging system into the chamber by gravity. Secure the prepared slide on the microscope.
d. Move the magnet several millimeters above the slide to apply sufficient stretching force to extend the DNA molecule (~2-4 pN). Then use the low magnification side port CCD camera and video monitor to locate a bead that appears to be fluctuating the expected distance from the surface (longer tethers exhibit larger fluctuations). Since the DNA tethers for our MT-smFRET experiments are long (~5 microns) this qualitative screening step is readily achieved by eye.
e. Next, apply positive or negative torque into molecule to check if the bead is anchored by a single DNA molecule. The extension of a single MT-smFRET

DNA tether that is torsionally unconstrained should not change as turns are introduced. Again, changes in the z-position of the bead are readily detected by eye during this step. The same check can be performed when utilizing a torsionally constrained DNA tether. However in this instance, one can use the differential response of positive and negative supercoiling to check for the presence of a single DNA tether[38].

f. After modifying the optics for observation of the FRET dyes, use the red (633 nm) laser to verify the presence of an active Cy5 dye located directly under the magnetic bead. Since the dye molecule is located ~5 microns away from the magnetic bead, it is necessary to focus the objective down toward the surface to observe the fluorescence signal.

g. Next, use the green (532 nm) laser to verify the presence of an active Cy3 molecule located directly under the magnetic bead.

h. Once the presence of an active smFRET pair has been verified, set the magnet assembly to the desired linear position, acquire movie (typically 10-30 frames per second) until dyes photo-bleach.

2. Applications of MT-smFRET

The first experiment combining MT with smFRET was used to dissect the force-induced stretching of single-stranded DNA. In this study, Shroff et al. built a smFRET-based force sensor consisting of a single stranded DNA attached with donor and acceptor dye molecules. They established the relationship between the applied stretching forces and smFRET efficiency (Figure 6b). Using this correlation, they

were able to determine the internal force sensed by a single-stranded DNA during partial conversion to double-stranded DNA.

Using a similar setup, research in our laboratory characterized the mechanical unfolding pathway of human telomere G-quadruplexes (GQs) [27]. In this experiment, a MT system was used to apply mechanical force to perturb the folding dynamics of telomere GQs, whereas smFRET reported on the structural transitions in a single telomere GQ (Figure 6c). By analyzing the effect of the stretching force on GQ folding and unfolding, we determined the physical distance to the transition state for the GQ folding pathway. These measurements revealed that the GQ is a relatively brittle DNA structure that readily unfolds after slight structural deformation.

Figure 6

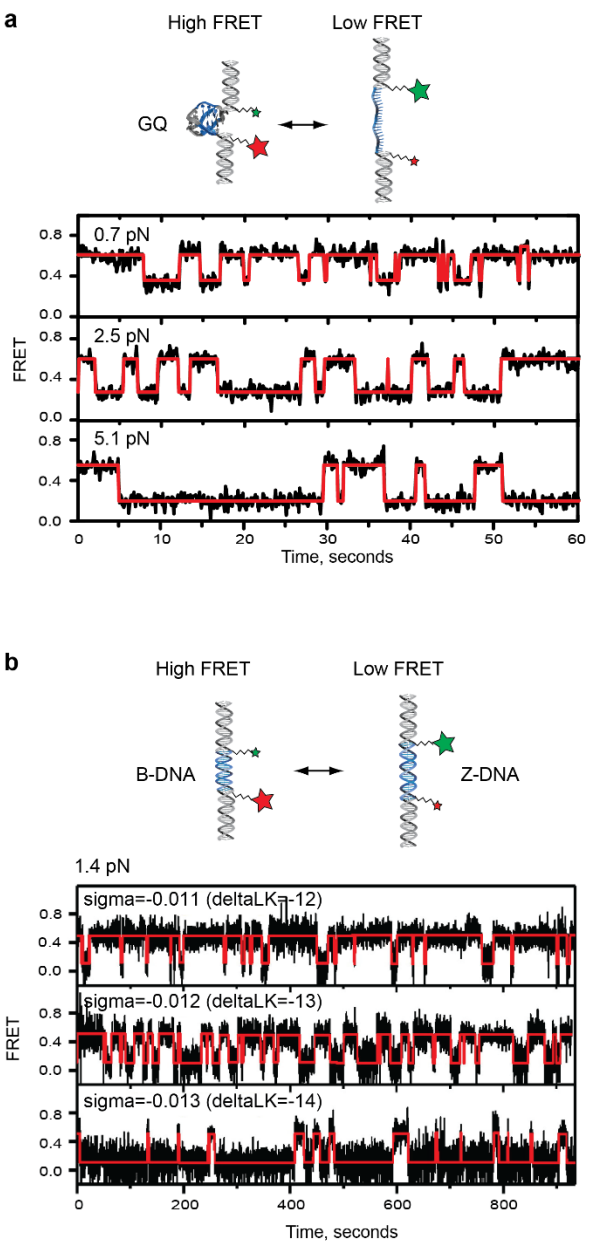


Figure 6. Examples of MT-smFRET experiments. (**a**) smFRET trajectories of a single GQ molecule flanked by donor and acceptor dyes held under three different forces (0.7, 2.5 and 5.1 pN). At low forces, the molecule spends the majority of the time in the high-FRET (folded) state. In contrast, a higher force (5.1 pN) substantially shifts the telomere DNA GQ folding equilibrium toward the low-FRET (unfolded) state. smFRET trajectory is plotted together with the idealized FRET trace (red)

generated by hidden Markov modeling using HaMMy fitting software (this figure is modified from Long et al. NAR) . (**b**) smFRET trajectory of a DNA molecule containing the GC core DNA flanked by donor and acceptor dyes held under 1.4 pN with three different superhelical density (-0.011, -0.012 and -0.013). When the molecule is underwound by magnetic tweezers, the FRET efficiency changes from a high to a low value, indicating the transition from the B-form to the Z-form. Single-molecule FRET trajectories are fitted with a hidden Markov model yielding idealized FRET traces (red) (this figure is adapted from Lee et al. PNAS).

The MT-smFRET instrument not only enables the detection and application of stretching forces on DNA molecules, but also can be used to study the response of DNA molecules under torque. Taking advantage of this application, Lee et al. detected the B-to-Z DNA transition under torsion and tension. In this study, MT was used to apply torque to duplex DNA molecules containing GC repeats (Figure 6d). By simultaneously tracking the smFRET efficiency from the donor and acceptor dyes flanking the GC repeats, as well as changes in the overall extension of the supercoiled DNA tether, they directly observed the B-to-Z conformational change under minute negative superhelicity at low tension, which highlights the highly cooperative and dynamic nature of the B-Z transition.

These studies provided unique insights into the mechanical properties of DNA structures, with implications for the mechanisms of DNA binding proteins that remodel nucleic acid structure, and demonstrated the general utility of the integrated MT-smFRET assay.

CONCLUSION

In summary, here we describe the MT-smFRET method to analyze mechanical properties of nucleic acids. The integration of these single-molecule

methods permits measurement of mechanically induced structural transitions with sub-nanometer spatial resolution over a wide range of stretching forces and torques. A nucleic acid SOI is embedded within a long DNA molecule, flanked by smFRET probes, and oriented between a microscope slide and a magnetic bead. In this way, stretching forces and torques allow physical manipulation of the SOI, while smFRET provides a method of interrogating the SOI structure and conformation. We anticipate that the MT-smFRET study of nucleic acid structure will be easily adapted for experiments aiming to interrogate the mechanical response of other biological macromolecules and systems.

REFERENCES

1 Brenner, M. D., Zhou, R. & Ha, T. Forcing a connection: impacts of single-molecule force spectroscopy on in vivo tension sensing. *Biopolymers* **95**, 332-344 (2011).

2 Bustamante, C., Chemla, Y. R., Forde, N. R. & Izhaky, D. Mechanical processes in biochemistry. *Annu Rev Biochem* **73**, 705-748 (2004).

3 Dulin, D., Lipfert, J., Moolman, M. C. & Dekker, N. H. Studying genomic processes at the single-molecule level: introducing the tools and applications. *Nature reviews. Genetics* **14**, 9-22 (2013).

4 Abels, J. A., Moreno-Herrero, F., van der Heijden, T., Dekker, C. & Dekker, N. H. Single-molecule measurements of the persistence length of double-stranded RNA. *Biophys J* **88**, 2737-2744 (2005).

5 Dawid, A., Croquette, V., Grigoriev, M. & Heslot, F. Single-molecule study of RuvAB-mediated Holliday-junction migration. *Proc Natl Acad Sci U S A* **101**, 11611-11616 (2004).

6 Kim, K. & Saleh, O. A. A high-resolution magnetic tweezer for single-molecule measurements. *Nucleic Acids Res* **37**, e136 (2009).

7 Kruithof, M. *et al.* Single-molecule force spectroscopy reveals a highly compliant helical folding for the 30-nm chromatin fiber. *Nature structural & molecular biology* **16**, 534-540 (2009).

8 Lionnet, T., Joubaud, S., Lavery, R., Bensimon, D. & Croquette, V. Wringing out DNA. *Phys Rev Lett* **96**, 178102 (2006).

9 van Loenhout, M. T., van der Heijden, T., Kanaar, R., Wyman, C. & Dekker, C. Dynamics of RecA filaments on single-stranded DNA. *Nucleic Acids Res* **37**, 4089-4099 (2009).

10 Stone, M. D. *et al.* Chirality sensing by Escherichia coli topoisomerase IV and the mechanism of type II topoisomerases. *Proc Natl Acad Sci U S A* **100**, 8654-8659 (2003).

11 Bustamante, C., Bryant, Z. & Smith, S. B. Ten years of tension: single-molecule DNA mechanics. *Nature* **421**, 423-427 (2003).

12 Neuman, K. C. & Nagy, A. Single-molecule force spectroscopy: optical tweezers, magnetic tweezers and atomic force microscopy. *Nature methods* **5**, 491-505 (2008).

13 Smith, S. B., Cui, Y. & Bustamante, C. Overstretching B-DNA: the elastic response of individual double-stranded and single-stranded DNA molecules. *Science* **271**, 795-799 (1996).

14 Smith, S. B., Finzi, L. & Bustamante, C. Direct mechanical measurements of the elasticity of single DNA molecules by using magnetic beads. *Science* **258**, 1122-1126 (1992).

15 Yu, Z. *et al.* A force calibration standard for magnetic tweezers. *Rev Sci Instrum* **85**, 123114 (2014).

16 Abbondanzieri, E. A., Greenleaf, W. J., Shaevitz, J. W., Landick, R. & Block, S. M. Direct observation of base-pair stepping by RNA polymerase. *Nature* **438**, 460-465 (2005).

17 Moffitt, J. R., Chemla, Y. R., Izhaky, D. & Bustamante, C. Differential detection of dual traps improves the spatial resolution of optical tweezers. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 9006-9011 (2006).

18 Carter, A. R., Seol, Y. & Perkins, T. T. Precision surface-coupled optical-trapping assay with one-basepair resolution. *Biophysical journal* **96**, 2926-2934 (2009).

19 Neuman, K. C., Lionnet, T. & Allemand, J. F. Single-molecule micromanipulation techniques. *Annual Review of Materials Research* **37**, 33-67 (2007).

20 Ha, T. Single-molecule fluorescence methods for the study of nucleic acids. *Curr Opin Struct Biol* **11**, 287-292 (2001).

21 Roy, R., Hohng, S. & Ha, T. A practical guide to single-molecule FRET. *Nature methods* **5**, 507-516 (2008).

22 Ha, T. *et al.* Probing the interaction between two single molecules: fluorescence resonance energy transfer between a single donor and a single acceptor. *Proc Natl Acad Sci U S A* **93**, 6264-6268 (1996).

23 Lang, M. J., Fordyce, P. M., Engh, A. M., Neuman, K. C. & Block, S. M. Simultaneous, coincident optical trapping and single-molecule fluorescence. *Nature methods* **1**, 133-139 (2004).

24 Zhou, R. *et al.* SSB functions as a sliding platform that migrates on DNA via reptation. *Cell* **146**, 222-232 (2011).

25 Hohng, S. *et al.* Fluorescence-force spectroscopy maps two-dimensional reaction landscape of the holliday junction. *Science* **318**, 279-283 (2007).

26 Lee, M., Kim, S. H. & Hong, S. C. Minute negative superhelicity is sufficient to induce the B-Z transition in the presence of low tension. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 4985-4990 (2010).

27 Long, X., Parks, J. W., Bagshaw, C. R. & Stone, M. D. Mechanical unfolding of human telomere G-quadruplex DNA probed by integrated fluorescence and magnetic tweezers spectroscopy. *Nucleic acids research* **41**, 2746-2755 (2013).

28 Shroff, H. *et al.* Biocompatible force sensor with optical readout and dimensions of 6 nm3. *Nano letters* **5**, 1509-1514 (2005).

29 Comstock, M. J., Ha, T. & Chemla, Y. R. Ultrahigh-resolution optical trap with single-fluorophore sensitivity. *Nature methods* **8**, 335-340 (2011).

30 Axelrod, D., Burghardt, T. P. & Thompson, N. L. Total internal reflection fluorescence. *Annu Rev Biophys Bioeng* **13**, 247-268 (1984).

31 Axelrod, D. Total internal reflection fluorescence microscopy in cell biology. *Methods Enzymol* **361**, 1-33 (2003).

32 Vilfan, I. D., Lipfert, J., Koster, D. A., Lemay, S. G. & Dekker, N. H. Magnetic Tweezers for Single-Molecule Experiments. *Handbook of Single-Molecule Biophysics*, 371-395 (2009).

33 Lipfert, J., Hao, X. & Dekker, N. H. Quantitative modeling and optimization of magnetic tweezers. *Biophysical journal* **96**, 5040-5049 (2009).

34 Cnossen, J. P., Dulin, D. & Dekker, N. H. An optimized software framework for real-time, high-throughput tracking of spherical beads. *Review of Scientific Instruments* **85** (2014).

35 Bustamante, C., Marko, J. F., Siggia, E. D. & Smith, S. Entropic elasticity of lambda-phage DNA. *Science* **265**, 1599-1600 (1994).

36 Bouchiat, C. *et al.* Estimating the persistence length of a worm-like chain molecule from force-extension measurements. *Biophysical journal* **76**, 409-413 (1999).

37 Akiyama, B. M. & Stone, M. D. Assembly of complex RNAs by splinted ligation. *Methods in enzymology* **469**, 27-46 (2009).

38 Strick, T. R., Allemand, J. F., Bensimon, D. & Croquette, V. Behavior of supercoiled DNA. *Biophys J* **74**, 2016-2028 (1998).

APPENDIX I

MT-smFRET Instrument Control, Data Acquisition, and Real-time Analysis

ABSTRACT

This chapter provides a detailed description of the Magnetic Tweezer-single molecule FRET (MT-smFRET) instrument software. Within the chapter are subsections that describe how to properly modify the software, a description of the front end functions of each VI, and a detailed programmatic description of two key software components: real-time single molecule bead tracking and alternating laser excitation (ALEX). All algorithms were written in LabVIEW to interface directly with previously existing software.

REAL-TIME SINGLE MOLECULE BEAD TRACKING

1) General software usage and overview of control programs

1.1) Backup process before modifying the existing software

Currently, when the "Start Magnetic Tweezers" button is pressed on the desktop, a batch file is run that opens the different magnetic tweezer control programs: Andor Camera control, Bead Panel Tracker, and Motion Control. All of these files are located in C:\Andor\ folder. These files should never be modified directly as they are used by the general user. The files that can be modified for testing purposes are in the "Public Share\MT Software\Beta Version" directory. If the proper syncing protocols are taken (see below), these files should be identical to execution files on the C drive (Figure 1.1). Before modifying any of these files, the entire directory (on the C drive) should be copied to the server directory "Public Share\MT

Software\Old Version\20xx-xx-xx" and "Public Share\MT Software\Beta Version", where the new folder in "Old Version" contains the current date. This backup process ensures that the current working software has a hard backup so the previous software can be reinstated in case any fatal errors are made during the editing process. Once backed up, any VIs in the "Beta Version" folder can be opened and modified. The control software on the "Public Share" can be run directly over the network to test the functionality of the code. Once the new modifications are working properly, the "Beta Version" software can overwrite the Current Version folder on the Public Share. The Current Version folder is present as a reference directory for the Current Software being used on the Magnetic Tweezers C drive. Modifications should only be made to this folder once the new beta software version has been thoroughly tested. When the Current Version folder is modified, the entire directory should overwrite the filed on the Magnetic Tweezers "C:\Andor\" folder through manual copying or the installed program SyncToy.

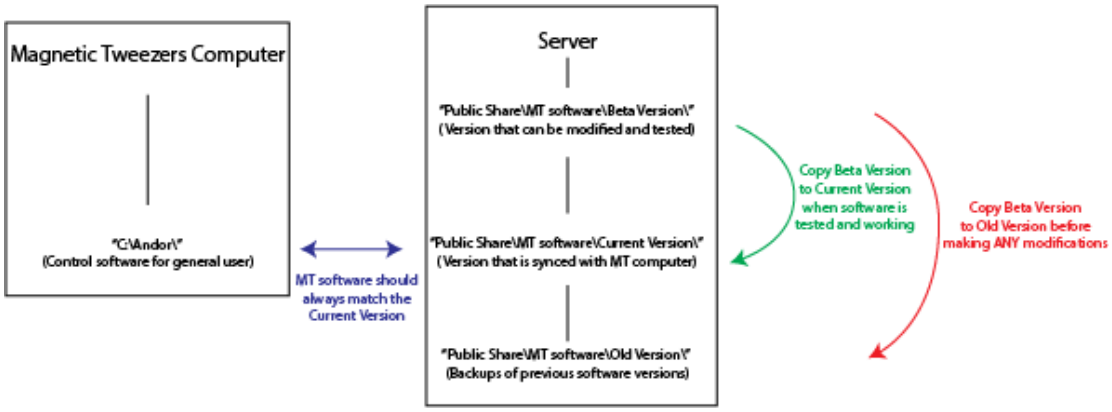


Figure 1.1. MT control software version architecture. The general user software resides on the MT computer C drive. The server contains a Beta Version folder which should be copied to a new Old Version folder before being modified and tested. Once

the Beta Version is tested and accepted, the newly modified files can overwrite the Current Software folder which is then synced with the MT computer.

1.2) Bead tracking integrated with the Andor Camera software

When the new bead tracking software was written, the previous front end control system and control file architecture was maintained in order to preserve usability for users and editors. Similarly, rewriting the file architecture would have been laborious and extremely involved.

1.3) Andor_887_beat.vi – Camera control and image acquisition

The Andor software was written before my time in the lab and specifically was written to control a smFRET microscope operating with an Andor camera. The code has been slightly modified to allow real-time tracking of single molecule tethers rather than solely capturing video data. This allows real-time manipulation and analysis of MT experiments and also reduces the size of the output file dramatically as image data is no longer saved. To access the real-time tracking mode, the Collection Mode pulldown must be changed to "Save Dat File" (Figure 1.2). While in this mode, the image data is sent to BeadPanelTracker-Math.vi (Section 1.5) where bead tracking can be controlled in real-time. The only other feature that was added in the MT version of andor is the ability to collect magnet vertical position data during movie collection mode. This was implemented to allow video capture of MT-smFRET traces while simultaneously writing the magnet position to a separate text file (Figure 1.2).

Figure 1.2. Andor Camera Control Software (andor_887_beta). Updated components of the control program include the Collection Mode choice and Save Magnet During Movie. Collection mode contains two types of acquisition, Save Movie File, which saves a raw image data file, and Save Dat File, which sends the image data to BeadPanelTracker-Math for real-time bead position analysis.

1.4) BeadPanelTracker.vi – General bead tracking control

The "BeadPanelTracker" vi is the front end control VI that the user interacts with during real-time tracking (Figure 1.3). In order to operate this panel, the "Andor Beta" control software must be running and collecting data in "Save Dat File" mode

(Figure 1.2). This software mode sends the image data for real-time bead tracking to the BeadPanelTracker-Math VI. In the BeadPanelTracker panel, left clicking on the real-time image will move the Reference cursor toward the mouse, which should be directed toward the center of the Reference bead. Right clicking the mouse moves the Tether cursor which should be directed toward the center of the Tether bead (Figure 1.3).

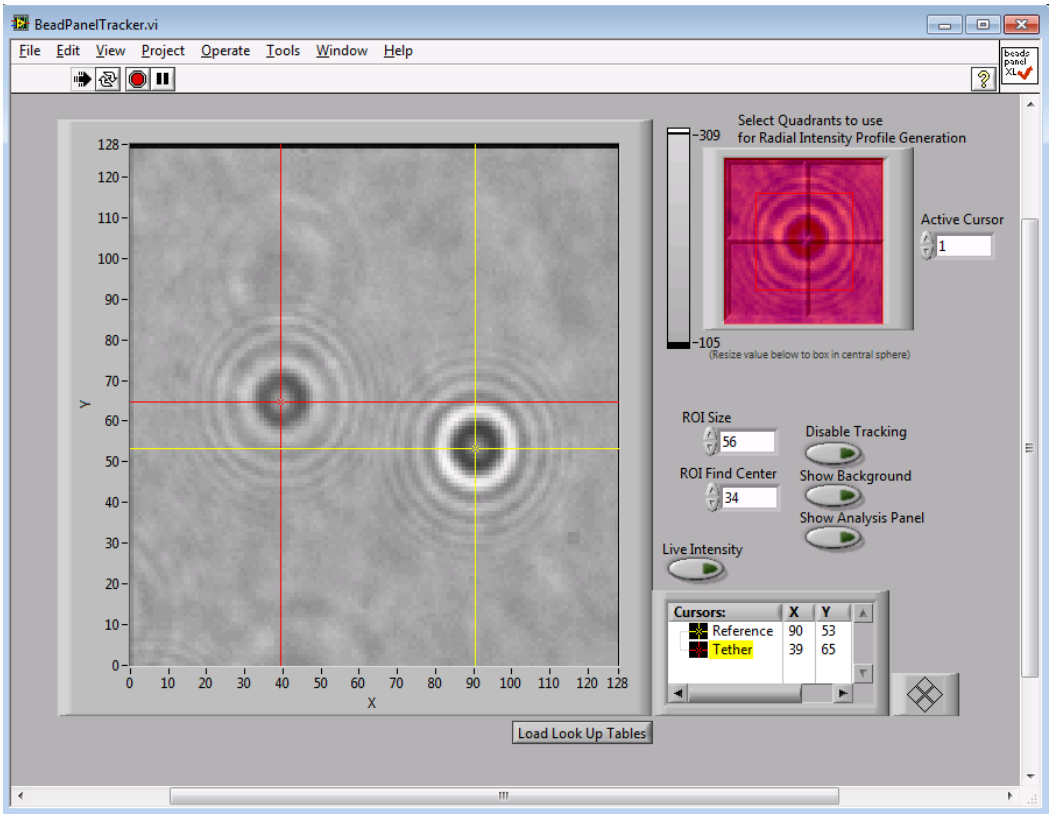


Figure 1.3. Real-time imaging and tracking front end control (BeadPanelTracker.vi). In this VI the user can direct x, y, and z tracking of the beads through clicking the camera image and modifying the ROI values.

Efficiency and accuracy of tracking for each bead is governed by two front-end parameters: ROI Size and ROI Find Center (ROI: Region of Interest, see Figure 1.3). ROI Size determines how much of the image is being used for z-position

tracking of the bead. The default value of 56 is typically a good choice. When modifying the ROI Size, a good rule of thumb is to make the ROI image (upper right Figure 1.3) fit all the ring pattern of the bead within the ROI when bead image is largest. Note that this parameter SHOULD NOT BE MODIFIED after calibration of z-position (see below). Modifying this parameter to a different number after calibration will cause issues with z-position tracking of tether molecules. ROI find center value determines the red box size in the ROI panel and defines the region for centroid tracking. For the most efficient tracking of the bead center, this box should be made to be slightly larger than the first major ring in the bead pattern at the largest bead pattern. A good example of a well-chosen ROI size can be seen in Figure 1.3. Although this value can be modified at any time during the experiment, ROI size is sufficiently compliant that no modification is typically necessary in order to achieve robust centroid tracking.

The other buttons in the BeadPanelTracker are instated for usability purposes (Figure 1.3). First, the Active Cursor button changes which bead is being represented in the upper right ROI panel. This allows modification of the ROI size and ROI find center parameters during visual inspection of either bead. Disable Tracking turns off the centroid tracking functionality, which can help prevent buffer overflow issues when searching for molecules and tracking is not needed. Show Background brings forward the BeadPanelTracker-Math panel which is the VI that controls all of the reading, writing, and analysis for real-time bead tracking. Show Analysis Panel brings forward the beadpositon VI, which includes many of the graphs that allow

visualization of the captured Z-position data and the associated subroutines. This panel is essential for visualizing data when performing real-time z-position tracking. Live intensity is a button solely used for visualizing intensity data during MT-smFRET experiments at the specified cursor position. Lastly, the Load Lookup Table button loads a set of predefined lookup tables for bead tracking purposes. This process is essentially a substitute for the calibration process (discussed later), but will only give a rough estimate of molecular extension as each set of beads should be recalibrated on a per molecule basis. In essence, this functionality is solely useful for the molecular search process so you can find a suitable tether without having to first calibrate.

1.5) BeadPanelTracker-Math.vi – Computational hub for bead tracking

BeadPanelTracker-Math is a subroutine of the andor software that performs x, y, and z tracking of bead images collected by the Andor control vi. When the Andor software is run in "Save Dat File" mode, the for loop inside the camera software is run in image capture mode and transfers the collected images to the BeadPanelTracker-Math vi (Figure 1.4). Inside this vi, the image data is processed into individual frames and then run through two subroutines, find_bead_center and beadposition, which find the bead x/y position and z position, respectively. The results of these tracking algorithms are then processed and written to file. Before the intricacies of the image processing is discussed, I will first describe the front end control that a typical user uses.

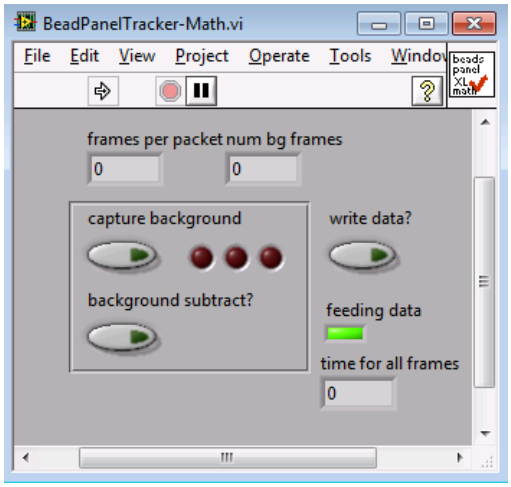


Figure 1.4. Data manipulation and master tracking VI (BeadPanelTracker-Math.vi). In this VI are two main functionalities, capturing and applying the background correction and writing the data to file. This VI contains subVIs that perform x, y, and z tracking.

When a user clicks the "Show background panel" in the BeadPanelTracker vi, the BeadPanelTracker-Math vi is opened and serves two main functions: background subtraction and saving tracking data to file. Background subtraction removes the background data from the image (for example if dust is caught on the objective, mirrors, etc). To remove the background, reposition the stage so that the imaging field contains no beads. Then click and hold the capture background button for approximately two seconds (Figure 1.4). Release the button and then press the background subtract? button to remove the background image during experimental acquisition. The write data? button will initiate writing x, y, and z tracking data into a tab delimited file with 14 columns. Each row represents a frame of tracking data, and the columns represent the listed values: magnetic rotational position, objective piezo position (nano-stepper motor), the linear stage position of the magnet (magnet height above slide), bead x position, bead y position, bead z position, and current force

applied to the bead. These seven values are recorded for the reference bead as the first seven columns of data. The last seven columns of data represent the corresponding values for the magnetic tether bead. It remains unclear why the program was initially written to save the instrument parameters in a doubly redundant fashion, but removing the redundancy would modify the output file structure for subsequent analysis programs, therefore the current architecture was maintained.

1.6) MotionControl.vi – Control of the magnets and the piezo motor-mounted objective

The original magnet control software (developed by Michael Stone) was modified to include two new functionalities: bead tracking calibration and automated magnet rotation. The bead tracking calibration is a general software feature that must be used for all real-time z-position tracking experiments. Automated rotation is a specific algorithm that integrates z-position signal with an automated rotation mechanism. This portion of the program is only useful when you have a repetitive signal, like nick/ligation events.

Proper execution of the bead calibration protocol is paramount for accurate z-position tracking of your reference and tether beads. The main portion of the MotionControl software that is responsible for bead position calibration is the NanoDrive section (Figure 1.5). This region of the software controls the piezo stepper-motor that the microscope objective is mounted upon. When the MotionControl program is started, the NanoDrive current position should read around 50 in the Current Position box. Changing the Nano Set Point to a different value and

clicking Set Position will move the NanoDrive to a new, well-defined position, effectively changing the image focus. Depending on the length of the tether molecule, your calibration parameters in the NanoDrive will change. The Current Position of the NanoDrive should always be set to 50 before starting. The Num Steps parameter changes how many calibration points you use and the Step Size sets how much distance (in micrometers) is between each calibration point. Together, these two parameters define your calibration range where Num Steps multiplied by Step Size will provide the effective calibration range. A good rule of thumb is to use a calibration range three times your molecular extension with a Step Size of 0.1 microns. So if the tether molecule is one micron long, then the Step Size is set to 0.1 microns and the Num Steps should be set to 30. For longer molecules, the step size can be increased to speed up the tracking algorithm as higher spacial resolution is not required. Step Time should be set to 100 minimally, otherwise the stepper motor will have trouble keeping up with the calibration procedure. Higher values are acceptable, but not required as empirical testing of these values produced similar calibrations. If the calibration should be saved to file (as in a real experiment) the Generate Lookup Table LED should be set to “on”. Once these parameters are set, the calibration is initiated by clicking Start Calibration. During this time, the user should not perform any other processes or instrument manipulations.

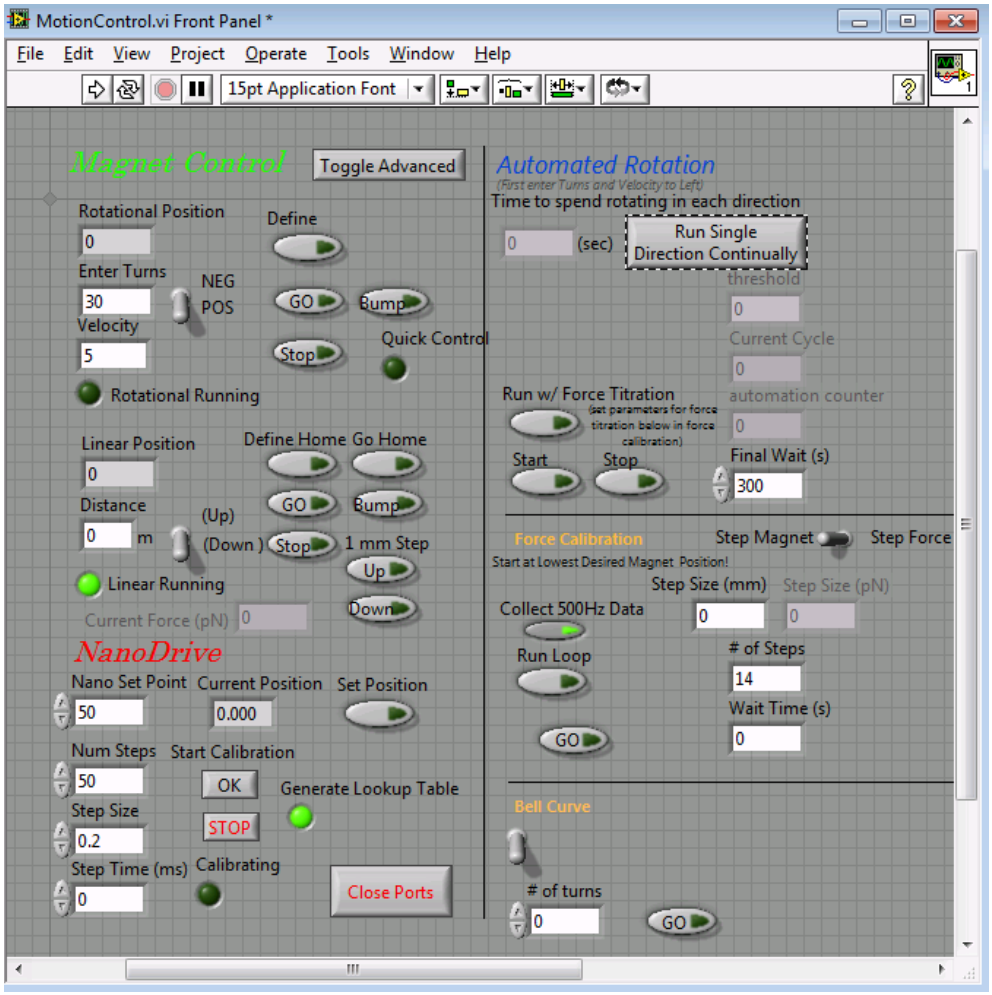


Figure 1.5. Magnet and objective stepper motor control system (MotionControl.vi). Motion control contains controls (left) for magnet height above slide, magnet rotation, and piezo-stepper height for objective positioning. The functions on the right are automated control mechanisms including: automated rotational control, automated magnet height stepping for force calibration, and bell curve analysis of torsionally constrained tethers.

2) Real-time tracking algorithm design, implementation, and work-flow

2.1) andor_887_beta – collects image data and forwards to the real-time tracking system

The andor control software was written by a former lab member of Michael Stone, therefore the intricacies of camera control and image capture will not be

discussed in this chapter. Despite this, there is an important interface between the andor software and BeadPanelTracker-Math delineated herein. The functionality of video recording has been modified for the “run ‘til abort” mode of the andor image collection (mode label, upper right, Figure 1.2). Originally the software would collect images from the camera buffer, save those images to file, and repeat until a user specified command. The small modification introduced allows a user to select a feature Camera Mode as “Save Movie File” or “Save Dat File” (bottom, Figure 1.2). Programmatically this change was introduced through a conditional statement inside the image capture/data write loop (Figure 2.1).

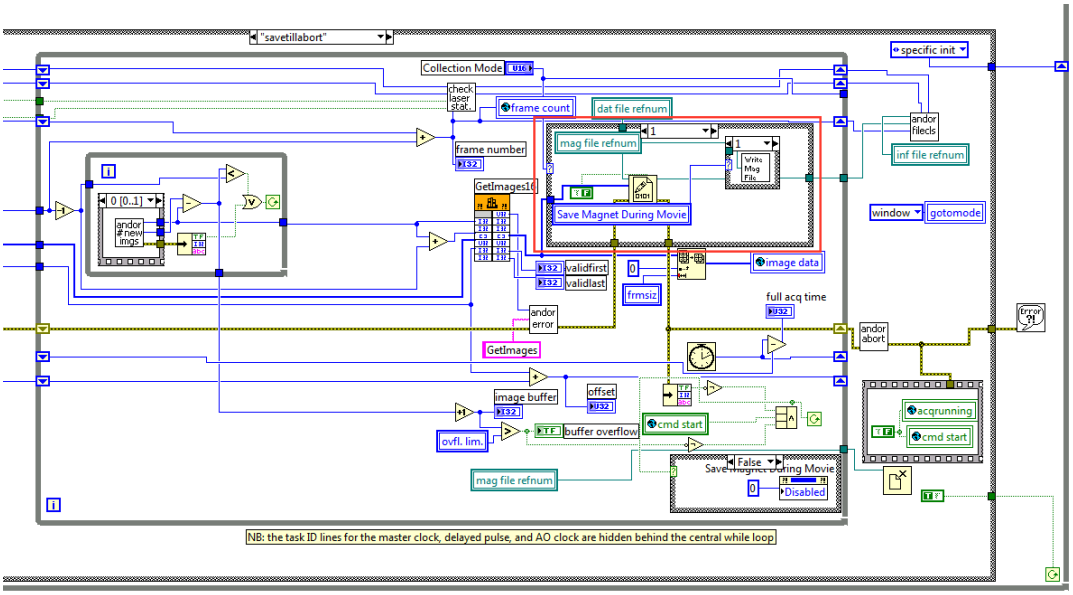


Figure 2.1. Modified Andor Software allows movie collection or bead tracking. In the main image collection loop, the conditional option “savetilabort” is the main image acquisition code that will run for most user experiments. Highlighted in a red box is the modified code that allows switch between collection modes “Save Movie” and “Save Dat”.

Choosing “Save Movie File” will bypass the real-time analysis software and directly write the image data to file (Figure 2.2). The file refnums are pointer

variables to the file references that are opened upon starting data collection. In the case of Movie file collection, the image data is written to file automatically but the user also has the choice to “Write Mag File” which is a subVI routine that writes the magnet position to file. This option is really only meaningful when in “Save Movie File” collection mode. Selecting this option creates a reference file for Force-FRET experiments so that the FRET data can be correlated to changes in force throughout an individual movie file.

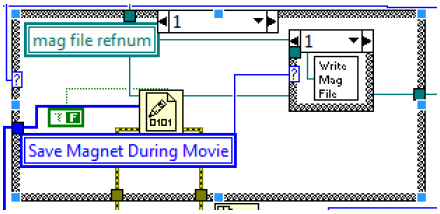


Figure 2.2. Save Movie File Mode. This version of collection causes the default data writing process to occur and a raw image data movie file is saved. The user can also save a magnetic height reference file for Force-FRET experiments.

During real-time bead tracking experiments, the “Save Dat File” mode will pass the image data to the tracking program for real-time analysis, which then returns the analyzed data in an array (including bead x, y, and z positions among other instrumental parameters) and is subsequently written to file (Figure 2.3).

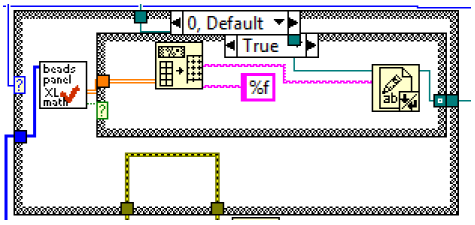


Figure 2.3. Save Dat File Mode. During acquisition in this mode, the andor software is still pulling the images from the camera, but instead of writing the raw data to file, it passes the image data to BeadPanelTracker-Math as a subVI routine for real-time processing. The image data is analyzed in this sub-routine and the results of the analysis are returned and written to file.

2.2) BeadPanelTracker-Math – The hub VI for real-time bead tracking

The complexity of this algorithmic architecture predates my work on this fragment of the program. I attempted to clean up as much of the code as I could while still being time efficient and maintaining the previous structure of the VI. This section will be dedicated to describing the functionalities of this VI and how it integrates with the necessary sub-routines.

To begin, the video data is sent into this routine as a one dimensional array of 16-bit intensity values. In most cases, the overarching conditional statement of "capture background" will read false, meaning the code that is displayed will run (Fig. 2.4). When the capture background button is pressed and held, the raw image data is collected, binned in the false portion of the conditional statement, and saved as a time averaged reference for background noise.

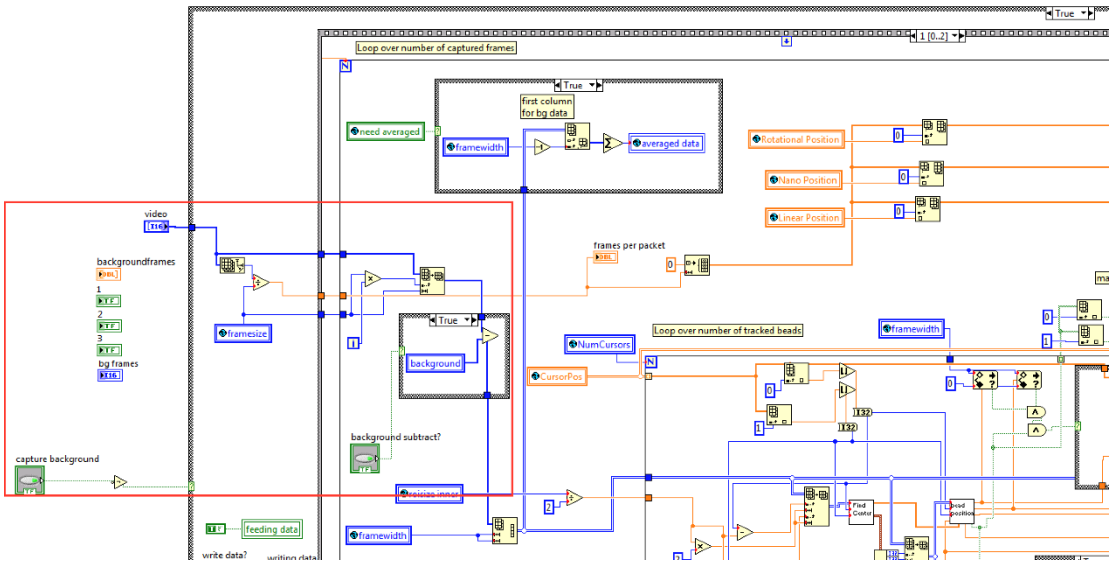


Figure 2.4. Grabbing individual frames inside a loop and background subtracting. In the highlighted region, the video data is parsed into individual frames

based on the frame size. The loop iterates over all the frames and on a per frame basis subtracts the background image. Future work could modify the code to do a global background correction rather than a per-frame background correction.

In the case of data analysis rather than background subtraction, input video data is checked for the number of frames in the image packet by taking the size of the image array and dividing it by the frame size. For example, if the frame size is 128 px x 128 px = 16,384 (default value) and there are 819,200 integer values in the video array then there are 819,200/16,384 = 50 full frames in the image packet. The for loop that this data is sent into will run 50 times, one iteration for each frame. In each iteration of the loop, a frame is grabbed from the "video" array by grabbing an array subset with the length framesize (grab a single frame) at the position framesize*i (current frame position in video array stack). The returned array subset is a single frame that is sent forward into the bead tracking algorithm.

Before the tracking begins the array is background subtracted (if selected) and reshaped into a two dimensional array with the width and height of "framewidth", this data represents the true camera image (Figure 2.5). Once a single frame image is read for analysis, it is fed into the loop that performs the tracking algorithm.

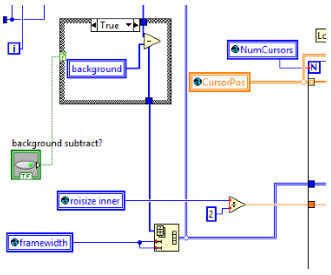


Figure 2.5. One dimensional data stream is reshaped into a real frame image. Input from the camera is a one dimensional array of integers, but the frame size is well defined so this data can be parsed and converted into a two dimensional array that is easily visualized.

Inside this nested for loop, the number of "Cursors" is looped over to perform x, y, and z tracking of the beads (Figure 2.6). The cursors are from the BeadPanelTracker VI (Section 1.4) and represent positional values of the bead centers, typically one reference bead and one tether bead, so in the default case this loop will have two iterations.

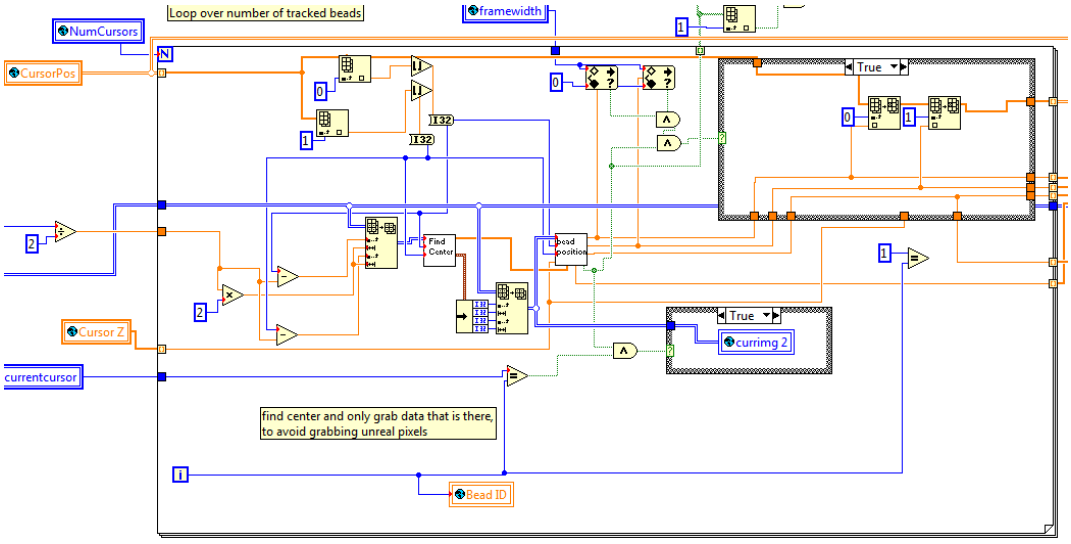


Figure 2.6. Bead position tracking. Inside this loop structure, each bead is iterated over. During each loop iteration, the previous iteration's bead position is used to grab an image subset that represents the likely bead center. The center is then tracked absolutely, and passed to a z-position determination program. Upon determination of the beads' x, y, and z parameters, error checking occurs to ensure proper calculation of bead position.

The first section of this loop is used to find the center of the bead for the current frame. Before discussing how the algorithm functions, a few concepts must be understood. Central to the x, y tracking is the idea that the previous frame provides the approximate location of the bead. In most cases the reference bead motion will be extremely minimal and the bead tether will only fluctuate a few pixels over the period

of one frame. In effect, this allows us to do local tracking of the bead center position through a computationally inexpensive method. Trying to find the bead center in the entire image would limit the runtime capabilities of the tracking program. The one caveat for this type of tracking is that the user must manually position each cursor (BeadPanelTracker VI) near each bead center at the start of an experiment, allowing the algorithm to begin tracking (Section 1.4). Programmatically, the loop reads in the previous position of the bead (x and y) and floors the values to integers (Figure 2.7).

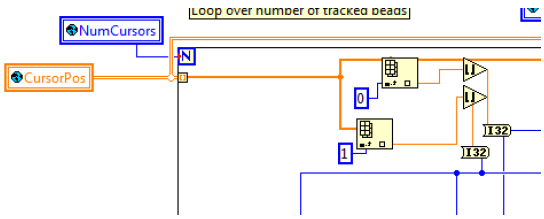


Figure 2.7 True bead x and y positions are floored for image capture. The x and y position of the bead in the previous frame are modified to an integer value so that a specific pixel value can be chosen for subsequent tracking analysis. This approximation should not change the outcome of the downstream calculations.

This process is necessary because the fractional values (or true bead centers) cannot be used when referencing the quantized array of intensity values in the video image. These values are used to set the "center" point of the bead within the current image.

Next, using the user defined ROI Find Center value from BeadPanelTracker (Section 1.4) VI, a subset of the image around the proposed bead center is retrieved (Figure 2.8). roisize inner is a global variable that represents the ROI Find Center value, in most cases this value will be something like 26. Inside the nested for loop, the array subset function takes image data from the reconstructed image array and grabs a two dimensional subset based on roisize inner. To grab the proper array

subset the following parameters must be used. The first parameter is the previous frame's bead x value minus ½ of roisize inner. In the x dimension it then grabs a number of columns equal to roisize inner. The next two parameters form the same operation, but in the y dimension. In effect, this array subset represents an image box with the width and height of roisize inner. The image is centered on the bead position from the previous frame. In this way, the subset image almost certainly contains the true bead center, with only a minimal amount of excess data. Using this image subset, the true bead center is found in the subset image using a version of a cross correlation function (see Section 2.3 for details).

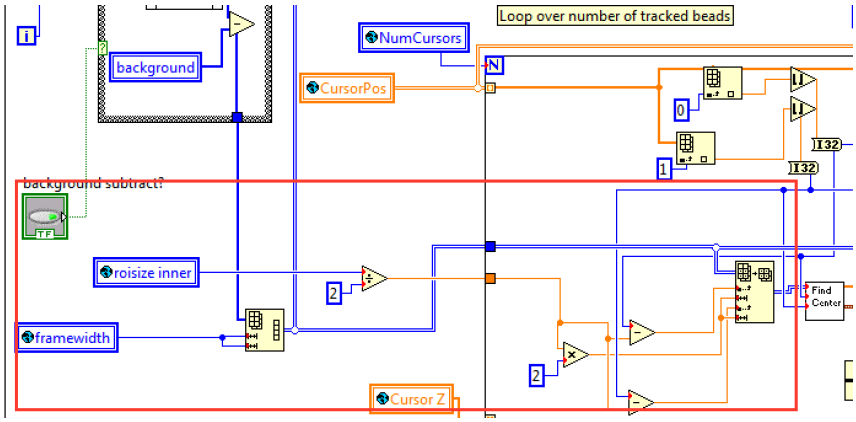


Figure 2.8. The image data is reshaped and parsed into a subset image. For centroid tracking it is advantageous to grab a local image subset of the approximate bead location and determine the true x and y position from this point. Such analysis minimizes computational expenses and prevents potential tracking issues associated with larger window sizes.

The Find Center subVI returns the change in x and y position relative to the previous cursor position (should be approximately less than ± 3, as the bead shouldn't have moved much over one frame). Similarly, this subVI returns a cluster of integer values which represent the proper image subset to be used for z position analysis.

This array subset is determined by finding the bead center for the current frame and grabbing a box around the center with the width and height of ROI Size from BeadPanelTracker (Section 1.4). This image subset is distinct from the one sent into Find Center because its center position now represents the true bead center for the current frame and the width/height is larger to encompass the entire diffraction pattern of the bead (Figure 2.9). The new image subset array is fed into the beadposition subVI where the z position of the bead is tracked using a lookup table algorithm (see Section 2.4). Other inputs include the bead's absolute x, y, and z positions from the previous frame and the calculated x and y offset values from the Find Center subVI.

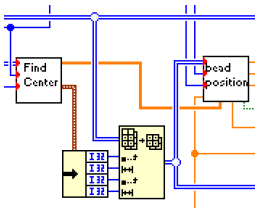


Figure 2.9. The calculated bead center recalculates the ROI for z position analysis. After determination of the true x and y positions, the image subset is redefined and optimized for z position analysis. During this process the image center (with respect to the true bead center) is reoriented and the image width/height is increased to encompass the entire interference patter of the bead.

Using these inputs, the absolute x, y, and z position are returned and written to file. The output values are doubles (decimal) that are ready to be written to file and passed on to the subsequent frame, but before that occurs the program performs error checking to make sure the values are acceptable (Figure 2.10). In the upper left region of this figure, the returned x and y positions are checked to be within the image frame. In the case of normal image capture mode, this would check if the x and y

values are between 0 and 128. In theory this is a suitable check, but if the bead center is near 0 or 128 in either dimension, the tracking will fail as part of the bead will be outside the image. If the check fails (and therefore the tracking fails) NaN values (not a number) are passed forward rather than the aberrant tracking values. This allows downstream data processing to recognize tracking errors.

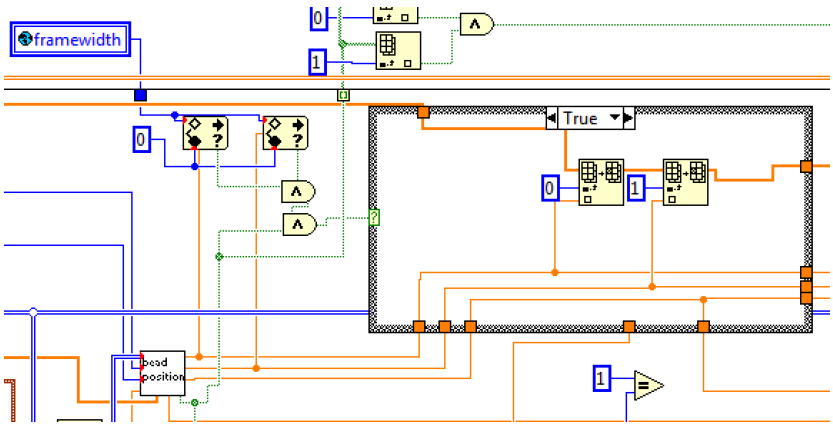


Figure 2.10. Error checking ensures the proposed bead center is within real limits. The program checks the proposed bead x and y positions and determines if the values are within the camera frame. If not, then the program assumes a tracking error has occurred and NaN values are written to file.

The output of the nested for loops will be a two dimensional array of data (Figure 2.11). The first three parameters are the rotational position of the magnets, the z-position of the piezo-stepper motor mounted objective, and the linear (z) position of the magnets. These parameters are defined and controlled by the MotionControl VI as global variables (Section 1.6). The next four parameters are the beads' x, y, z parameters, as well as the current force. When exiting the for loop that performs the bead tracking, the data is exported on a per frame basis, creating a one dimensional array for each bead, where index 0 contains the values for bead 0 (reference) and

index 1 contains the values for bead 1 (tether). The current force parameter is purely an estimate of the force based on magnet linear position (assuming MyOne beads for tethering) unless the force calibration parameters are manually adjusted in the beadposition subVI. All of these components are brought together and interleaved into a one dimensional array of values where the parameters read in this order (Rotational position, Nano position, linear position, Bead0 x, Bead0 y, Bead0 z, Force, Rotational position, Nano position, linear position, Bead1 x, Bead1 y, Bead1 z, Force). I am aware of the redundancy in the information written to file, but this architecture was chosen before my involvement and was therefore maintained. This data is exported outside of the frame for loop, creating a two dimensional array of data that contains bead information in the rows, where each row represents an individual frame. Finally, the algorithm completes by updating the necessary global variables.

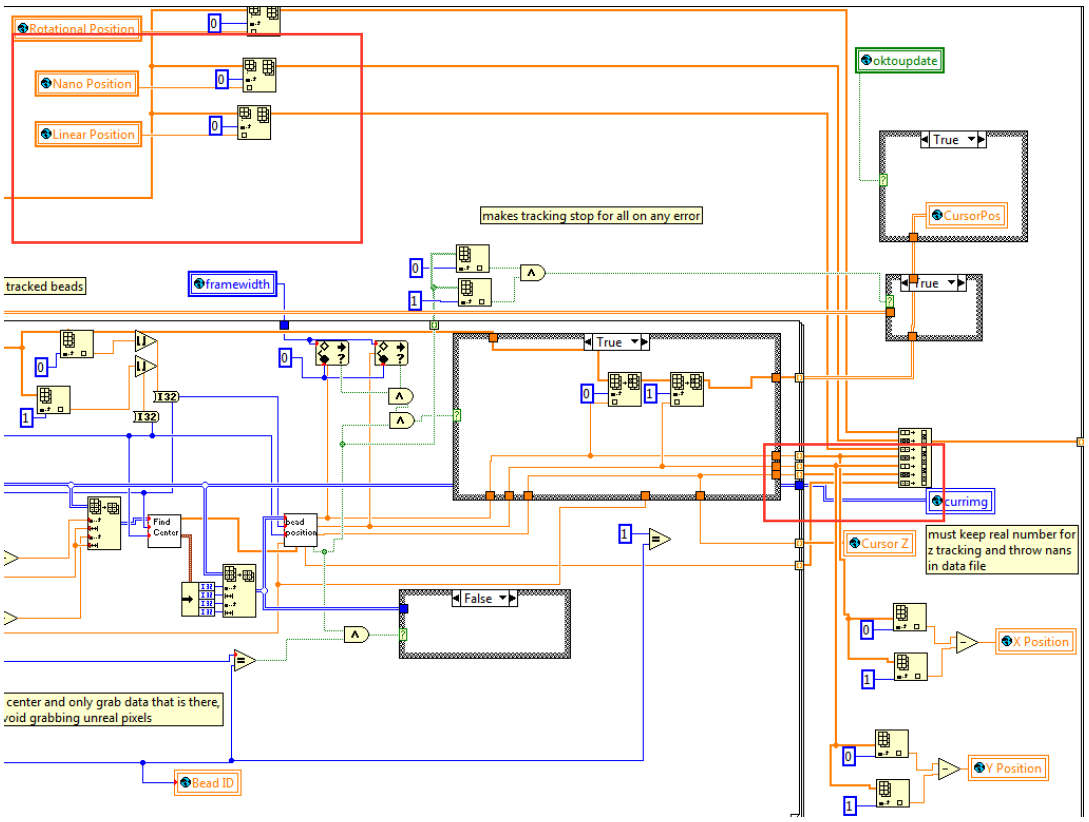


Figure 2.11. Data is exported for writing to file. The output of BeadPanelTracker-Math is a two dimensional array of values that contains instrumental parameters as well as x, y, and z positions of the tether and reference beads. Each row of the two dimensional array represents these data points for an individual frame.

2.3) find_bead_center – Determining the absolute bead center through a modified cross-correlation analysis

The purpose of the find bead center algorithm is to locate the absolute x and y position of the bead center in the current frame, to sub-pixel accuracy. In order to achieve this, the x and y centers are determined individually. As the interference pattern produces a symmetrical pattern, the bead center can be found by performing a cross-correlation analysis on the inverse image. To simplify the calculations, a two dimensional image can be collapsed into two separate one dimensional arrays by vertical binning (for x intensities) and horizontal binning (for y intensities). The

output of such binning will produce a two dimensional pattern that represents a three dimensional object (Figure 2.12, see bead image vs Raw binning, white x binning and red y binning).

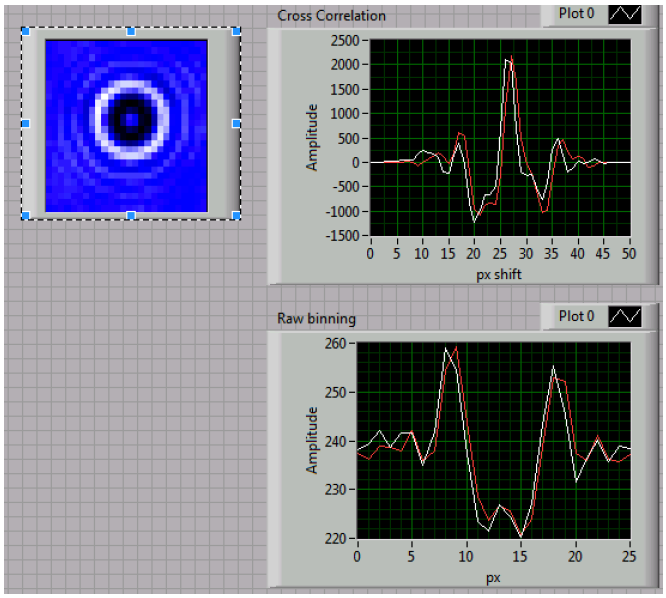


Figure 2.12 Finding the bead center. In this figure, the raw bead image is displayed in the upper left corner. Vertical and horizontal binning of the image data produces a two dimensional representation of the bead, in the x (white) and y (red) dimensions respectively. Cross correlation analysis on these two dimensional representations allows tracking of the center positions with minimal computational overhead.

The image transformation into the binned arrays is performed quite simply as a sum of the columns and rows (Figure 2.13). Let's consider the example of determining the one dimensional y data array. The first, outer most for loop runs twice, once for x-dimension (i_{outer}=0) and once for the y-dimension (i_{outer}=1). The second level loop has an iteration count of n=roisize inner. As the ROI data input has a dimension length equivalent to roisize inner, the second level loop will run for each column (i_{outer} = 0) and each row (i_{outer} = 1). On the second iteration of the outer loop, i_{outer}=1, the condition statement in the second level (inner) for loop will be false and

the code that is visualized will be run. Each iteration of the inner loop then extracts an entire row at position i_{inner} of the image data. The row is summed and averaged to provide a single point in the bin array. In this way, all inner loop iterations will create a one dimensional array, where each element is the sum of a row or column in the image (Figure 2.12, see raw binning). The data exported from the outer loop will be a two dimensional array, where index 0 contains a one dimensional array of summed columns, and index 1 contains a one dimensional array of summed rows.

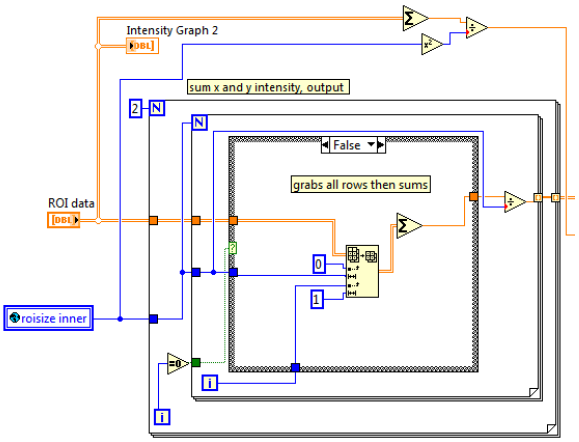


Figure 2.13. Summing the two dimensional image array into a pair of one dimensional x and y intensity arrays. The x and y binned values are accomplished through a nested loop structure. The outer loop provides two iterations where the x then y binned arrays are computed. The inner loop is run over the entire number of columns or rows, depending on the outer loops iteration number. Each column or row is grabbed, summed, and exported to an array of summed columns or rows.

Once the row sum and column sum are generated, the image is normalized to reduce noise during cross correlation. In order to achieve this, all the pixel intensities in the image are summed and divided by the total number of pixels, thereby reducing the average signal per pixel to zero (Figure 2.14). Without this normalization process,

the cross correlation returns a maximum when the two arrays in the cross correlation are completely overlapped, as this results in the highest amount of signal overlap. When the values are normalized, the cross correlation only returns positive correlation when the true signal regions overlap (Figure 2.12, see cross correlation).

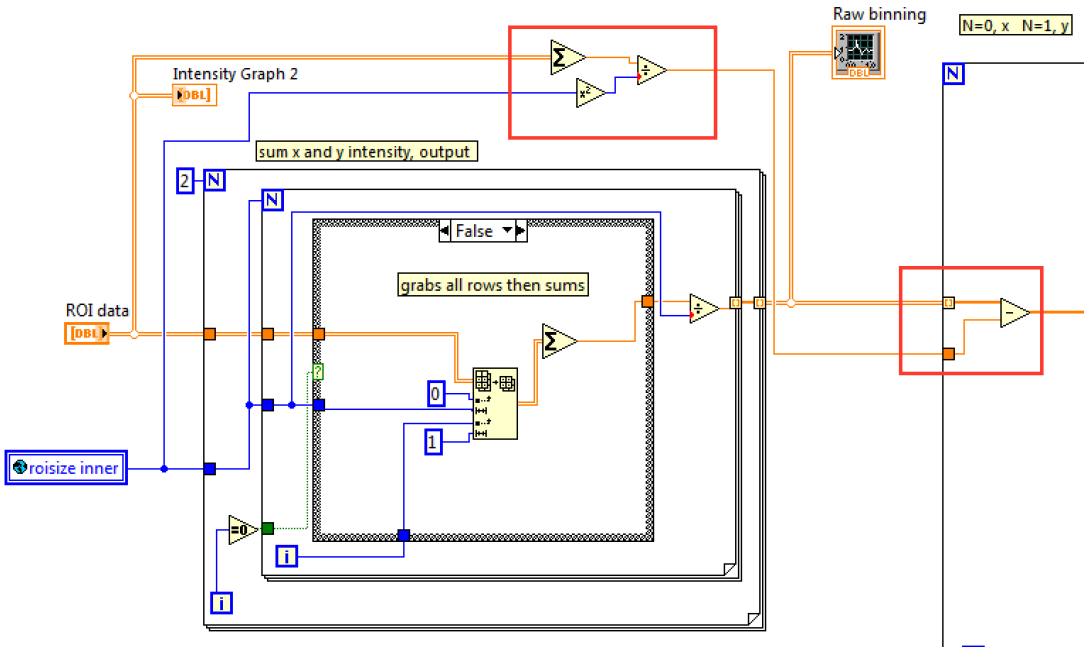


Figure 2.14. Normalizing the image data. The average pixel intensity of the entire ROI is subtracted from each binned row. This places the average intensity value at zero and allows the cross correlation to find the real bead center.

The cross correlation is simply performed by comparing one of the sum arrays (per loop iteration), to the inverted version of that same array (Figure 2.15, left red box). In short, the cross correlation takes the inverted array and compares the amount of signal overlap on a per index basis. For example, let's consider the sum of the columns and therefore the X array and the X_{inv} (X inverted). The first comparison made by a cross correlation will be determining the amount of signal overlap between X[0] and X_{inv}[N] where 0 is the first index of X and N is the last index of X_{inv}. The

second iteration will compare X[0,1] to X_{inv}[N-1,N]. This process continues until the comparison X[N] and X_{inv}[0] are compared. In essence, this function slides the X_{inv} array over the X array from one index of overlap, to total array overlap, back to one index of overlap on the opposite edge of X. If the center position of the bead is in the exact center of the image, then the cross correlation will return the highest product value at 1/2N shift, or where the two arrays are completely overlapped. I encourage the reader to look at the extensive online resources for cross correlation if a more detailed description is required.

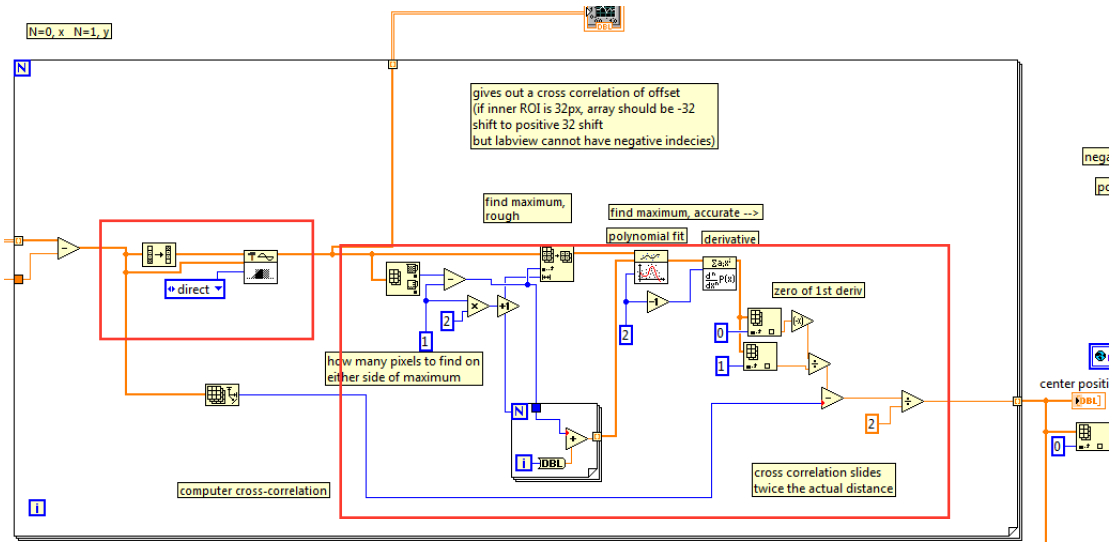


Figure 2.15. Performing the cross correlation and finding the bead center. The cross correlation is performed through a built-in VI (left red box). The cross correlation result is used in a second order polynomial fit and a subsequent derivative to find the peak center. The cross correlation peak center is then converted to the real bead center, with respect to the ROI image.

The cross-correlation returns an array with a length twice the size of ROI Find Center and is subsequently tested for the bead center. In order to extrapolate the bead center from the cross-correlation analysis, the maximum of the cross correlation is

found (Figure 2.12 Cross Correlation, example shows maximum around x = 26). This maximum provides a rough center position, but only contains quantized pixel accuracy. Once the maximum index is found, a subset array of the two neighboring values (one on each side) of the maximum is acquired. A fitting function can be applied to these data to find a more accurate, sub-pixel resolution bead center. These three points are fit to a second order polynomial and the fitting parameters are returned. I have tested multiple values in place of the "1" value in the right red box and, surprisingly, grabbing less values for the polynomial fit provides more accurate centroid tracking (Figure 2.15). I encourage the user to modify this number and do extensive analysis on this value as a parameter for bead tracking, as my testing was performed on-the-fly. Once the rough maximum is fit with a second order polynomial, the first order derivative of the fit parameters is obtained. This process is used to find the absolute center, as the first derivative will provide an exact minimum of the fit trough. The returned derivative parameters are representative of the equation $y = mx + b$, where parameter 0 is b and parameter 1 is m. Therefore, solving the equation under the conditions $y = 0$ will provide the absolute maximum of the first order integral. This absolute center (relative to the input image) must be obtained for downstream processing, as the output of the centroid process to this point represents the x and y position in the cross correlation rather than the initial array. Take the example in Figure 2.12, the center position of the cross correlation is obviously at ~26. If you look at the raw data in this figure, you see that the actual bead center is at ~ 13. This is case of perfect tracking where no motion occurred during the previous

frame, so it is simple to see how subtracting the initial array length (26) from the cross correlation peak center (~26), will give you a center position of 0, meaning the center position did not change during this frame. A subtle but important point is to understand why this output value is subsequently divided by two. To understand this concept let us consider an extreme case where the bead moved by 5 pixels in the x direction (refer to Figure 2.12 during example). In the raw data, the bead center will be 5 pixels shifted from the center, maybe around 18. When the cross correlation is performed, the X_{inv} array must shift the full 25 pixels to get to make the arrays overlap entirely (1/2N). Then to reach a fit maximum, the array must shift 5 pixels for the X_{inv} bead center to reach the image center and another 5 pixels for the X_{inv} bead center to overlap perfectly with the X bead center. This would place the center of the cross correlation fit at x = 35. When we subtract the array length from the cross correlation center we receive a value of 10. This value is twice the actual offset and is a property of the cross correlation function, so dividing the offset value by two will provide the actual distance change with respect to the true bead image rather than the cross correlation analysis.

The next portion of the find center algorithm defines the image bounding box of the bead for subsequent z-position analysis (Figure 2.16). Originally the algorithm was written without the region boxed in red. For the purposes of most tracking instances this code would work fine. The code in the red box takes the previous frame x and y position, adjusts the bead center based on the cross correlation analysis, then defines the bounding box for the z position analysis with a width and height of

roisize. This ROI is centered exactly at the x and y center. The ROI x and y positions (with respect to the entire camera image) are defined as the ceiled versions of the calculated upper left hand corner of the ROI. The width and height are the defined by the roisize. As I mentioned before, this new ROI bounding box for z position tracking will work in most occasions, but when the bead is close to the camera edge the tracking will fail. For example, consider if one bead is close to the edge of the camera and 1/3 of the interference pattern is not being captured. The user defined ROI will try to grab the entire ring pattern, in this case 1/3 of that data is non-existent. Attempting to grab non-existent data will throw an error in LabVIEW. You may wonder why the bead would be close to the camera bounds. In some experiments, reference beads and tethers are hard to find in the same image, so some bead pairs are just barely close enough together for imaging in the same camera frame. The code in the red box facilitates these experiments. The delineated conditional statements check if the proposed ROI (based on the bead center) is within the camera frame. In our example where 1/3 of the bead is off the screen, the code will define the bead bounding box as the real 2/3 image data rather than including the non-existent 1/3 of the data. After the checks are complete the proper bounding box is saved in an array cluster and exported for the z position analysis.

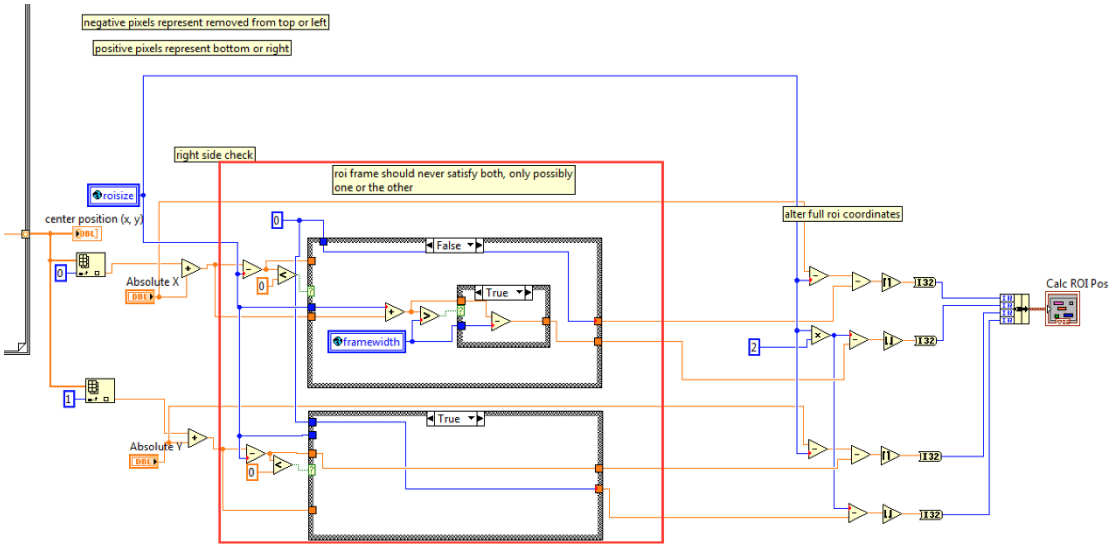


Figure 2.16. Building the ROI bounding box for z position analysis. Using the newly defined bead center, the ROI determines the position of the bead with respect to the camera. If the bead is close to the edge of the camera viewing field, then the ROI is corrected for the lack of data.

2.4) beadposition – Computing the tether bead z-position using a look-up table analysis

After the beads x and y positions are found through find_bead_center, the z-position of each bead is calculated through the beadposition VI by analyzing the diffraction pattern of the beads (Figure 2.17). The input image for this VI is a center positioned image subset with width and height defined by the user in the BeadPanelTracker.vi (ROI Size). The x and y positions from the previous frame along with the calculated lateral motion of the bead (during the current frame) are also fed into the vi for center position reference during z position calculation. Once the z-position is calculated (see below for details), beadposition outputs the current frame/current bead absolute x, y, and z position, along with the estimated current

force and an error flag that will disable tracking if an error occurs during z-position analysis.

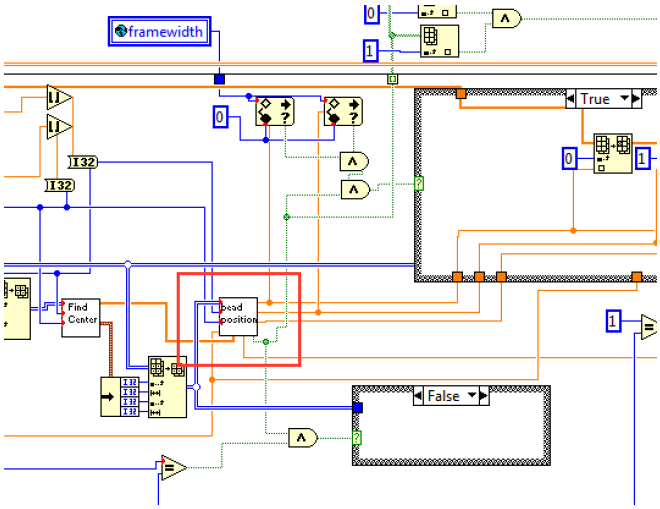


Figure 2.17. Beadposition is used to grab a bead ROI and find the z-position. The Find Center algorithm returns the exact bead position in the full image. These values are used to compute a new ROI for determining the RIP (and z-position) of the current bead under study.

Beadposition determines the z-position of a bead by generated a radial intensity profile (RIP) of the bead diffraction pattern and comparing that RIP to a set of calibration values. Therefore, in order to understand the z-position tracking it is necessary to first understand how a RIP is calculated, what it represents, and how RIPs are converted to a calibration set, also known as a lookup table. A RIP is a one dimensional representation of the bead diffraction pattern, much like the raw binning profile in Figure 2.12. Instead of binning in the x and y dimensions, the pixel intensities are binned radially from the center of the bead. In Figure 2.18 is an image of a typical bead, with a determined center position (red). If you imagine taking a cross section of the bead image from the bead center along the blue line until point R,

you would achieve a radial intensity profile seen in the right panel. Of course, this profile would be quite noisy due to scarcity of data along the individual line. If the blue line was rotated anywhere between 0 and 360 degrees, approximate the same RIP would be achieved. To maximize resolution of the RIP, all the possible cross sections can be obtained and combined to form a single, high-resolution RIP. How this is achieved in a digital image is discussed later (Section 2.5, CalculateRadialIntensityProfile).

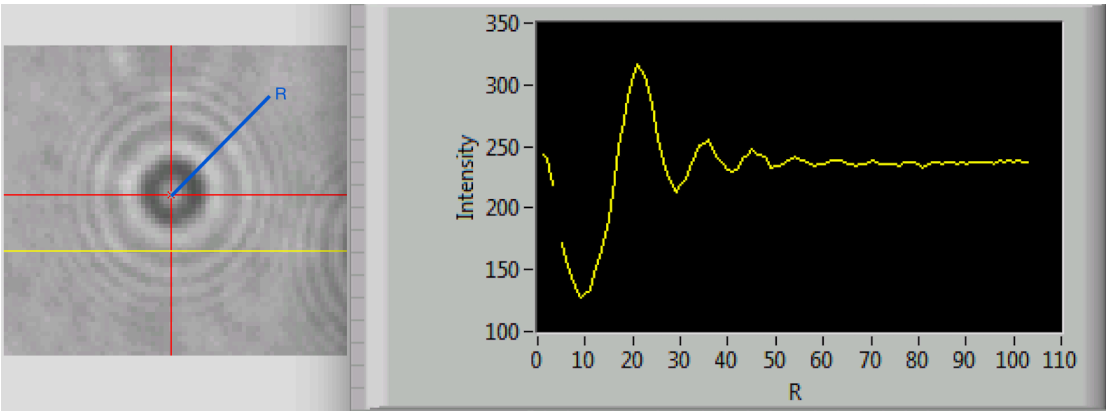


Figure 2.18. Radial Intensity Profile (RIP) is calculated by binning all the ROI data into a one dimensional array. Using the bead center and the newly generated ROI, the bead image (left) is used to create a corresponding one dimensional array (right) for real-time analysis. This methodology is much more computationally efficient than total image comparison.

The exact diffraction pattern that is created by the bead is a result of defocusing away from the bead focal plane. Focusing out of the bead plane will provide an interference pattern that changes as a function of distance from the focal plane. This interference pattern is extremely sensitive to distance changes and can accurately report on bead motion in the z-plane. For a bead fixed on the slide surface, the motion will be negligible on the order of minutes. Therefore, using an extremely

high resolution piezo-stepper motor, one can modulate the interference pattern by quantized values (Figure 2.19).

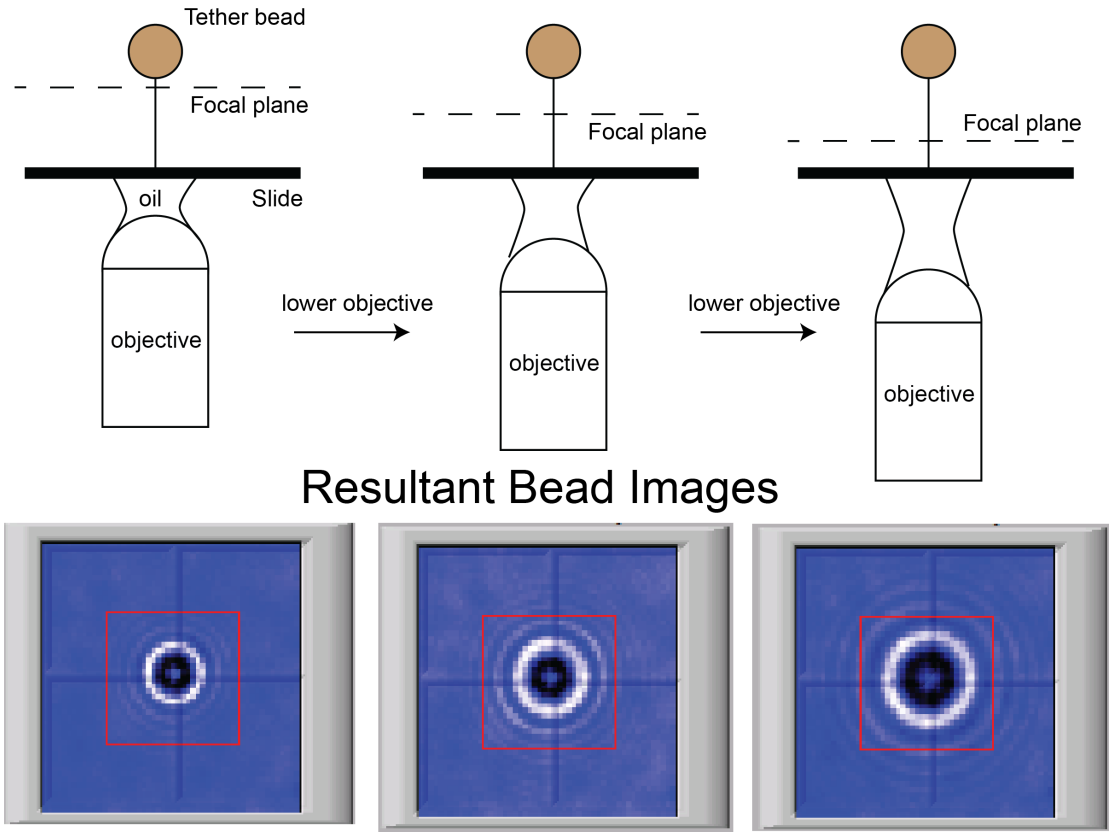


Figure 2.19. Moving the objective changes the focal plane and therefore the diffraction pattern of the bead. Using a stepper motor on the objective, the focal plane of the captured image can be changed. As the focal plane changes, different bead diffraction patterns can be captured and correlated to specific bead heights, which can later be used for bead tracking.

In the beadposition VI, a stacked sequence governs the program execution. The first frame of the stacked sequence (0 of 4) is purely for analytical purposes and creates a timer to determine how the timespan for z-position analysis. Frame 1 initializes the radial intensity array and radial intensity array hits for use during calculate RIP (Figure 2.20). The former array stores the RIPs during the given frame

and the hits array is used to tell what position in the RIP each pixel corresponds to during RIP calculation (see Section 2.5, CalculateRadialIntensityProfile). Other functions in this frame include resetting the bead zero point after calibration and plotting the approximate force decay.

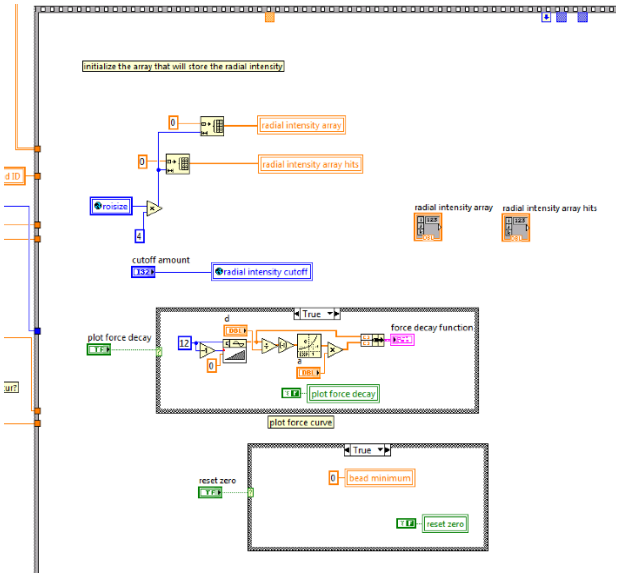


Figure 2.20. Frame one of beadposition initializes variables. Each time frame one of beadposition is run, it reinitializes the radial intensity arrays so that a new round of z-position tracking can be performed. This frame also functions to reset calibrated zero and plot force decay.

The second frame of the sequence implements the CalculateRadialIntensityProfile VI, which computes the current frame/current bead RIP for subsequent use in z-position determination (Figure 2.21). The inputs to this function include: a quadrant selector that allows the user to define which quadrants of the ROI to use in BeadPanelTracker, the image ROI of the bead defined by ROI size, the center position of the bead within the image ROI, and the empty radial intensity

arrays. The output of this function is a radial intensity profile, as seen in the right panel of Figure 2.18. During this frame the bead display pane is also updated.

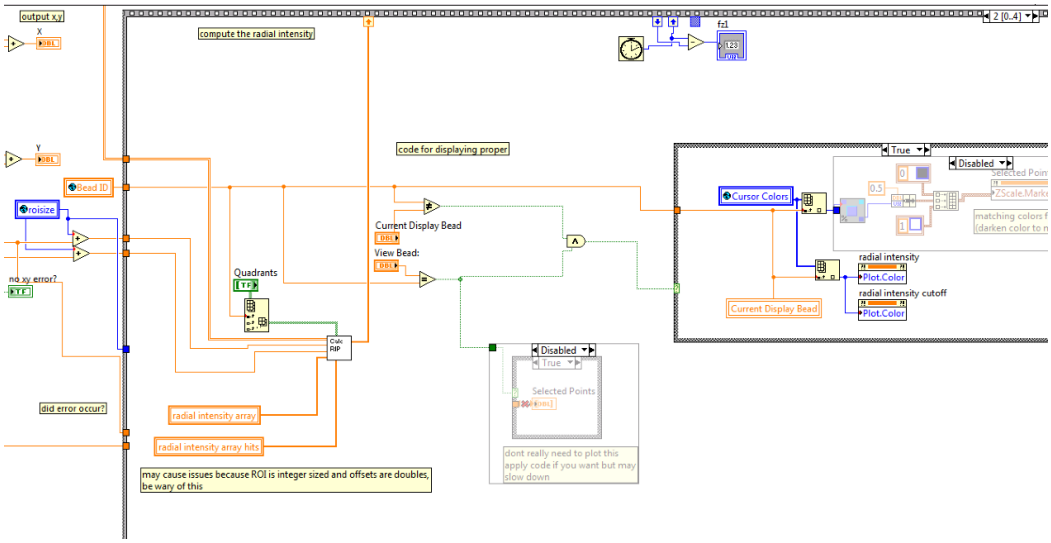


Figure 2.21. Calculating the radial intensity profile of the image ROI occurs in frame two. In order to track the z-position efficiently, the highly symmetric bead image is converted to a one dimensional array in the Calc RIP sub VI. The inputs to this function provide empty arrays for data storage, the bead center, and the image ROI of the bead.

The third frame takes two different forms, depending on whether real-time tracking is occurring or the lookup table is being calibrated (Figure 2.22). Understanding the z-position tracking algorithm requires an understanding of the lookup table, which is a result of calibration. Therefore, we will consider the

calibrating case first and how the code operates during the calibration process.

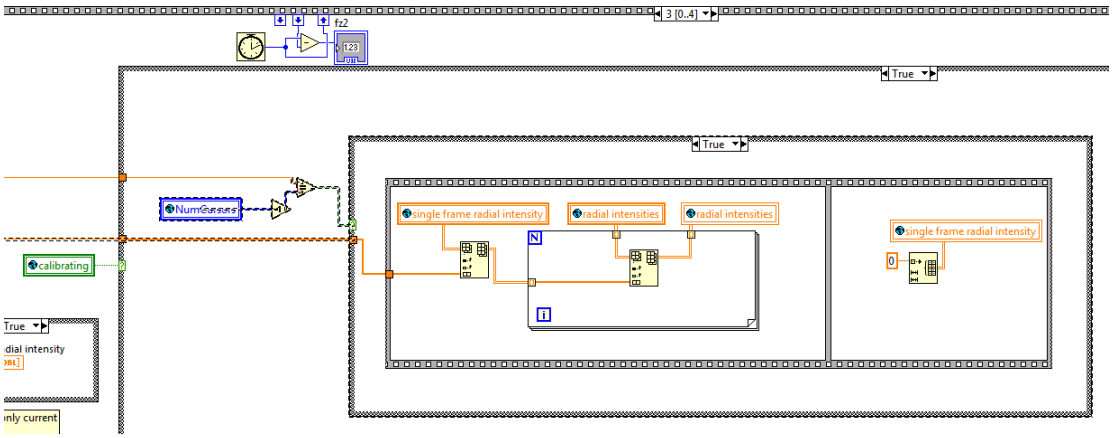


Figure 2.22. Frame three of the bead position mode operating in "calibrating" mode stores the radial intensity profiles in a global variable. During the calibration process, instead of tracking the bead position this VI uses frame three to save the RIP generated in frame two in a global variable (radial intensities). This variable stores the RIPs during the calibration process (discussed in detail below).

By stepping the objective through multiple positions, one can capture the change in the bead diffraction pattern over discrete distances thereby generating a lookup table of distances. This lookup table acts as a reference for bead motion during real-time analysis. The step size and duration of time spent will determine the resolution of bead tracking. In theory, very small step sizes and very long durations at each step would provide the best accuracy for a calibration. In reality, limitations of the piezo-stepper motor and drift in the bead position through thermal fluctuations make an ideal step size of around 100 nm with a 200 ms acquisition time for a 5 micrometer calibration range (determined experimentally). How the calibration process is achieved programmatically will be discussed below.

Calibrating begins when the user enters calibration parameters and clicks the Start Calibration button in the Motion Control VI (Figure 2.23).

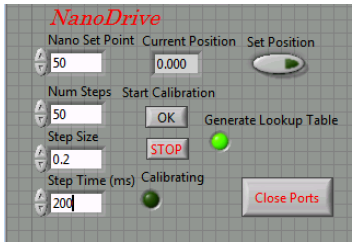


Figure 2.23. NanoDrive controller allows calibration of the bead images for real-time z-position tracking of beads. The nano drive parameters (Num Steps, Step Size, and Step Time (ms)) control the calibration process for RIPs. The Num Steps defines how many discrete states are to be acquired during calibration. The Step Size defines how many microns the stepper motor should move the objective during each individual step. Lastly, the Step Time defines how long the capturing process occurs at each step. This parameter essentially defines how many frames are being used to define a RIP at a given objective height. The Start Calibration button will begin the z-position calibration using the chosen parameters.

When this button is clicked, a loop in the Motion Control VI is activated. Frame 0 in this loop is a waiting period for the allotted step time (Figure 2.24). This step is in place so that the beadposition VI has time to acquire and store the frames (and subsequent RIPs) for this calibration position. For example, consider if the Step Time was set to 200 ms and the camera framerate was set to 500 Hz (2 ms exposure). beadposition VI would run 100 times, saving the captured RIPs to a global variable "radial intensities" during frame three of the beadposition stacked sequence (Figure 2.22). During this period, the instrument (and most importantly the objective) is stationary so all of the collected frames represent the same objective position and therefore the same bead z-position.

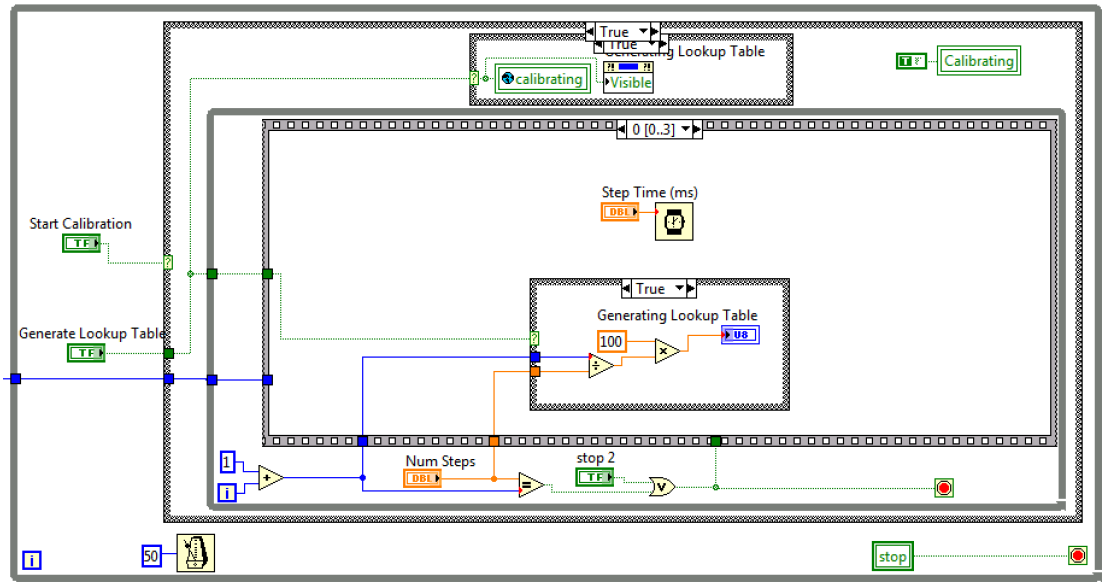


Figure 2.24. Frame zero of the Motion Control calibration loop makes the calibration process wait while RIPs are saved. During the allotted wait time seen above, the beadposition VI is collecting RIPs at the current objective position and storing them to a global array. Longer wait periods will result in more data being used for the final lookup table values. If errors are being thrown during the calibration process, a simple fix is to try and increase the wait time slightly.

Once the allotted time span is completed, the next stacked frame (Frame 1) in Motion Control averages the bundle of RIPs, fixes any erroneous values in the averaged RIP, and adds the calculated RIP to the lookup table global array (Figure 2.25). At the same time, the objective position is saved in a Lookup Table Heights array as a reference of the absolute position of the focal plane.

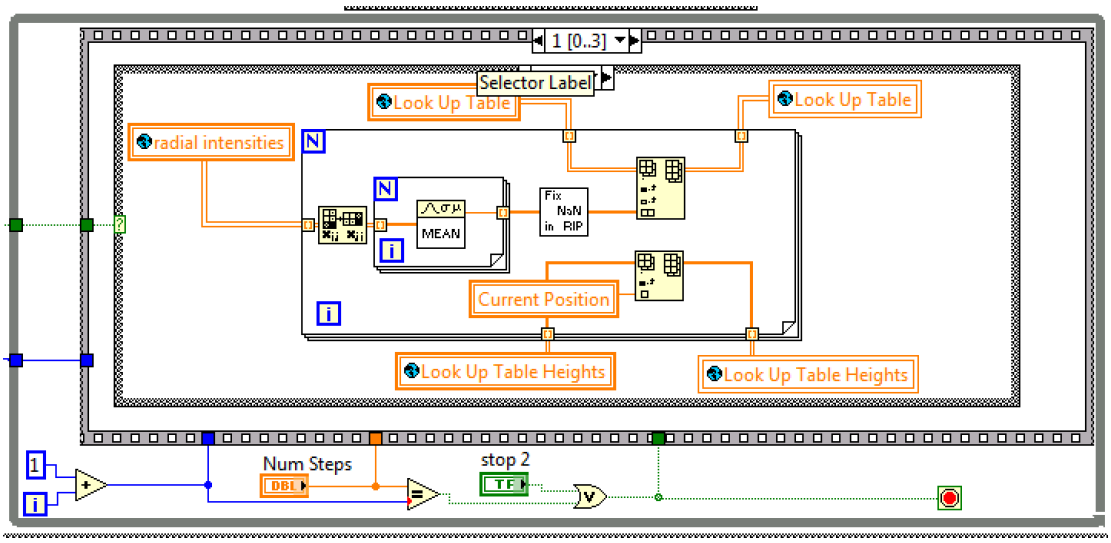


Figure 2.25. Frame 1 of the Motion Control calibration loop stores the captured RIP in the lookup table as a single calibration point. All of the RIPs calculated in the beadposition VI are averaged to create a final RIP for the given objective height and then stored in the Lookup Table array. The Lookup Table Heights array is used to define the current objective position (focal plane) for the corresponding RIP.

Once the calibration RIP for the current objective position is saved, the next lookup table RIP can be captured once Frame 2 of the Motion Control loop moves the objective by Step Size (Figure 2.26). The piezo motor requires a connection to the initialized link, so that link is wired in from the initiation center. The motor is then stepped up incrementally during each iteration of the loop. By example, if the initial Nano Set Point was 50.0 μm and the set step size was 0.2 μm, then the objective would be moved to 50.2 μm during this frame.

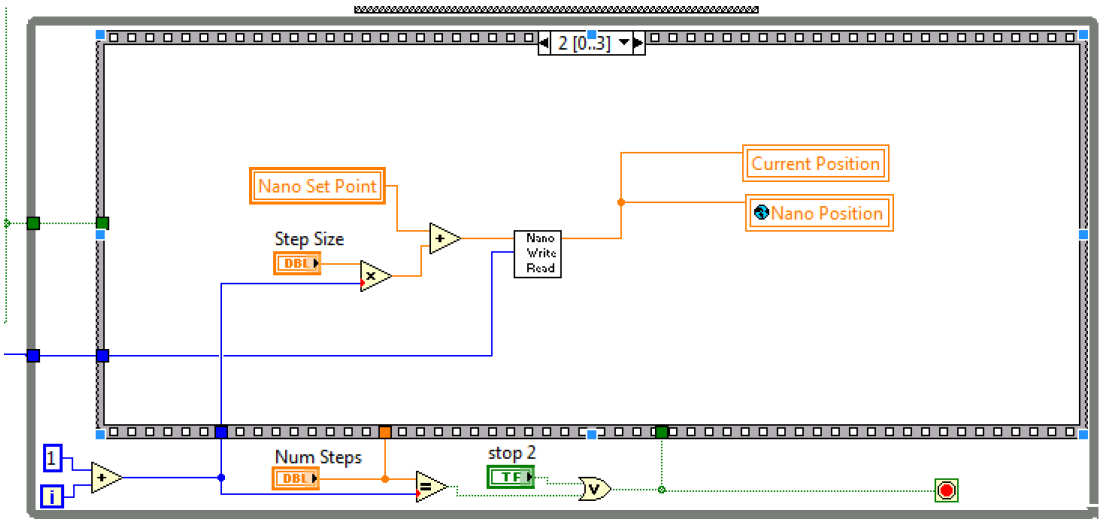


Figure 2.26. Frame 2 of the Motion Control calibration loop moves the objective to the next calibration position. Once the previous calibration RIP is saved in the Lookup table, the objective is moved to the next calibration point.

During the last frame of the sequence (Frame 3), the temporary array for RIPs is cleared for the next objective position and the global LUT is converted to a local LUT in the beadposition VI (Figure 2.27). The global LUT is only converted to a local LUT once all of the steps of the objective position motion have been completed. Note that this conversion from global to local allows faster program execution during real-time analysis, but it requires that the current vi/folder architecture to be maintained. If the beadposition or MotionControl VI must be moved or renamed, the Setup Local LUT VI must be modified. Once this operation completes, the while loop containing the calibration sequence is terminated. Once the inner loop is complete, the Start Calibration conditional is checked and the check will return false (as the Mechanical action is switch until released), and the calibration will complete, signaling both the Motion Control VI and beadposition VI that the calibration is complete (Figure 2.24). I should note that most of the calibration code could be consolidated to the

beadposition VI, allowing usage of only local variables (for radial intensities) which may optimize calibration performance. Proper care would need to be taken in order to ensure that flagging the calibration process would correctly identify each step initiated by the MotionControl VI. Currently the calibration code is integrated into the MotionControl VI because only one instrument initialization is allowed at a time for the piezo stepper.

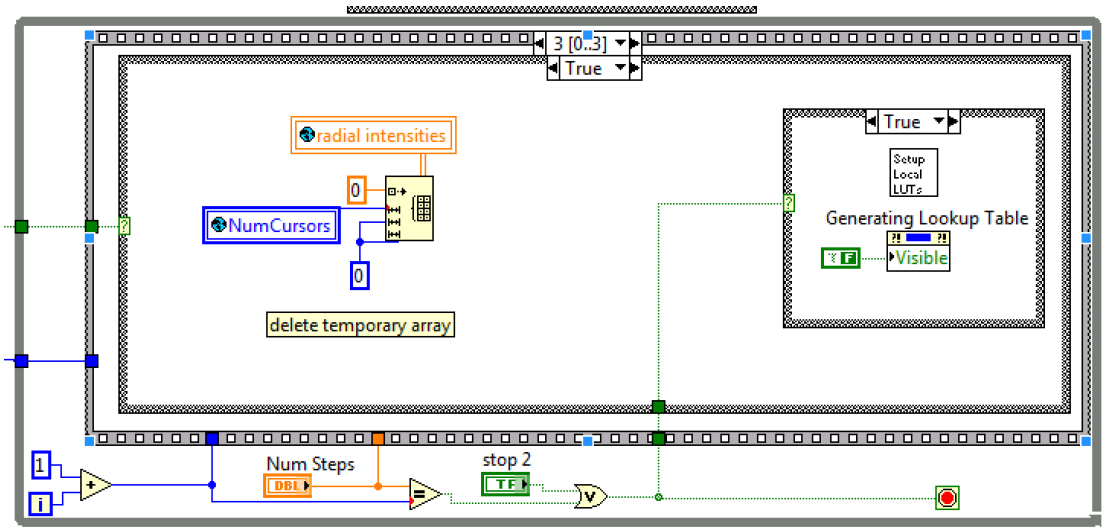


Figure 2.27. Frame three of the Motion Control calibration loop reinitializes the RIP global and builds a local Lookup Table for the beadposition array if the calibration is completed. The global variable radial intensities saves the RIPs generated by beadposition during the calibration. Each calibration point should have a distinct set of RIPs so the array is cleared for the next loop iteration. If the last calibration point has been obtained, the Setup Local LUTs VI will move the global Lookup table values to local variables in the beadposition VI for efficient tracking of bead z-positions.

Throughout the calibration procedure, beadposition Frame 3 was functioning purely for calibration procedure, but once the calibration is completed, the "calibrating" conditional will return false and real-time tracking begins (Figure 2.28). Tracking initiates by retrieving an array subset of the RIP for the current frame/bead.

The cutoff value (typically around 8) removes the lower indices of the RIP array as there are some NaN values (nonexistent data) in this area. The NaN values are a result of the low resolution of our camera. With a higher resolution camera, the NaN values would be confined to a much smaller region toward the zero index of the RIP array. Despite this shortcoming, the remaining RIP profile is sufficient for z-position tracking. The modified RIP is normalized for comparison to each RIP in the lookup table which were generated during the calibration procedure (see above). The comparison takes the form of a Pearson chi-squared test, which in this case measures the similarity of the current RIP to each RIP in the lookup table. In order to determine the similarity, the current RIP is subtracted from the expected (lookup table) RIP and then squared to remove any sign from each comparison index. The sum of these comparison values will then give you a value greater than or equal to zero, where zero is a perfect match between the RIPs.

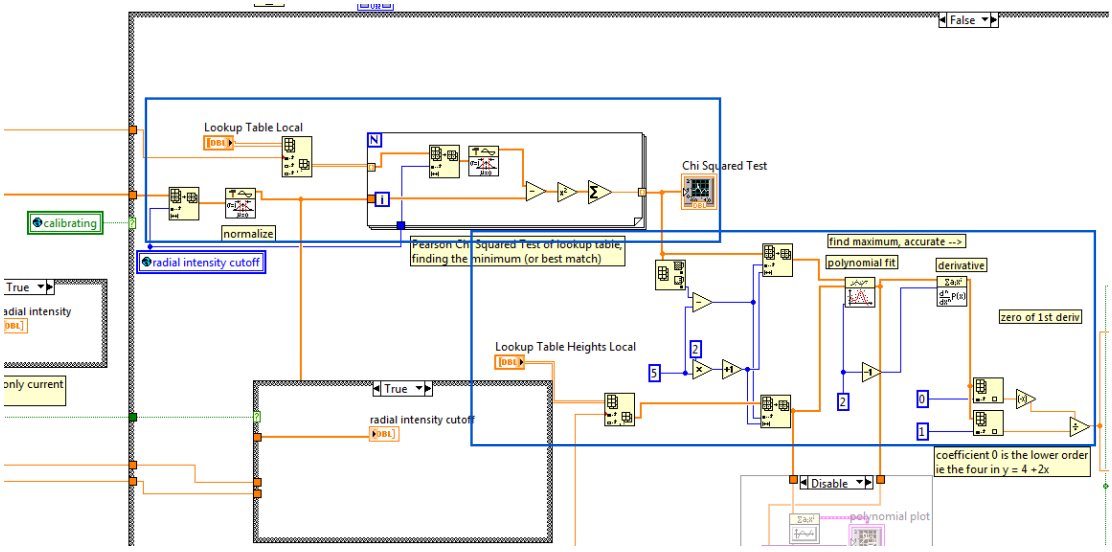


Figure 2.28. beadposition frame three will use the calibrated lookup tables to determine the current bead z-position. A Pearson chi-squared test is used to

determine the Lookup Table RIP that is most similar to the current bead RIP (rough z-position). The local minima of the chi-squared test is fit, with neighboring values, to determine a parabolic function for the chi-squared minima. This process is used in order to define the exact z-position of the bead.

A one dimensional array of comparison values is output from the chi-squared test, where lower values represent similarity between the current RIP and the lookup table (Figure 2.29). As one might expect, the minimum of the chi-squared tests can be obtained, but it will most likely provide an inexact match between the current RIP and a single RIP in the lookup table. To account for the quantized nature of the lookup table, the lowest 10 values from the chi-squared test can be fit to a second order polynomial to determine the exact bead z-position. This form of interpolation is identical to the find_bead_center VI, which then takes the derivative of the function to calculate the local minimum of the chi-squared test.

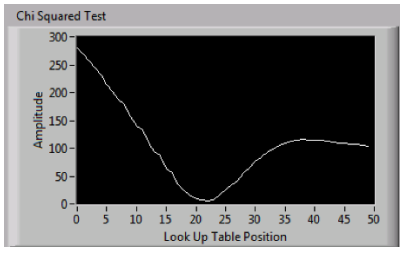


Figure 2.29. A visual representation of the Pearson chi-squared test for determining z-position. The minima of this function defines the position of the current RIP in the Lookup table. Using the calibration values, the exact z-position of the current bead can be determined.

After the absolute position has been determined, the algorithm manipulates the absolute positions of both beads (reference and tether) to provide a relative tether length for real-time analysis (Figure 2.30). If the calibration took place with a torsionally unconstrained molecule, then reference zero position of the molecule

would read at full extension if the value was not corrected. In this case, a correction for the true z-position of the tether can be applied by allowing the bead to bounce on the slide surface while the Define Zero button is pressed. By doing this, the true zero position of the tether bead is defined (where the bead is touching the slide surface) and then corrected for during subsequent real-time analysis frames. In last conditional of Figure2.30, the code corrects for relative z position, plots the data from the real-time analysis for user visualization, and ensures no errors are flagged for the relative bead heights.

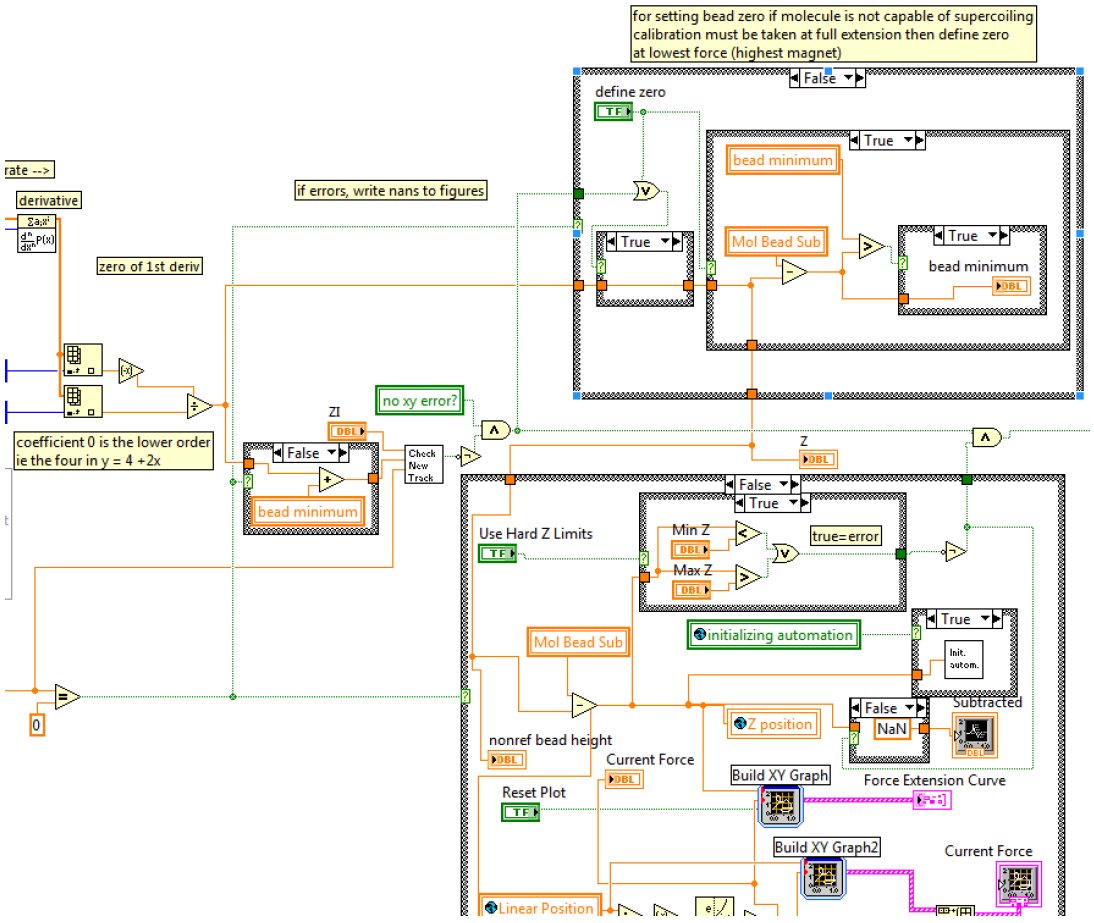


Figure 2.30. Final portions of the beadposition tracking algorithm incorporate z-positional correction and build plots for the user interface. If the calibration was

performed with an extended molecule, then the zero position of the tether bead must be determined. This is done by allowing the bead to bounce on the surface and defining a "bead minimum". The bead minimum is then used to correct the relative difference in z-position between the reference bead and the tether during real-time analysis. Graphs are then plotted providing the absolute and relative positions of the beads, as well as the force extension curve and the current force curve.

2.5) ComputeRadialIntensityProfile – Computing a radial representation of the bead interference pattern in a digital image.

Radial Intensity Profiles (RIPs) are one dimensional array representations of the bead interferences pattern obtained from the camera (Figure 2.18). RIPs are generated by converting the two dimensional array of intensity values in the Cartesian coordinate system to a pseudo-radial coordinate system centered about the bead center. In this vi, the inputs are: single frame ROI (image data of the bead), radial intensity array (stores one dimensional radial representation of two dimensional image), radial intensity array hits (saves how many pixels are found on along the radial representation at position R), and ref x and ref y (reference the absolute bead center position in the single frame ROI) (Figure 2.31).

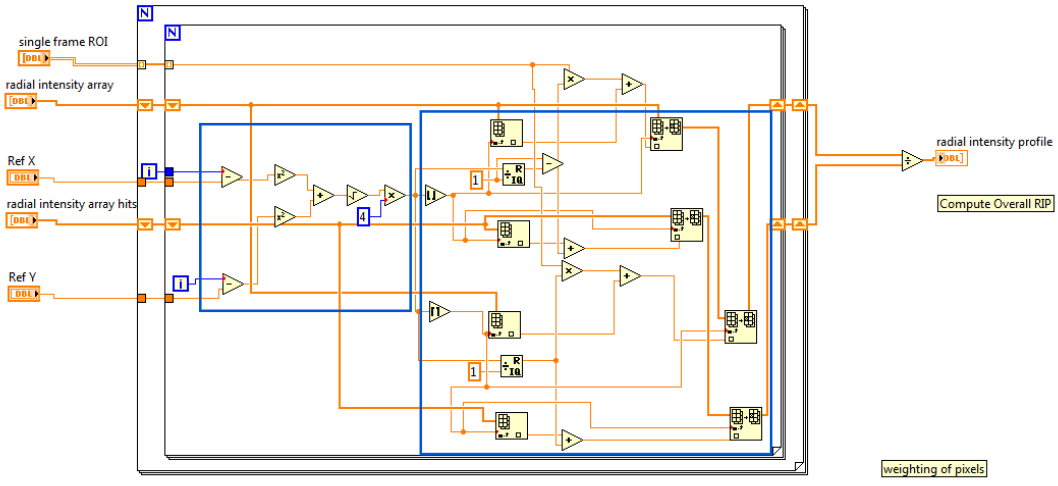


Figure 2.31. Generating the RIP is done by finding each pixel's radial position and binning the pixel intensity. Each pixel is looped over and process during the nested loop visualized in the figure. Using Pythagorean's theorem, the radial position of the current pixel is determined. The contribution of each pixel to discrete radial positions is determined and applied through a weighting function. The data is saved in a one dimensional radial representation of the image, which is averaged and corrected after each pixel contribution is calculated and saved.

Inside the algorithm is a for loop nested in a for loop, where each loop iteration processes a single pixel in the single frame ROI. This loop is very similar to that found in find_bead_center.vi in that the operation is performed on a per pixel basis. During the algorithm, each pixel with integer Cartesian coordinates in the ROI are converted to a subpixel position along the radial axis R (Figure 2.18). This conversion is done by taking the absolute center point within the single frame ROI and using the Pythagorean theorem ($a^2 + b^2 = c^2$) to calculate R for that specific pixel.

For example, if the center position of bead is 23.5,23.2 (Ref X, Ref Y) and we are looking at the first pixel in the ROI [0,0], the initial subtracted values would be -23.5 (a), -23.2 (b). Then the values are squared, added, and square rooted to determine c ($c = \sqrt{a^2 + b^2} = \sqrt{(-23.5)^2 + (-23.2)^2} = 33.02$. The value is then multiplied by 4 as a scaling factor so that the data is spread over four times R, thereby giving subpixel resolution in the RIP (132.09). Next the value must be placed into a RIP array with discrete integer values of R. For example, R does not contain an index at 132.09. Therefore the next part of the algorithm defines how to weight the data into the RIP array. In the case of 132.09, the value is rounded up and down (132 and 133). The intensity value at (0, 0) would need to be placed in the proper array index. In order to account for the sub-integer resolution, the remained of 132.09 is used for

value scaling. In this example, 0.91 is multiplied by intensity at (0, 0) and then added to the RIP array at position 132. Similarly, 0.09 is multiplied by 132.09 and added to RIP position 133. In this way the data can be divided up among the corresponding RIP integer values. A secondary array is saved to denote what fractional intensities were saved in the RIP during the loop procedure (radial intensity array hits). This allows averaging of each RIP index at the end of the looping process. The loops continue over the entire ROI and sum the corresponding R values in the RIP array. At the end of the loops the array values are averaged based on how much data was stored in each element and the completed RIP is returned.

3) Alternating laser excitation (ALEX) implementation through DAQmx

3.1) Integration of a laser control program with the current Andor software

In order to integrate the laser control into the current Andor control software, a subVI was created to initiate the laser control program. The foreground controller of Andor has a laser mode selection (Figure 3.1). Currently, the software is written to recognize the “manual” option or “green/red seq” option. If manual is selected, then (as expected) the lasers remain under manual user control. If green/red seq is selected, then data acquisition (in “Save ‘til abort” camera mode) will trigger automated control of the laser. The “# green” and “# red” variables act as the control mechanism for the user. Before beginning data acquisition, the user should set the values to define the laser sequence. For example, if the user sets #green to 1 and #red to 2, then data acquisition will initiate a laser sequence where the green laser is on for one

frame, the red laser is on for the subsequent two frames, and the process then repeats until data acquisition is complete.

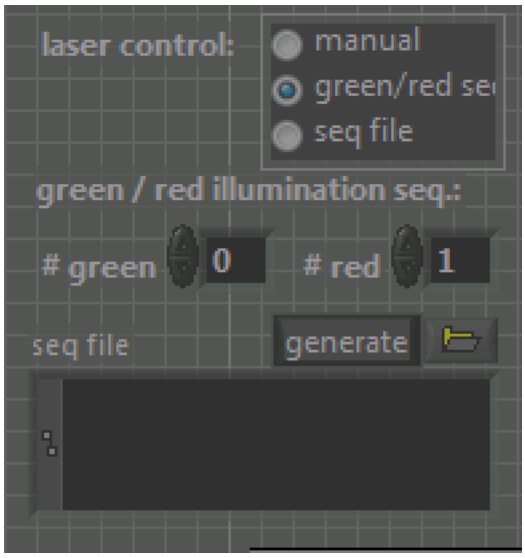


Figure 3.1. Andor camera control contains a built in controller for the laser shutters. The "laser control" radio button switch can be set to manual (for manual laser control) or green/red seq for setting a defined shutter sequence during acquisition. The number controls (# green and # red) set the sequence number for green frames on then red frames on. If # green is set to 1 and # red is set to 2, then the green laser will be on (red off) for one frame and the red will be on (green off) for the next two frames, then the program repeats this sequence until acquisition completes.

In the block diagram, this laser control is initiated by a sub VI in the "Save 'til abort" condition (Figure 3.2). When the "man lsr cntrl" option reads false (front end laser control panel: green/red seq selected), the program will run a start_laser_sequence VI. Although it is not immediately obvious, this VI is not the laser control VI, but rather acts as an invoker of the laser control vi (Figure 3.3). The program was written in this fashion so that the laser control VI will run continuously in an invisible panel until data acquisition is complete.

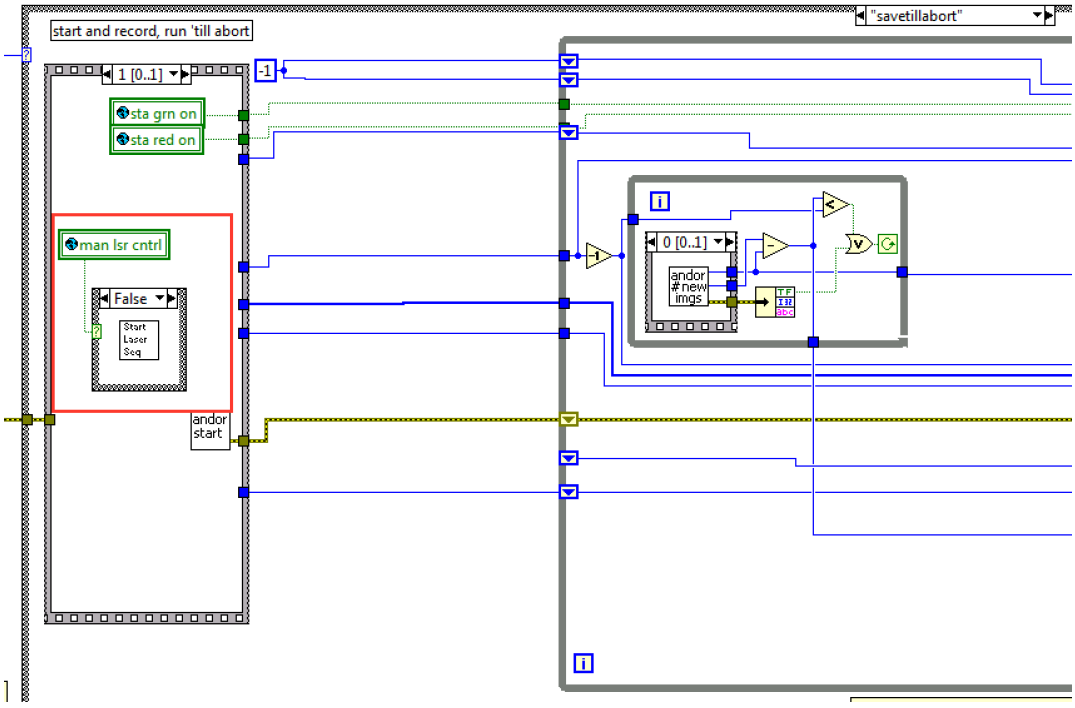


Figure 3.2. Starting camera acquisition will start the laser sequence. If the manual laser control is false, then the user has selected an automated laser sequence. The VI in this conditional will call the alternating shutters VI which controls the shutters based on the user input.

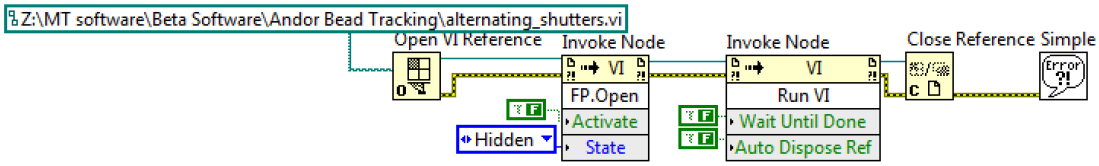


Figure 3.3. Start Laser sequence invokes the alternating shutter VI in the background. Invoking the alternating shutter VI allows background execution of alternating shutters so that the program can be run without display to the user.

3.2) Automated alternating shutter control through a DAQmx controller

The actual automated laser control code operates by using a shutter sequence to modify the current laser states when a camera fire event is detected (Figure 3.4). The control software operates as described: the camera sends a fire event (which is

detected by the DAQmx board) which triggers a new shutter event. The red green sequence input by the user defines how the DAQmx deals with this signal. Let's consider the example where the user input #green 1 and #red 2. The first camera fire event will trigger the green laser to turn on. The second camera fire event will trigger the green laser to turn off and the red laser to turn on. The third camera fire event will trigger the green laser to turn off (already in an off state, so nothing happens) and the red laser to turn on (already in an on state, so nothing happens). Once the user terminates data acquisition, the program controlling the lasers will terminate, thereby terminating the sequence. Originally, I tried to write a program that ran directly through LabVIEW runtime to acquire the camera fire signal and control the shutters, but this attempt was unsuccessful. Surprisingly, the shutter control worked effectively through this methodology, but the camera fire event creates an extremely short signaling event (microseconds in length). This is problematic because sampling rate of the LabVIEW runtime environment (~1 KHz) ported through the DAQmx board is incapable of detecting the rising edge of such a short signal. Fortunately, the DAQmx board is capable of sampling at MHz rates when run autonomously, and is therefore capable of detecting the camera fire event with extreme precision.

In order to run the DAQmx board autonomously from LabVIEW, a DAQ command sequence is generated, fed to the DAQmx board, and run until the program is aborted (Figure 3.4). There are five I/O lines running to the DAQ breakout box: PFI1 (camera fire BNC connector), line0:1 (shutter trigger lines that reset shutter state), and line2:3 (shutter observer lines that return the current state of the shutter).

From the shutter input/output lines, DAQmx tasks are created for digital output (writing new shutter states to lines) and digital input (reading shutter states). These handlers are sent into sampler clocks. The clocks determine how often the DAQmx board performs its task on the I/O lines. In this case, the clock sample rate is set to 10 MHz, which is much faster than the length of time required to resolve the camera fire event.

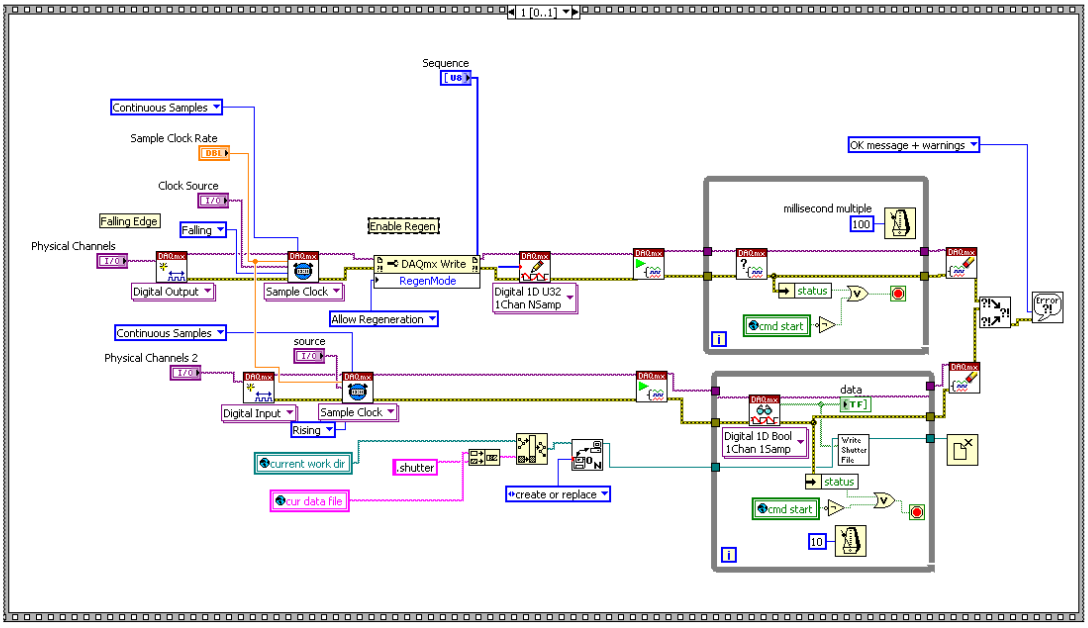


Figure 3.4. Alternating Shutters controls digital input and output to read and write the shutter states. The digital output line is set to trigger off the camera fire (analog input) falling edge. This signal is tested for at a 10 MHz rate. The predefined user sequence is used to write the shutter sequence to the digital output and continues to output the sequence until acquisition completes. The digital input line monitors the shutter state and reads the shutter state on the rising edge of the camera fire event. This process continues until data acquisition completes.

The digital output is programmed to watch the clock source (PF1: camera fire line) for a falling edge, or the end of a camera fire signal. During this time, the DAQ is told to collect continuous samples, meaning it should collect until an abort

command is sent. For the digital output line, the DAQmx Write is modified so that the property RegenMode is set to allow regeneration. This property change is important because it allows the sequence signal to be repeated once the sequence array is exhausted. Once the property is changed, the write parameter is set to write the sequence from the sequence array as a Digital 1D U32 1 Channel N Sample. This specific writing process will send a signal over the fed in lines (line0:1) corresponding to the binary representation of the sequence value. For example, the sequence array defaults to sequence [0] = 2 and sequence [1] = 1. In this case, the first falling edge of the clock channel will trigger a written value of 2 to the shutter lines. The value 2 will be written to the DAQ as a binary number, where 2 is equivalent to 10 in binary representation. For this command then, line0 will be sent a 1 command (digital on) and line1 will be sent a 0 command (digital off). So the first camera fire event will result in green shutter open, red shutter closed. The second value of the sequence array (decimal format 1) will be sent to the DAQ board as binary representation 01. In this case, line0 will be written with 0 command (digital off) and line1 will receive 1 (digital on). Of course, this corresponds to green shutter closed, red shutter open. So with the default sequence array, the camera will provide the repeating sequence [green: on, red: off], [green: off, red: on]. Finally, the DAQ start task VI is used to initiate the signaling on these digital output lines. The loop following this command serves to keep the DAQ lines operating until an error is thrown by the DAQ or the camera acquisition is completed, after which the DAQ task is terminated.

The digital input lines (for reading the shutter state) are initiated similarly to digital output lines with the difference of trigger point. In the case of the digital input lines, the Sampling clock will trigger a read event on the digital output lines at the rising edge of the camera fire event. Modifying the trigger edge prevents convolution of the shutter control and shutter state reading. A keen reader might wondering why even read the shutter state at all, shouldn't the state be implied by the digital input triggering? In an ideal case, this supposition would be true but in some cases the DAQ fails to send a signal or the shutters fail to recognize a signal leading to a desynchronization of the shutter sequence. If the shutter sequence is continually read, then if, for instance, two green frames occur in a row instead of the intended single green frame, the user will be able to parse out the error in instrument control during data analysis. Once the input clock is set and the DAQ task is started, a loop controls the data reading of the DAQ FIFO (First-In First-Out). Since the loop runs with a 10 ms timer, the DAQ input buffer should never overflow before they are read (as long as the camera frame-rate does not exceed 100 fps). These values are written to a text file so that the user can later parse the fluorescence time traces based on the camera shutter states. After data acquisition is complete, this digital input channel is closed and the program terminates.