UNIVERSITY OF CALIFORNIA

SANTA CRUZ

ASSEMBLY AND DYNAMICS OF THE *TETRAHYMENA THERMOPHILA* TELOMERASE HOLOENZYME

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

DOCTOR OF PHILOSOPHY

in

MOLECULAR, CELL, and DEVELOPMENTAL BIOLOGY

by

Benjamin M. Akiyama

June 2014

The Dissertation of Benjamin M. Akiyama is approved:

_____

Professor Michael D. Stone, Chair

_____

Professor Harry F. Noller

_____

Professor Alan M. Zahler

_____

Professor Melissa S. Jurica

_____

Tyrus Miller
Vice Provost and Dean of Graduate Studies

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER I

CHAPTER II

CHAPTER III

CHAPTER IV

CHAPTER V

CHAPTER VI

CHAPTER VII

LIST OF TABLES

CHAPTER III

CHAPTER IV

Abstract

Benjamin Akiyama

Assembly and dynamics of the *Tetrahymena thermophila* telomerase holoenzyme

Telomerase is an enzyme that maintains telomeres, the protective structures at the ends of chromosomes. Telomerase dysfunction is associated with premature aging disorders and inappropriate telomerase activation is associated with ~90% of all cancers, which has motivated studies to better understand the function of this crucial enzyme. Telomerase has two essential components, the protein component telomerase reverse transcriptase (TERT) and the telomerase RNA (TER). TERT functions by reverse transcribing a small template region of TER into telomere DNA. TER also has several conserved secondary and tertiary structures known to be important for telomerase function.

In this thesis research, I describe our attempts to better understand how conserved elements of TERT and TER assemble to form a functional enzyme. In Chapter II, using the well-characterized model system *Tetrahymena thermophila*, I describe our characterization of the assembly protein p65. p65 functions in *Tetrahymena* by binding TER and reorganizing important RNA motifs, placing them in the correct orientation to facilitate TERT binding. We identified a C-terminal domain of p65 which is necessary and sufficient for this activity. We then further characterized this domain, demonstrating that it binds a conserved region of the RNA and affects basepairing interactions in an important stem-loop region of the RNA. This work was originally published in the journal *RNA*.

In Chapter III, I further explain our efforts to characterize TERT-TER interactions. In this work we used a biochemical technique known as site-directed hydroxyl radical probing to identify sites of interaction between protein and RNA within the enzyme. Our work uncovered a specific interaction between a conserved RNA element known as the template boundary element (TBE) and a conserved protein element known as the CP2 motif. Mutagenesis experiments identified the most crucial amino acid determinants of the CP2 motif and demonstrated that CP2 motif mutations severely affect telomerase activity. This work originally appeared in the *Journal of Biological Chemistry*.

Another focus of my thesis work has been telomerase dynamics. Telomerase activity involves coordinated motions between protein, RNA, and DNA subunits within the holoenzyme. Understanding these dynamics is crucial to our understanding of telomerase function. In order to study dynamics within the telomerase holoenzyme, we used a technique known as single-molecule Förster resonance energy transfer (smFRET). FRET describes the distance-dependent energy transfer between a donor and acceptor fluorophore at sub-nanometer distances. When these fluorophores are site-specifically incorporated into a molecule of interest, real-time changes in FRET can be used to measure conformational changes within the molecule.

In Chapter IV, I describe a smFRET telomerase binding assay we developed to measure conformational motions during telomerase catalytic activity. We identified reciprocal motions in telomerase RNA on either side of the template that occur during telomerase extension activity. Using smFRET, we demonstrated that RNA on either

side of the template undergoes a cycle of expansion and compactions as template RNA is pulled through the active site. This work was originally published in *Nature Structural and Molecular Biology.*

In Chapter V, I describe smFRET experiments which demonstrate that DNA bound to telomerase is an equilibrium between two conformations: an active state and an alternative state. Furthermore, we discovered that a conserved domain of TERT, the telomerase essential N-terminal (TEN) domain, is responsible for stabilizing the active state and TEN domain mutations which favor the alternative state disrupt the rate of telomerase processivity. Our work conclusively identifies a new DNA binding mode for telomerase and demonstrates an unappreciated function for a conserved telomerase protein domain.

In Chapter VI of this thesis, I describe preliminary work we have performed to further study transitions between the active and the alternative state described in Chapter V. I describe efforts to develop new methodologies to study FRET dynamics across the DNA primer, as well as our efforts to study protein-DNA and protein-RNA motions. These experiments are too preliminary to yield conclusive results, however they demonstrate great promise for future smFRET studies of telomerase.

In the final chapter of this thesis, I describe potential future experiments on telomerase. Having extensively discussed the TEN domain in Chapters V and VI, I focus on potential studies on real-time measurements of telomerase extension activity. I also discuss experiments on further structural characterization of protein-

RNA interactions in the telomerase holoenzyme. I feel both of these topics merit further study in order to better understand the mechanism of this important enzyme.

Acknowledgements

I am incredibly grateful for the support, guidance, and encouragement of Professor Michael Stone, who has been an excellent mentor and advisor. I would also like to thank my thesis committee—Professors Harry F. Noller, Alan M. Zahler, and Melissa S. Jurica—for all of their advice and assistance. I have been very fortunate to have excellent colleagues and friends in the Stone Lab who have provided an amazing work environment, and I am also thankful to have shared my time as a graduate student with an amazing group of friends. Finally, I would like to thank my parents, Harry and Linda Akiyama, for all of their love, encouragement, and support.

Chapter I: Introduction

In the 1970s, scientists working on uncovering the mechanism of DNA replication made an interesting discovery. Based on the existing model of DNA replication, in which DNA strands are elongated in fragments originating from another 3' DNA end, the ends of linear chromosomes should be shortened with each round of replication. It was recognized that some mechanism must exist to protect against DNA shortening over many generations and James Watson coined the term "the end replication problem" to describe the problem of how cells might protect against this effect[1]. The Russian theoretical biologist Alexey Olovnikov made another crucial connection, suggesting that the gradual shortening of linear DNA during DNA replication could be connected to the limited number of cell divisions observed in cell culture and to the phenomenon of aging[2].

In order to solve the mystery of the end replication problem, scientists first had to discover what unique properties exist at the ends of chromosomes. Elizabeth Blackburn, working in the laboratory of Joseph Gall, set out to sequence the ends of the chromosomal DNA. In order to accomplish this, she utilized the ideal model organism, the ciliated protozoan *Tetrahymena thermophila*. This single-celled organism has a unique biology, in which it contains two nuclei, a macronucleus and a micronucleus. The macronucleus contains thousands of chromosomal fragments[3]. As a result, this represented a natural choice to study DNA ends as it contained an enriched source of chromosomes. It was discovered that the DNA at the ends of

chromosomes contains a unique sequence (TTGGGG in *Tetrahymena*) which is repeated many times in succession[4].

These repeated G-rich DNA sequences were subsequently found in many other organisms, including humans which contain the sequence TTAGGG. Later experiments demonstrated that the repeated sequences form a scaffold for DNA binding proteins, forming structures known as telomeres. The telomeres protect the ends of the chromosomes from the DNA damage response[5]. These sequences also explain how they could resolve the end replication problem. If the cell could maintain a long stretch of a non-coding DNA sequence at the ends of chromosomes, it could keep the coding DNA sequences from being shortened during DNA replication, so long as the telomere DNA sequence remained long enough to protect the coding DNA.

Having discovered the telomeric DNA sequence, the next step was to identify the factor which synthesized the telomere. Carol Greider, working in the laboratory of Elizabeth Blackburn, managed to biochemically purify an enzyme from *Tetrahymena thermophila* that was capable of adding telomeric DNA sequences to a DNA primer *in vitro*[6]. It was discovered that a 159 nucleotide RNA co-purified with this activity, suggesting that the enzyme, known as telomerase, contained both a protein and an RNA component. Carol Greider was able to clone and sequence this RNA, and identified that it contained a 6-nt region that was complementary to the telomere DNA[7]. Subsequently, Joachim Lingner, working in Thomas Cech's laboratory, identified the protein component of telomerase[8].

Over the next several years the molecular details of telomerase began to emerge. To summarize, the essential telomerase components contain two factors: the protein telomerase reverse transcriptase (TERT) and the telomerase RNA (TER). TERT functions by repetitively reverse transcribing a template region of TER into telomere DNA, elongating the telomere to prevent telomere loss[7]. In addition to the template, telomerase RNA contains many other conserved motifs (Figure 1A). These include a site of strong protein interaction, known as the template boundary element (TBE) and a three-stranded RNA structure known as a pseudoknot[9]. Pseudoknots are ubiquitously found in TERs, however it is unclear the exact functional role of this conserved tertiary RNA structure.

The TERT protein also has a conserved domain architecture (Figure 1B). TERT consists of four domains: the telomerase N-terminal (TEN) domain, the RNA-binding domain (RBD), the reverse transcriptase (RT) domain, and a C-terminal extension (CTE). The conserved domains have distinct functions. The TEN domain cross-links with the telomeric DNA substrate, suggesting a site of DNA interaction[10]. The telomerase RBD binds with high affinity to TER, suggesting it is the major site of RNA interactions[11]. Finally, the reverse transcriptase domain has strong homology to viral reverse transcriptases and contains the active site for telomerase catalytic activity[8].

The importance of telomerase was further highlighted when a sensitive PCR-based assay for measuring telomerase activity in living samples was developed. This assay was used to measure telomerase in human cancer biopsies and in normal tissue

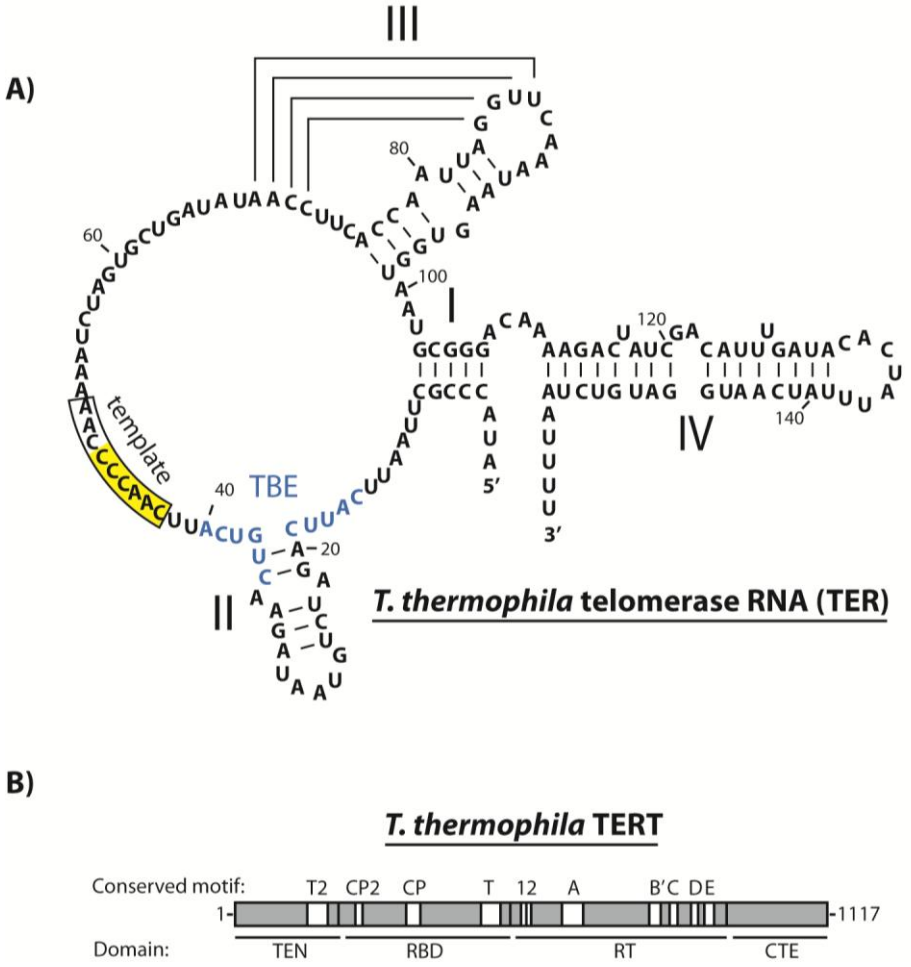


Figure 1. Diagram of the architecture of *Tetrahymena thermophila* TER and TERT. **A)** Secondary structure of *T. thermophila* TER, demonstrating conserved features such as the template (yellow box), the template boundary element (blue), and the pseudoknot (stem III). **B)** Domain architecture of *T. thermophila* TERT, demonstrating the conserved telomerase essential N-terminal (TEN) domain, RNA binding domain (RBD), reverse transcriptase (RT) domain, and C-terminal extension (CTE). The figure also demonstrates the location of conserved motifs within these domains.

and it was discovered that telomerase was over-expressed in ~90% of human cancer biopsies compared to control samples[12]. This suggested that telomerase could be an important drug target for future cancer therapies, promoting further interest in understanding this unique enzyme. It also became apparent that mutations to

telomerase components can cause disease, such as the heritable diseases dyskeratosis congenita and aplastic anemia, which are marked by the progressive loss of proliferative tissues[13]. Taken together, these pieces of information demonstrate that telomerase is essential for the natural proliferation of rapidly dividing cells, but can also be exploited by cancer cells for their unchecked proliferative potential.

Telomerase catalytic activity itself can be divided into two separate activities: a process called nucleotide addition processivity (NAP) and a process called repeat addition processivity (RAP) (Figure 2A). NAP is analogous to a canonical reverse transcriptase reaction, in which the template RNA is used to copy one nucleotide at a time into a substrate DNA. RAP is more complicated. When the end of the template is reached, the enzyme pauses, denatures the DNA-RNA duplex, and repositions the template RNA[7]. RAP requires the coordinated motions of DNA, RNA, and protein components and the mechanism by which this occurs is unknown. RAP is also the rate-limiting step of telomerase activity, and telomerase activity assays demonstrate strong bands every six nucleotides corresponding to the pauses found at the end of each round of NAP (Figure 2B).

Recently, several high resolution structures have emerged of telomerase protein and RNA components. These include structures of the human RNA pseudoknot and the *Tetrahymena* RNA components of the TBE and stem IV[14-17]. For TERT, structures have been solved for the *Tetrahymena* TEN domain and RBD[18,19]. The most complete TERT structure to date is the full-length structure of TERT from *Tribolium castaneum*[20]. This model organism lacks a TEN domain and does not have

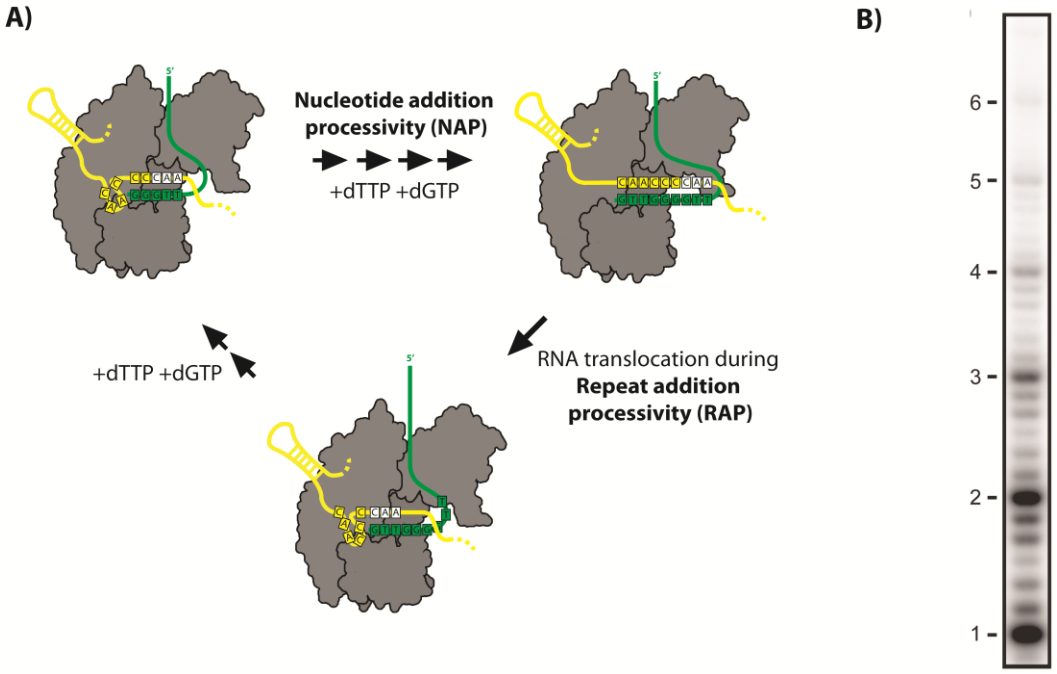


Figure 2. Diagram of telomerase catalytic cycle. **A)** Telomeric DNA (Green) binds to telomerase RNA (yellow) through the template and alignment residues (yellow and white boxes). In the presence of dNTPs, telomerase undergoes nucleotide addition processivity (NAP), successively elongating the DNA primer one nucleotide at a time using the RNA as a template. When the end of the template is reached, the enzyme encounters a steric block, pauses, and translocates the template RNA to position the enzyme to copy another repeat. This process is known as repeat addition processivity (RAP). **B)** Telomerase extension assay, demonstrating pauses at each round of RAP (labeled with numbers).

a well-characterized associated RNA. However, the structure contains the RBD, RT, and CTE domains of this important protein. The structure revealed that TERT adopts a doughnut shape. The RT domain contains a conserved "primer-grip" organization with a central channel to grip the DNA-RNA duplex. The RBD is positioned at the N-terminus of the RT, appearing to be optimally positioned to bind the TBE near the template RNA exiting the active site of the enzyme. The structure also reveals a

narrow channel near the active site of the enzyme where dNTPs can diffuse into the active site.

The emergence of high resolution structural work has been important to our understanding of telomerase function. However, many gaps remain in our understanding. Largely these gaps exist because although we now have structural information for many of the important fragments of TERT and TER, we do not have an appreciation for how the pieces fit together. As a result many crucial questions remain about telomerase assembly. A large portion of this thesis focuses on understanding how telomerase RNA and protein fragments associate in the *Tetrahymena thermophila* holoenzyme. We have chosen the *Tetrahymena* model system due to the availability of biochemical information and high resolution structures in the enzyme[11,18,19,21-23]. It is important to note that both the TERT and TER subunits of telomerase have many conserved features between the *Tetrahymena* and human systems, suggesting that knowledge about *Tetrahymena* TERT-TER interactions should also be applicable to humans[24].

In Chapter II of this thesis, I focus on how the assembly protein p65 promotes the association of TERT and TER subunits. p65 is an assembly protein in *Tetrahymena thermophila* which acts by inducing a bend in stem IV RNA, positioning protein-binding sites in the RNA for TERT association[25]. p65 is a La domain protein, containing adjacent La and RRM motifs, which are known to associate with 3' uridine residues on pol III transcribed RNAs[26]. p65 also contains at the time uncharacterized N- and C-terminal domains[27]. We set out to identify the

functions of the various N-terminal, La, RRM, and C-terminal domains of p65. We discovered that the C-terminal domain of p65 is necessary and sufficient to induce the structural rearrangement associated with p65 function. Furthermore, using RNase protection mapping we determined that the p65 C-terminal domain directly binds a conserved GA bulge on stem-loop IV important for TER function. Furthermore, the C-terminal domain serves to reorganize stem-loop IV residues important for TERT binding. A subsequent high-resolution crystal structure of p65 in complex with an RNA substrate confirmed many of our findings[28].

In Chapter III we continue our studies of TERT-TER assembly, in this case focusing on which regions of TERT associate with which regions of TER. Previous work had demonstrated that the RBD of TERT is responsible for the majority of TER interactions in *in vitro* binding studies[11]. Biochemical studies had also identified conserved regions of TER important for RBD association[22,23]. However, it was not appreciated at the time which regions of the TERT RBD interacted with which regions of TER. This information could be very important in future structural modeling efforts. To address this question we used a technique known as site-directed hydroxyl radical probing. In this technique, a single cysteine in the protein of interest is labeled with a Fe-EDTA moiety. In the presence of ascorbate and hydrogen peroxide, hydroxyl radicals are generated, cleaving nearby RNA residues. This technique had previously been applied in the ribosome to map protein-RNA contacts[29].

Our results identified a site of strong interaction between a conserved RBD motif known as the CP2 motif and the TBE of TER. Both regions had previously been implicated to be important for function in mutagenesis experiments[21,22], however it had not been previously demonstrated that they were direct binding partners. We went on to identify the most critical residues in the 12 amino acid CP2 motif and demonstrate the effect of these mutations in telomerase extension assays. We have since used our results to design a new RBD construct for crystallography studies, containing the crucial CP2 motif and initial crystallography studies have yielded crystals grown from protein-RNA complexes.

Another major focus of my thesis work has been on telomerase dynamics. Telomerase activity involves the relative motions of DNA, RNA, and protein components. By understanding these motions we can begin to construct a model of telomerase in action, with the ultimate goal of creating a "movie" of telomerase activity. In order to study telomerase dynamics, we have developed a single-molecule Förster resonance energy transfer (smFRET) assay for telomerase. FRET descrives the transfer of energy from a donor fluorophore to an acceptor fluorophore over distances of 0-100 Å. Because FRET is highly distance dependent, occurs at molecular-scale distances, and can be measured in real time, it represents an ideal method to study dynamics in biomolecules such as telomerase[30]. By studying FRET at the single-molecule level, we can study asynchronous or complex populations of molecules, eliminating averaging effects. Single-molecule studies are also beneficial

because they require less material and telomerase is not highly expressed in *in vitro* reconstitution reactions.

Chapter IV of this thesis focuses on understanding telomerase RNA motions of RNA on either side of the RNA template. Initial RNA mutagenesis studies on telomerase suggested that the region 3' of the template RNA, known as the template recognition element (TRE), has a critical length requirement for early telomerase catalytic intermediates, but not for late intermediates. This critical length dependence for different intermediates in the catalytic cycle, suggested that the RNA backbone of the TRE may undergo a cycle of compression and expansion during the telomerase catalytic cycle. We designed a smFRET assay to measure the length across both the TRE and across the template to the 5' end of the template (the TBE). This work demonstrated that reciprocal cycles of compression and expansion in the telomerase RNA backbone occur on either side of the template during catalysis, a model dubbed "the RNA accordion model." We also proposed that a strain in the RNA backbone could potentially be used to drive RNA translocation during RAP, giving a further functional significance to the RNA accordion.

Chapter V of this thesis studies another aspect of telomerase dynamics. smFRET studies of telomerase uncovered an equilibrium between two states for DNA bound to telomerase, an active state and a newly-discovered alternative state. Further characterization of these two states suggested that the primary role of the TERT TEN domain is to stabilize the active state of the enzyme. This was found to be particularly true for early catalytic intermediates in the telomerase extension cycle which have a

shorter number of basepairs between the telomeric DNA substrate and the template RNA to hold the DNA in the active site.

Having established the role of the TEN domain in stabilizing short DNA-RNA duplexes in the active site, in Chapter VI I describe our efforts to further characterize the transitions between the active and alternative states. To do so we have incorporated new smFRET techniques to measure motions across the DNA, between DNA and protein, and between protein and RNA labeling sites. This work has shown promising preliminary results, but requires additional data to fully develop our conclusions. The technology we have developed will likely be useful in other smFRET studies of telomerase as well.

In the final chapter, I describe ideas for future experiments on the telomerase holoenzyme. This includes my thoughts on monitoring telomerase activity in real time using smFRET, in addition to potential experiments on the role of the RNA pseudoknot in telomerase activity. Finally, I also describe some potential crystallography studies.

References

1. Watson, J.D. Origin of concatemeric T7 DNA. *Nat New Biol* **239**, 197-201 (1972).

2. Olovnikov, A.M. A theory of marginotomy. The incomplete copying of template margin in enzymic synthesis of polynucleotides and biological significance of the phenomenon. *J Theor Biol* **41**, 181-90 (1973).

3. Prescott, D.M. The DNA of ciliated protozoa. *Microbiol Rev* **58**, 233-67 (1994).

4. Blackburn, E.H. & Gall, J.G. A tandemly repeated sequence at the termini of the extrachromosomal ribosomal RNA genes in Tetrahymena. *J Mol Biol* **120**, 33-53 (1978).

5. Palm, W. & de Lange, T. How shelterin protects mammalian telomeres. *Annu Rev Genet* **42**, 301-34 (2008).

6. Greider, C.W. & Blackburn, E.H. Identification of a specific telomere terminal transferase activity in Tetrahymena extracts. *Cell* **43**, 405-13 (1985).

7. 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).

8. Lingner, J. et al. Reverse transcriptase motifs in the catalytic subunit of telomerase. *Science* **276**, 561-7 (1997).

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

10. Lue, N.F. A physical and functional constituent of telomerase anchor site. *J Biol Chem* **280**, 26586-91 (2005).

11. Lai, C.K., Mitchell, J.R. & Collins, K. RNA binding domain of telomerase reverse transcriptase. *Mol Cell Biol* **21**, 990-1000 (2001).

12. Kim, N.W. et al. Specific association of human telomerase activity with immortal cells and cancer. *Science* **266**, 2011-5 (1994).

13. Vulliamy, T.J. & Dokal, I. Dyskeratosis congenita: the diverse clinical presentation of mutations in the telomerase complex. *Biochimie* **90**, 122-30 (2008).

14. Theimer, C.A., Blois, C.A. & Feigon, J. Structure of the human telomerase RNA pseudoknot reveals conserved tertiary interactions essential for function. *Mol Cell* **17**, 671-82 (2005).

15. Richards, R.J., Theimer, C.A., Finger, L.D. & Feigon, J. Structure of the Tetrahymena thermophila telomerase RNA helix II template boundary element. *Nucleic Acids Res* **34**, 816-25 (2006).

16. Richards, R.J. et al. Structural study of elements of Tetrahymena telomerase RNA stem-loop IV domain important for function. *RNA* **12**, 1475-85 (2006).

17. Chen, Y. et al. Structure of stem-loop IV of Tetrahymena telomerase RNA. *EMBO J* **25**, 3156-66 (2006).

18. 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**, 218-25 (2006).

19. Rouda, S. & Skordalakes, E. Structure of the RNA-binding domain of telomerase: implications for RNA recognition and binding. *Structure* **15**, 1403-12 (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. Miller, M.C., Liu, J.K. & Collins, K. Template definition by Tetrahymena telomerase reverse transcriptase. *EMBO J* **19**, 4412-22 (2000).

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

23. Lai, C.K., Miller, M.C. & Collins, K. Roles for RNA in telomerase nucleotide and repeat addition processivity. *Mol Cell* **11**, 1673-83 (2003).

24. Lin, J. et al. A universal telomerase RNA core structure includes structured motifs required for binding the telomerase reverse transcriptase protein. *Proc Natl Acad Sci U S A* **101**, 14713-8 (2004).

25. Stone, M.D. et al. Stepwise protein-mediated RNA folding directs assembly of telomerase ribonucleoprotein. *Nature* **446**, 458-61 (2007).

26. Bayfield, M.A., Yang, R. & Maraia, R.J. Conserved and divergent features of the structure and function of La and La-related proteins (LARPs). *Biochim Biophys Acta* **1799**, 365-78 (2010).

27. Prathapam, R., Witkin, K.L., O'Connor, C.M. & Collins, K. A telomerase holoenzyme protein enhances telomerase RNA assembly with telomerase reverse transcriptase. *Nat Struct Mol Biol* **12**, 252-7 (2005).

28. Singh, M. et al. Structural basis for telomerase RNA recognition and RNP assembly by the holoenzyme La family protein p65. *Mol Cell* **47**, 16-26 (2012).

29. Heilek, G.M., Marusak, R., Meares, C.F. & Noller, H.F. Directed hydroxyl radical probing of 16S rRNA using Fe(II) tethered to ribosomal protein S4. *Proc Natl Acad Sci U S A* **92**, 1113-6 (1995).

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

CHAPTER II: The C-terminal domain of *Tetrahymena thermophila* telomerase holoenzyme protein p65 induces multiple structural changes in telomerase RNA

(originally published in the journal *RNA*)

Abstract

The unique cellular activity of the telomerase reverse transcriptase ribonucleoprotein (RNP) requires proper assembly of protein and RNA components into a functional complex. In the ciliate model organism *Tetrahymena thermophila,* the La-domain protein p65 is required for *in vivo* assembly of telomerase. Single-molecule and biochemical studies have shown that p65 promotes efficient RNA assembly with the telomerase reverse transcriptase (TERT) protein in part by inducing a bend in the conserved stem IV region of telomerase RNA (TER). The domain architecture of p65 consists of an N-terminal domain, a La-RRM motif, and a C-terminal domain (CTD). Using single molecule Förster resonance energy transfer (smFRET), we demonstrate the $p65^{CTD}$ is necessary for the RNA remodeling activity of the protein and is sufficient to induce a substantial conformational change in stem IV of TER. Moreover, nuclease protection assays directly map the site of $p65^{CTD}$ interaction to stem IV and reveal that in addition to bending stem IV, p65 binding reorganizes nucleotides that comprise the low affinity TERT binding site within stem-loop IV.

Introduction

Telomeres protect the ends of linear chromosomes from damage due to homologous recombination and other DNA repair mechanisms [1]. They consist of a repetitive DNA sequence, $(T_2AG_3)_n$ in humans and $(T_2G_4)_n$ in *Tetrahymena thermophila*, which forms a scaffold for DNA-binding proteins that shelter the DNA. Intact telomeres are essential for cell viability; short telomeres are associated with cellular senescence and telomere loss can lead to genomic instability and cell death [2].

Telomeres are progressively shortened during cell division due to the end replication problem; therefore, eukaryotic cells require the specialized enzyme telomerase to maintain telomere length over the course of multiple rounds of cell division. Telomerase dysfunction is associated with the disorders Dyskeratosis Congenita (DKC) and Aplastic Anemia (AA), which are marked by symptoms that disproportionately affect proliferative tissues [3]. Conversely, telomerase over-expression can also have negative consequences, helping to confer a high level of proliferative potential upon cells as evidenced by the presence of telomerase activity in ~90% of human cancer cell lines [4].

Telomerase is a ribonucleoprotein (RNP) enzyme, minimally composed of two elements: a protein subunit called telomerase reverse transcriptase (TERT) and the telomerase RNA (TER). TERT functions by repetitively reverse transcribing a short template region of TER into telomeric DNA sequence [5]. In addition to containing a template, all known TERs share a conserved structural organization

including: a template boundary element (TBE), pseudoknot domain, and a stem terminal element (STE) [6,7].

Due to the naturally low abundance of both TERT and TER, telomerase RNP assembly presents a serious challenge to the cell. In humans, many mutations associated with the telomerase dysfunction disorder DKC act at the level of telomerase assembly co-factors and not in TERT or TER [3]. Indeed, assembly co-factors play an essential role in telomerase assembly in many different model organisms. In vertebrates the dyskerin/NHP2/NOP10 complex is required for telomerase biogenesis, in yeast it is Sm proteins that facilitate telomerase assembly, and in ciliates La domain proteins are required for assembly [8-11].

In the well-studied ciliate model organism *Tetrahymena thermophila*, the La-domain protein p65 facilitates telomerase assembly *in vivo* and is required for telomere length maintenance [11]. *In vitro* biochemical studies showed p65 improves the K_d of the interaction between TERT and TER [12], mapped the site of p65-TER interaction to stems I and IV of the RNA [13], and demonstrated the presence of p65 could rescue the catalytic activity of certain telomerase mutants [14]. A recent study investigating the molecular mechanism of p65-induced telomerase RNP assembly employed single molecule Förster resonance energy transfer (smFRET), an approach that exploits the distance-dependent energy transfer from a donor fluorophore to an acceptor fluorophore to directly measure conformational changes in biological macromolecules. This work revealed a hierarchal mechanism for telomerase RNP assembly, wherein p65 binds TER and induces a bend in TER stem IV. This positions

separate TERT-binding elements in the RNA in the correct orientation for assembly, facilitating the TERT-TER interaction [15]. Moreover, these studies demonstrated an essential role for the evolutionarily conserved GA bulge within TER stem IV during p65-directed telomerase RNP assembly, consistent with a structural characterization of this GA bulge by NMR spectroscopy that demonstrated conformational flexibility in this region of TER [16].

To further dissect the RNA remodeling activity of p65, we expressed various p65 domain constructs and analyzed their RNA binding properties using smFRET and RNase protection mapping. The results directly demonstrate that the novel C-terminal domain (CTD) of p65 is required for p65-mediated TER conformational rearrangement. RNase protection experiments precisely mapped the site of the $p65^{CTD}$-TER interaction to stem IV. Interestingly, RNase probing experiments also revealed p65 binding to TER induces a substantial reorganization of nucleotides that comprise a low affinity TERT binding site within TER stem-loop IV [17]. This result suggests the mechanism of p65-facilitated telomerase assembly involves precise RNA conformational changes beyond the previously described bending of TER stem IV.

Results:

The p65 C-terminal domain is essential for telomerase RNA remodeling activity

Previous studies of p65 have identified four domains: an N-terminal domain, a La domain, an RNA recognition motif (RRM) domain, and a C-terminal domain (CTD) (Fig. 1A) [12,13]. Neither the N-terminal domain nor the C-terminal domain has

significant sequence homology with known protein motifs. La motifs are frequently found in conjunction with RRM motifs and often bind poly-U tracts at the 3' end of RNA Pol III transcripts [18]. Since *Tetrahymena* TER contains a 3' U-tract and is transcribed by Pol III [5], it is likely the p65 La-RRM domains bind to telomerase RNA in the vicinity of the 3' terminus. A previous study suggested the uncharacterized C-terminal domain and not the La-RRM motifs were required to increase the affinity between p65-bound TER and the RNA binding domain of TERT [13]. However, no direct test of the role of $p65^{CTD}$ in promoting conformational rearrangements of TER has been reported.

We set out to directly assay which domains were required for p65-mediated conformational rearrangement within TER. To this end we expressed and purified full-length p65 ($p65^{FL}$), a p65 N-terminal domain truncation ($p65^{\Delta N}$), a p65 C-terminal truncation ($p65^{\Delta C}$), and the CTD alone ($p65^{CTD}$) (Fig. 1A and 1B). Gel shift assays demonstrated that all constructs effectively bound TER, though $p65^{\Delta N}$ and $p65^{\Delta C}$ had a slightly reduced affinity for TER, while the $p65^{CTD}$ had a further reduced affinity (Fig 1C). We next employed the previously established smFRET assay of p65 activity [15], which measures the efficiency of energy transfer between a donor dye and an acceptor dye incorporated at specific sites within TER. A donor fluorophore was placed on U139 of the native TER sequence and an acceptor dye was placed on U10, flanking the putative p65 binding site (Fig. 2A). To facilitate our single-molecule studies a 5' extension was incorporated into the RNA, which was then hybridized to a

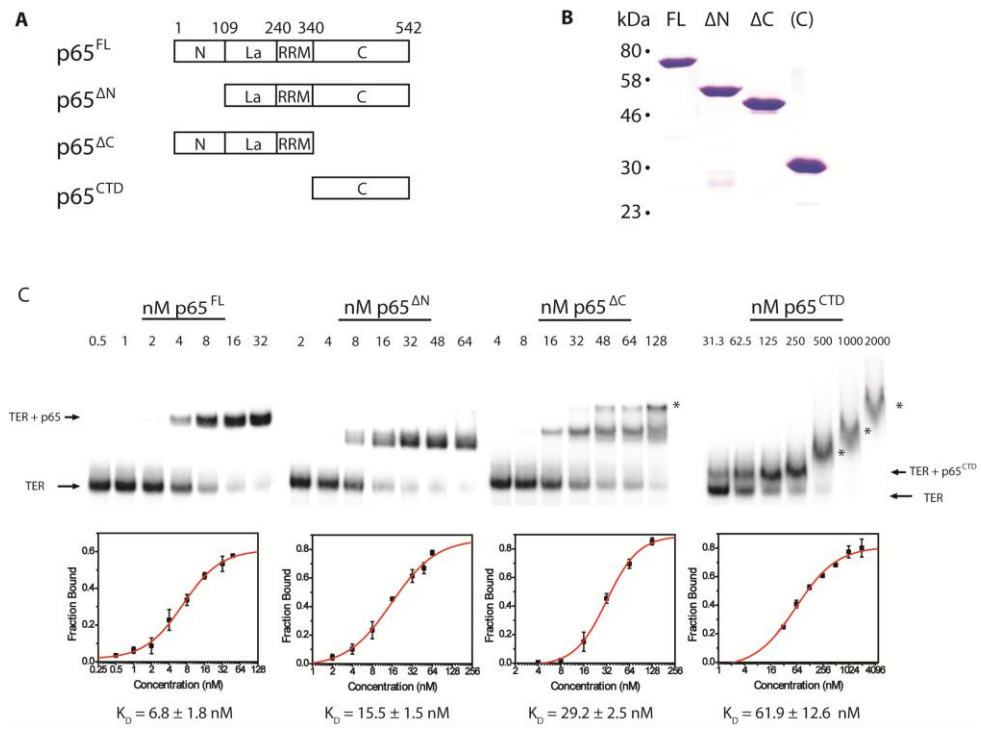


FIGURE 1: A) Schematic diagram demonstrating the domain organization of p65. p65 has four domains: an N-terminal domain, a La motif, an RNA recognition motif (RRM) domain, and a C-terminal domain. Protein constructs were expressed for full-length p65 ($p65^{FL}$), p65 lacking the N-terminal domain ($p65^{\Delta N}$), p65 lacking the C-terminal domain ($p65^{\Delta C}$), and only the C-terminal domain of p65 ($p65^{CTD}$). **B)** A Coomassie stained SDS-PAGE gel of the four purified p65 constructs. Numbers on the left indicate the position of molecular weight markers. **C)** Affinities of p65 constructs for telomerase RNA as determined by gel shift assay. Radiolabeled TER was incubated with increasing concentrations of $p65^{FL}$, $p65^{\Delta N}$, $p65^{\Delta C}$, and $p65^{CTD}$ and run on a native polyacrylamide gel, top panels. Gels were quantified to determine the fraction bound (bottom panels). Error bars represent the standard deviation of triplicate measurements. The red line is a fit of the data used to determine the dissociation constant (Kd) of each p65 construct for its TER substrate. The data was fit to the equation $F = [(F_{max})(c^n)]/[(K_d)^n(c^n)]$, where F represents the fraction bound, c represents the concentration, K_d is the dissociation constant (the concentration at which 50% of the RNA is bound), F_{max} represents the maximal value of F, and n represents the Hill coefficient. The following Hill coefficients were obtained: $p65^{FL}$ n=1.3, $p65^{\Delta N}$ n=1.3, $p65^{\Delta C}$ n=2.1, $p65^{CTD}$ n=0.9. Asterisks mark higher-order protein-RNA complexes formed at high protein concentrations.

short biotinylated DNA oligonucleotide, immobilized onto a streptavidin coated microscope slide, and imaged using prism-type total internal reflection microscopy [19]. The dye-labeled RNA construct was generated using DNA-splinted RNA ligation of chemically synthesized RNA fragments [20] and was shown previously to support wild-type levels of telomerase activity when reconstituted with TERT and p65 *in vitro* [15]. We define FRET as, $I_A / (I_A + I_D)$, where I_A is the intensity of the acceptor dye and I_D is the intensity of the donor dye. In this assay, lower FRET values represent larger inter-dye distances indicative of a more extended stem I and stem IV RNA conformation, while higher FRET values correspond to smaller inter-dye distances and a more compact stem I and stem IV RNA structure.

As expected, smFRET measurements of the immobilized RNA yielded a dominant FRET distribution centered at 0.26 FRET (Fig. 2B), consistent with previously reported results on a similar construct [15]. Also as anticipated, adding increasing amounts of full-length p65 ($p65^{FL}$) to the immobilized RNA shifted a large fraction of the TER molecules to a 0.45 FRET state (Fig. 2C). Having confirmed the smFRET assay provides an accurate measure of p65-mediated conformational rearrangement, we next set out to test which domains of p65 were necessary for its RNA remodeling activity. $p65^{\Delta N}$ retained the ability to induce a conformational change in TER that was quantitatively very similar (FRET ~ 0.43) to the FRET change observed with $p65^{FL}$ (compare Fig. 2C and Fig. 2D). In contrast, $p65^{\Delta C}$ failed to induce any detectable FRET change (Fig.2E), demonstrating that the $p65^{CTD}$ is necessary for p65-mediated structural rearrangement of TER. Importantly,

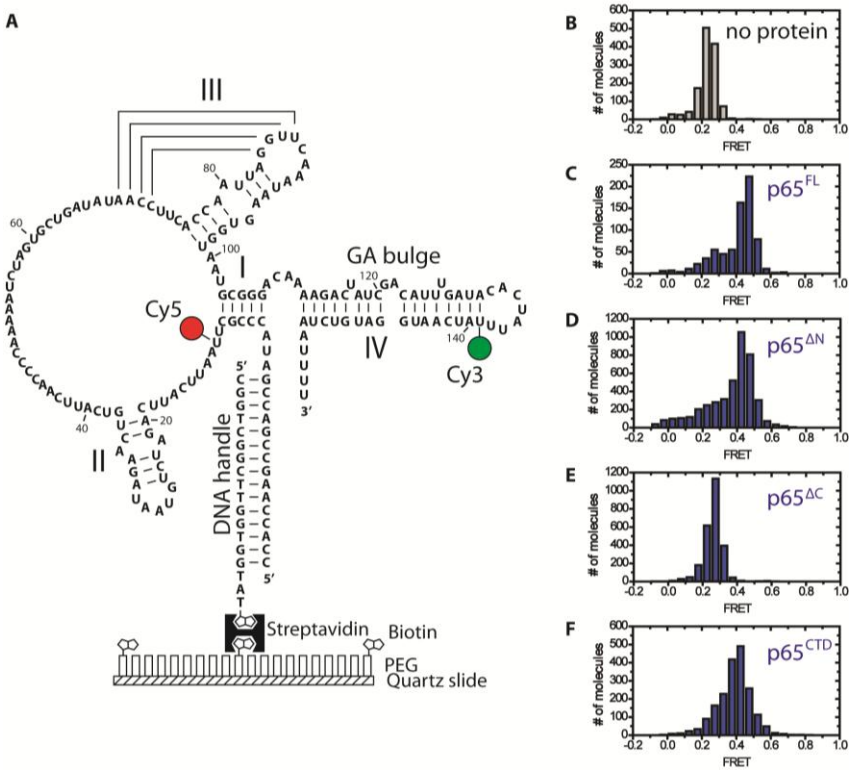


FIGURE 2: A) Diagram of telomerase RNA construct used in the current study. Telomerase RNA was labeled at U139 with a donor fluorophore (Cy3) and at U10 with an acceptor fluorophore (Cy5) for smFRET studies. The RNA construct was deposited on a quartz slide for TIRF microscopy by means of a 5' extension designed to hybridize with a biotinylated DNA handle. The biotinylated handle was immobilized on the quartz slide via a biotin-streptavidin linkage. **B)** Histogram demonstrating the FRET distribution of dye-labeled RNA molecules in the absence of p65. The distribution is centered at 0.26 FRET. **C-F)** Histograms of the FRET distribution of dye-labeled RNA molecules in the presence of 10 nM $p65^{FL}$ **(C)**, 32 nM $p65^{\Delta N}$ **(D)**, 64 nM $p65^{\Delta C}$ **(E)**, or 750 nM $p65^{CTD}$ **(F)**. FRET is defined as $I_A / (I_A + I_D)$, where I_A is the intensity of the acceptor dye and I_D is the intensity of the donor dye. The protein concentrations used were determined by EMSA to have a large fraction of p65-RNA complexes (Fig. 1C).

gel shift analysis determined the concentrations used in our smFRET assay were sufficient to support nearly quantitative binding of $p65^{\Delta C}$ to TER (Fig. 1C).

We next tested whether the $p65^{CTD}$ alone was sufficient to induce a TER conformational rearrangement. This construct induced a stable 0.40 FRET state upon protein binding (Fig. 2F and Supplemental Fig. S1), demonstrating that the C-terminal domain alone does induce an RNA structural change. The difference in FRET values between $p65^{CTD}$ and $p65^{FL}$ suggests the conformational states induced by the respective proteins are similar but not identical to one another. The origin of the difference between the $p65^{FL}$ and $p65^{CTD}$- induced FRET states may reflect the contributions of other p65 domains to the overall RNA fold. Alternatively, the 0.40 FRET state may represent a time averaged measurement of rapid RNA conformational dynamics between the 0.26 and 0.45 FRET states. Importantly, the 0.40 FRET state does not appear to be an artifact of high protein concentrations since concentrations of $p65^{\Delta C}$ as high as 250 nM induced no conformational change as observed by FRET (Supplementary Fig. S2). Interestingly, we observed that addition of $p65^{CTD}$ to $p65^{\Delta C}$-bound TER complexes shifted the FRET distribution to the 0.40 FRET state at lower concentrations than experiments conducted with $p65^{CTD}$ alone (Fig. 3A and Fig. 3B). This result demonstrates that $p65^{CTD}$ can be added in *trans* to the $p65^{\Delta C}$ protein and suggests non-covalent interactions between the $p65^{CTD}$ and the remainder of p65 stabilize C-terminal domain binding.

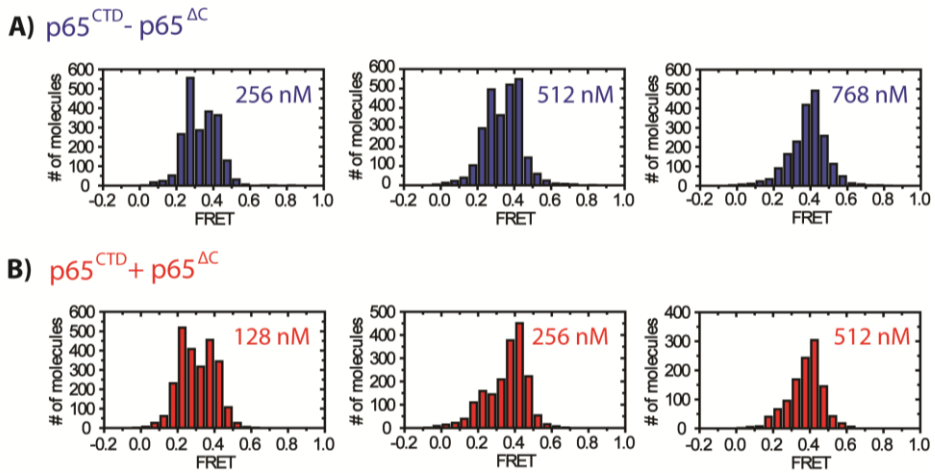


FIGURE 3: Comparison of RNA folding activity between $p65^{CTD}$ and $p65^{CTD}$ co-incubated with $p65^{\Delta C}$**.** The smFRET assay was performed on TER incubated with the indicated concentrations of $p65^{CTD}$ in the absence **(A)** or presence **(B)** of 64 nM $p65^{\Delta C}$. Incubating TER with $p65^{\Delta C}$ increased the affinity of $p65^{CTD}$ for TER, as evidenced by the reduced concentrations of $p65^{CTD}$ required to reach the higher FRET state. This suggests that non-covalent interactions between $p65^{\Delta C}$ and $p65^{CTD}$ stabilize the $p65^{\Delta C}$-$p65^{CTD}$-TER ternary complex.

$p65^{CTD}$ binds stem IV and reorganizes bases within stem-loop IV

To further dissect the mechanism of $p65^{CTD}$-induced RNA folding, we next sought to directly identify the $p65^{CTD}$ binding site using RNase protection experiments. We observed that full-length TER was prone to bind multiple p65 proteins at the high protein concentrations required for our footprinting analysis (Supplemental Fig. S3). For this reason, we employed a previously described RNA construct containing only the putative p65 binding site, stem I and stem IV of TER (Fig. 4A) [13]. The stem I-IV TER construct binds $p65^{FL}$ with comparable affinity to the full-length TER and permitted titrations with higher concentrations of the p65 constructs prior to forming higher order protein complexes that could lead to

misleading footprinting results. Furthermore, the shorter RNA construct facilitated the unambiguous resolution and quantification of all of the relevant nucleotides on a single gel. We digested the stem I-IV TER construct with RNase ONE, a nominally single stranded RNA nuclease that lacks sequence specificity. RNase ONE cleavage of the stem I-IV construct produced cleavage products corresponding to predicted unpaired nucleotides. Surprisingly, we also observed cleavage products corresponding to the top strand of stem IV, but no cleavage was observed for the bottom strand. We hypothesize that the bulged residues present in the top strand of stem IV may permit RNase ONE to cleave in this region.

Curiously, addition of p65 to our binding reaction, even at concentrations insufficient to drive formation of the p65-TER complex, had a stimulatory effect on RNase ONE cleavage (Supplemental Fig. S4). As a result cleavage efficiencies were substantially reduced in the absence of p65 which precluded us from using this condition as a baseline for protection mapping. Therefore, for each individual experiment a protein concentration that showed little to no binding in our gel shift analyses (Supplemental Fig. S5) was used as the reference for quantifying our RNase protection patterns. Importantly, this approach yielded an unambiguous protection trend upon p65 titration across the range of protein concentrations used in this study.

The stem I-IV TER molecules were 5'-end labeled with ^{32}P, incubated with RNase ONE in the presence of increasing concentrations of each protein construct ($p65^{FL}$, $p65^{\Delta N}$, $p65^{\Delta C}$, or $p65^{CTD}$) under single hit conditions, and then resolved on denaturing PAGE sequencing gels. In the case of $p65^{FL}$ a clear footprint emerges at

higher protein concentrations, demonstrating that p65 binds across the GA bulge on stem IV, consistent with this evolutionarily-conserved RNA structural feature being an essential determinant of p65 activity (Fig. 4B) [15,21]. The quantified results reveal that the footprint extends from the GA bulge to the 4-nt linker joining stems I and IV (Fig. 5A, blue squares). No protection was observed in stem I and the lack of RNA cleavage in the 3' poly-U tract precluded our ability to measure any protein protection in this region of the RNA. Strikingly, we observe a strong hyper-sensitivity to cleavage in stem-loop IV (Fig. 4B, asterisk and Fig. 5A, red circles), suggesting that in addition to re-positioning stem-loop IV for TERT-TER interaction, p65 also remodels base-pairing contacts in this region, exposing stem-loop IV to single-stranded RNA digestion. In contrast to this result, a previous RNase ONE probing study with $p65^{FL}$ and the native full-length TER sequence showed protection in the same stem-loop IV residues [14]. We therefore repeated our experiments using full-length TER and found the hypersensitivity to RNase ONE cleavage persisted under our binding and cleavage conditions (Supplemental Fig. S6). The source of the discrepancy between the two studies is unclear, but may be due to differences in the binding and cleavage conditions.

Experiments with stem I-IV bound by $p65^{\Delta N}$ yielded an identical protection pattern to full-length p65 (Fig. 4C and 5B). In contrast, $p65^{\Delta C}$ showed modest hyper-sensitivity throughout stem IV rather than protection, as well as a reduction in the degree of hyper-sensitivity in stem-loop IV (Fig. 4D and 5C). Lastly, protection mapping experiments performed in the presence of $p65^{CTD}$ yielded protection and

hypersensitivity patterns that while slightly less pronounced, are qualitatively very similar to the patterns observed for $p65^{FL}$ (Fig. 4E and 5D). Taken together, these results demonstrate the $p65^{CTD}$ is responsible for the protection of stem IV across the GA bulge, consistent with our smFRET experiments. Furthermore, the nuclease

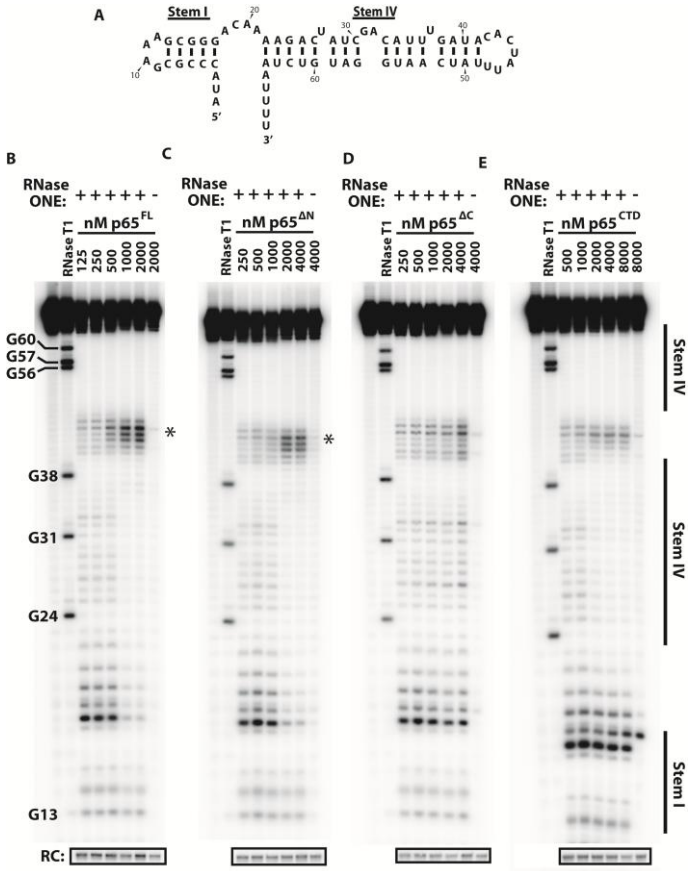


Figure 4: A) Shorter RNA construct used in RNA footprinting studies containing the putative RNA binding site of p65, stems I and IV of *Tetrahymena* telomerase RNA. A GAAA tetraloop has been introduced in stem I in place of the remainder of the RNA. **B-E)** RNase ONE digestion of stem I-IV construct RNA in increasing concentrations of $p65^{FL}$ **(B)**, $p65^{\Delta N}$ **(C)**, $p65^{\Delta C}$ **(D)**, and $p65^{CTD}$ **(E)**. RNase T1 digestion was used to generate a ladder to identify nucleotide position. $p65^{FL}$, $p65^{\Delta N}$, and $p65^{CTD}$ showed strong protection against cleavage in stem IV, suggesting this is the binding site of the C-terminal domain of p65. These constructs also demonstrated hyper-sensitivity in stem-loop IV (asterisks) upon protein binding, suggesting that p65 remodels this region of the RNA. R.C. = recovery control.

mapping experiments reveal a heretofore unappreciated activity of p65: the remodeling of bases in stem-loop IV which comprises a known low affinity TERT binding site.

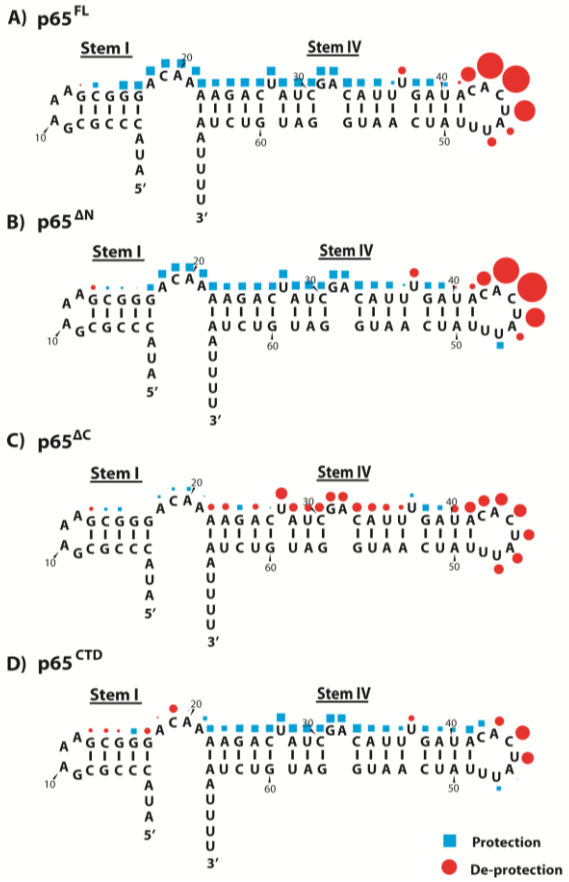


Figure 5: Quantification of RNase ONE protection results displayed in figure 4 for $p65^{FL}$ **(A)**, $p65^{\Delta N}$ **(B)**, $p65^{\Delta C}$ **(C)**, and $p65^{CTD}$ **(D)**. The Semi-automated Footprinting Analysis (SAFA) software package was used to quantify cleavage at individual nucleotides in the RNA. Cleavage was compared at each nucleotide for the highest and the lowest protein concentration to determine the percent protection (blue squares) or de-protection (red circles) upon protein addition. The relative size of the symbols in the figure was scaled to the amount of protection or de-protection observed at that nucleotide. Comparison of the four constructs suggests that the C-terminal domain binds stem IV of telomerase RNA. In addition, binding of p65 to stem IV sensitizes stem-loop IV nucleotides to RNase ONE cleavage.

Discussion

Model of p65 function

Our results demonstrate that the C-terminal domain of p65 is essential for the RNA conformational change associated with p65 activity. smFRET and RNase protection assays reveal that the C-terminal domain alone induces a related, but not identical, set of conformational rearrangements within TER to those observed with full-length p65. Thus while the La-RRM domains have no detectable activity with regard to the RNA remodeling activity of p65, they appear to be important for stabilizing the C-terminal domain induced conformational change and helping to improve the affinity of p65 for its RNA substrate. This suggests a model wherein the La-RRM domains bind the poly-U tract of TER, positioning the C-terminal domain for its lower affinity interaction with stem IV and remodeling the RNA for TERT interaction (Fig. 6). Our RNase footprinting results suggest that p65 disrupts contacts within stem-loop IV as part of its activity, in addition to its previously-identified bending activity on stem IV. NMR spectroscopy structures of stem-loop IV of *Tetrahymena* telomerase RNA [16,22] reveal that in the absence of p65 stem-loop IV residues are highly ordered, with residues C132 and U138 forming a non-canonical basepair and residues C132, A133, and C134 participating in base stacking interactions. These nucleotides show the greatest de-protection upon addition of p65, suggesting that disruption of these interactions may be responsible for the observed RNase hypersensitivity of these residues. This result is also in accord with the recent report that mutations in stem-loop IV which do not alter p65 affinity for TER can

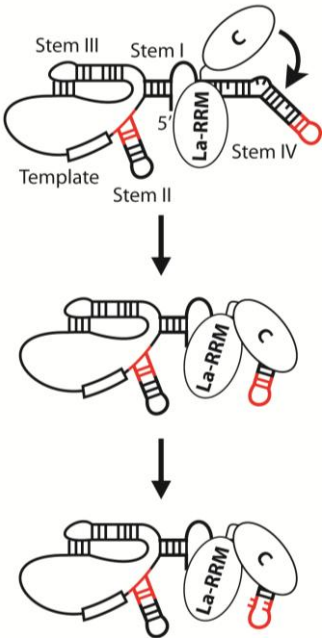


Figure 6: Model of p65 function. The La-RRM domains and the C-terminal domain of p65 interact with separate sites on telomerase RNA (the 3' poly-U tract and stem IV, respectively). La-RRM binding of the poly-U tract helps to target the C-terminal domain to stem IV, which in turn induces the conformational change associated with p65 activity. This positions TERT-interacting elements (red) in the optimal conformation for assembly. p65 binding also induces a re-arrangement in stem-loop IV nucleotides further facilitating TERT-TER assembly.

disrupt the ability of p65 to facilitate TERT-TER interactions [23].

Our smFRET assay and RNase footprinting results have not revealed the function of the N-terminal domain of p65. Our results, in combination with previous studies on the effect of N-terminal deletion on the K_d of the TERT-TER interaction [13], suggest that the N-terminus of p65 is dispensable for the *in vitro* activity of the protein. Other studies have suggested that p65 may have binding activity outside of stem IV, either in stem I or elsewhere in the RNA [13,14]. It may be that the N-terminal domain helps set up those contacts. An additional possibility is that the N-terminal

domain is essential *in vivo*, perhaps playing a regulatory role such as directing nuclear localization or interaction with other proteins.

Materials and Methods

Single-molecule FRET assays

Fluorophore-labeled telomerase RNAs were constructed as previously described [20]. Briefly, synthetic RNA fragments harboring reactive amine modifications at specific residues (Dharmacon) were labeled with mono-reactive Cy3 and Cy5 (GE Lifesciences) and purified by reverse-phase HPLC. A central RNA fragment that did not contain a modification was also generated by targeted RNase H (NEB) cleavage of full-length *in vitro* transcribed telomerase RNA using targeted cutting oligos (IDT, Supplemental Table S1). Fragments were assembled by DNA-splinted RNA ligation. Purified fluorophore-labeled RNAs were annealed to the biotin-labeled DNA handle and immobilized on PEGylated quartz slides and the appropriate p65 construct was flowed over the slide. Slides were imaged on a prism-type total internal fluorescence microscope using an Andor IXON CCD camera with a 100 ms integration time. FRET studies were performed in 20 mM Tris pH 8.0, 100 mM NaCl, 1 mM $MgCl_2$, 0.1 mg/ml BSA, 0.1 mg/ml yeast tRNA, 2 mM Trolox, 10% glucose, 10% glycerol, 4 µg/ml catalase, and 1 mg/ml glucose oxidase. FRET is defined as the ratio of $I_A/(I_A+I_D)$, where I_A is the acceptor intensity and I_D is the donor intensity. Histograms were compiled from the average FRET value obtained from molecules over a 2 second observation time. Molecules with 0.0 FRET, obtained due

to a small fraction of molecules containing bleached acceptor fluorophores, were excluded from histograms by omitting molecules that displayed lower than a minimum threshold value of acceptor intensity.

Electrophoretic Mobility Shift Assays

EMSAs were performed as previously described [24]. Body-labeled TER was transcribed by T7 RNA polymerase (NEB) using α-^{32}P UTP (Perkin-Elmer) and PAGE purified. 100 pM TER was incubated with various dilutions of the appropriate p65 construct and incubated at 30°C for 20 minutes. The binding conditions were 20 mM Tris pH 8.0, 100 mM NaCl, 1 mM $MgCl_2$, 0.1 mg/ml BSA, 0.1 mg/ml yeast tRNA, 1 mM DTT, 10% glycerol. The protein-RNA complex was then loaded on to a 5% 37.5:1 acrylamide: bis-acrylamide gel containing 0.5X TBE and 4% glycerol. The gel was run in 0.5X TBE at 200V at 4°C for 3 hours, dried, and imaged using an Amersham Phosphor Screen and a GE Typhoon Scanner. Band intensities were quantified using Imagequant and were used to determine K_d values. Percent bound complexes were plotted in Origin and K_d values were determined by fitting to the equation $F = [(F_{max})(c^n)]/[(K_d)^n(c^n)]$, where F represents the fraction bound, c represents the concentration, K_d is the dissociation constant (the concentration at which 50% of the RNA is bound), F_{max} represents the maximal value of F, and n represents the Hill coefficient.

Protein expression and purification

Proteins were expressed from a pET28 vector behind a 6x-his tag. $p65^{\Delta N}$ contained residues 109-542 of p65, $p65^{\Delta C}$ contained residues 1-340, $p65^{CTD}$ contained residues 302-542. All p65 constructs were expressed in *E. coli* (BL21 DE3). Cells were lysed using a cell disrupter and cell debris pelleted by centrifugation. p65 was purified from lysate on nickel affinity resin (GE Lifesciences) by means of an N-terminal 6X His-tag. The eluant was further purified by ion exchange chromatography on a Hi-trap Q column (GE lifesciences) and by size exclusion chromatography on a Superdex 200 column (GE lifesciences). All proteins were eluted into a final buffer containing 20 mM Tris pH 8.0, 100 mM NaCl, 1 mM $MgCl_2$, 1 mM DTT, and 10% glycerol, flash frozen in liquid nitrogen, and stored at -70°C for future use.

RNase Protection assays

Stem I-IV TER and full-length TER were synthesized by T7 RNA polymerase (NEB), treated with Turbo DNase (Ambion) and Calf Intestinal Phosphotase (NEB), and PAGE purified. Purified RNA was end-labeled using phosphonucleotide kinase (NEB) and ^{32}P-labeled γ-ATP (Perkin-Elmer) and PAGE purified. Purified p65 construct proteins were incubated with 5.0 ng/µl 5'-end-labeled TER construct for 15 minutes at room temperature in the presence of 20 mM Tris ph 8.0, 100 mM NaCl, 1 mM $MgCl_2$, 0.1 mg/ml yeast tRNA, 1 mM DTT, and 10% glycerol. RNase ONE (Promega) was added at a final concentration of 0.00067 U/µl and incubated at room

temperature. After 15 minutes a recovery control was added and the reaction was immediately quenched by phenol:chloroform extraction and ethanol precipitation. RNase T1 (Ambion) digestion was used to generate a reference ladder. The precipitated RNA was fractionated on a 10% 19:1 acrylamide:bis-acrylamide denaturing sequencing PAGE gel. The gel was imaged using an Amersham Phosphor Screen and a GE Typhoon Scanner.

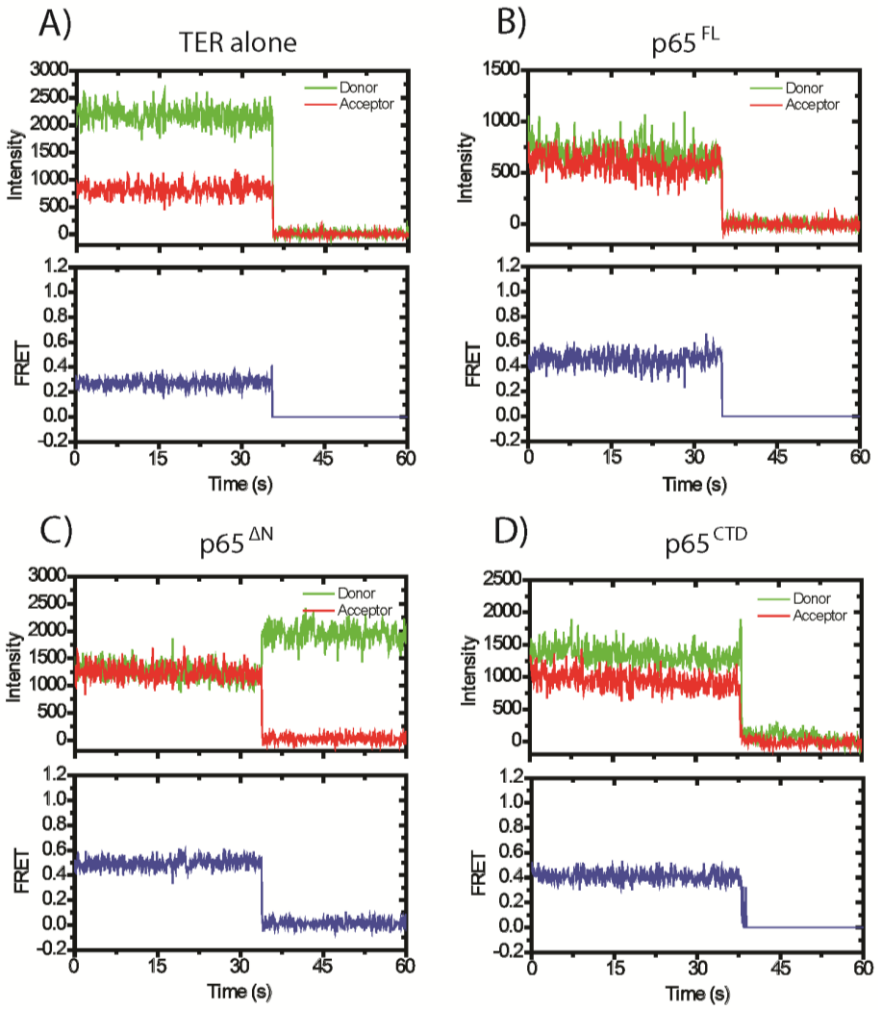
RNase ONE Footprinting Quantification:

RNase footprinting gels were quantified with SAFA [25]. Nucleotide bands were assigned using a T1 reference ladder and SAFA's user interface. SAFA fit each lane's intensity profile as the sum of a series of Lorentzian distributions and quantified the relative intensity of each band as the area under each distribution. Band intensities from each lane were then normalized to the amount of material observed in the recovery control. Normalized band intensities from each quantified band from the low protein (LP) control lane were compared against the high protein (HP) lane. For protected residues, the extent of protection at each band was quantified as 1 - (HP/LP). For deprotected residues, the extent of hypersensitivity was defined as (HP/LP) - 1. This factor was used to scale the sizes of the shapes in Figure 5. Lane intensity profiles for full-length TER (Supplemental Fig. S6) were generated in Imagequant and plotted in Origin.

Supplementary Table 1. Oligonucleotides used to construct dye-labeled TER

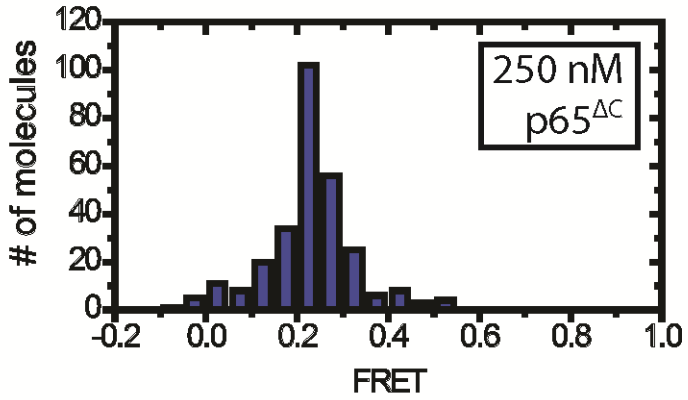
5’ TER Fragment	5’- CCACCAAGCCGACCGAUACCCGCU (5-N-U) AAUUCAUUCAGAUCUGUAAUA - 3’
3’ TER Fragment	5’ P - ACAUUUGAUACACUAUU (5-N-U) AUCAAUGGAUGUCUUAUUUU - 3’
Cutting Oligo Stem II (Underlined residues are 2’-OMe)	5’ - TGACAGTTCTATTACAGATCTG - 3’
Cutting Oligo Stem IV (Underlined residues are 2’-OMe)	5’ - ATCAAATGTCGATAGTCTTTTG - 3’
Stem II DNA splint	5’- GTTGAATGACAGTTCTATTACAGATCTGAA - 3’
Stem IV DNA splint	5’- TAGTGTATCAAATGTCGATAGTCTTTTGTC - 3’
Biotinylated DNA handle	5’ - CGGTCGGCTTGGTGGATAT-biotin -3’

Supplemental Fig. S1



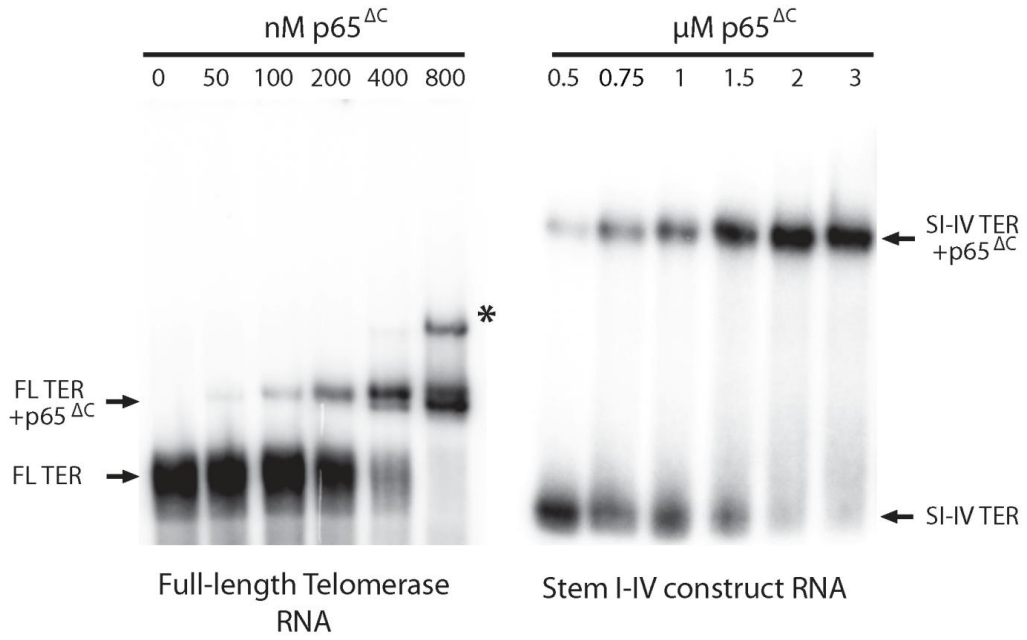
Supplemental Fig. S1. Representative single molecule FRET trajectories. smFRET was recorded on individual molecules of TER labeled at U139 with Cy3 and at U10 with Cy5. The intensities of the donor dye (green trace) and the acceptor dye (red trace) were measured over time and used to calculate FRET (blue trace). These are representative traces of TER incubated with no protein (A), $p65^{FL}$ (B), $p65^{\Delta N}$ (C), and $p65^{CTD}$ (D). $p65^{\Delta C}$ traces resembled TER alone and are not pictured. The abrupt loss of signal at the end of the trace corresponds to dye photobleaching.

Supplemental Fig. S2



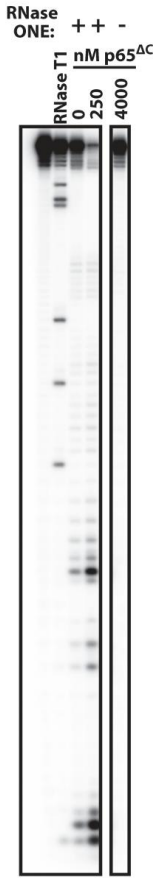
Supplemental Fig. S2. **Increasing concentrations of p65$^{\Delta C}$ fail to induce a structural change in TER.** Dye-labeled TER was incubated with 250 nM p65$^{\Delta C}$ and smFRET was measured. Histogram indicates the number of molecules at the given FRET values. The FRET distribution is centered at 0.23, similar to the 0.26 FRET distribution observed for unbound TER. Thus even high concentrations of p65$^{\Delta C}$ sufficient to form multiple interactions with TER (see Fig. 1C) do not induce the conformational re-arrangement associated with p65 function in the absence of the C-terminal domain.

Supplemental Fig. S3



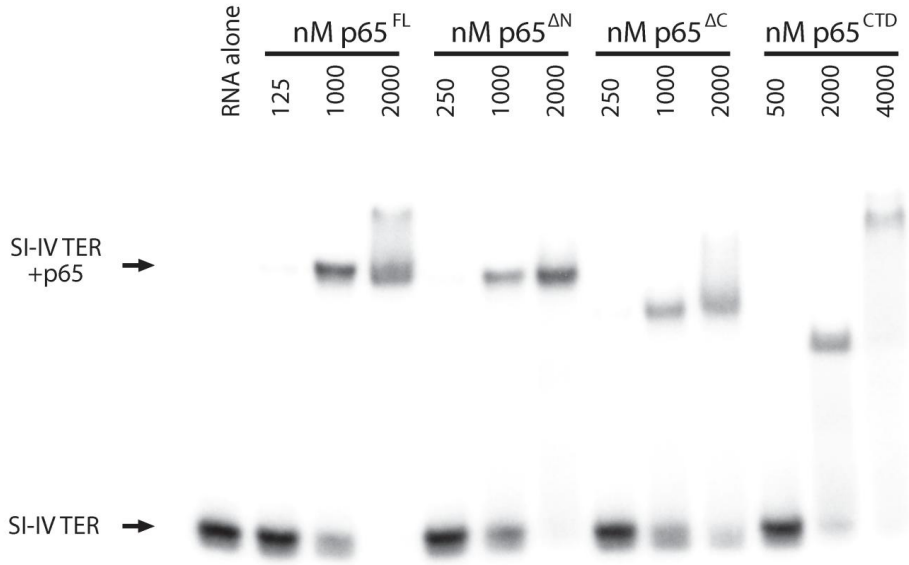
Supplemental Fig. S3. Stem I-IV construct RNA permits higher protein concentrations for RNase footprinting analysis. 5 ng/μl of radiolabeled full-length telomerase RNA (left) or stem I-IV construct RNA (right, see Figure 4A for sequence) were incubated in increasing concentrations of $p65^{\Delta C}$ and separated on a native polyacrylamide gel. Full-length telomerase RNA formed multimers at 800 nM $p65^{\Delta C}$ (indicated by the asterisk), consistent with multiple proteins binding a single RNA. These multimers could confuse the RNase footprinting results. Stem I-IV construct RNA forms no multimers at $p65^{\Delta C}$ concentrations up to 3 μM. EMSAs from full-length TER and Stem I-IV construct RNA were obtained on separate gels and do not represent an accurate reflection of the relative positions of the two constructs run on the same gel.

Supplemental Fig. S4



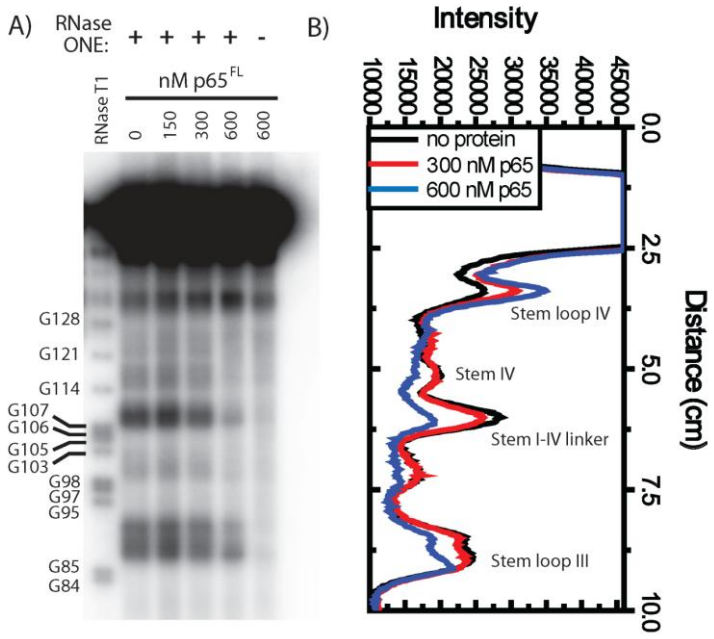
Supplemental Fig. S4. Addition of sub-stoichiometric quantities of p65 constructs stimulate RNase ONE cleavage efficiency. 5 ng/µl of end-labeled stem I-IV construct RNA was incubated with no protein or 250 nM $p65^{\Delta C}$ and cleaved with identical concentrations of RNase ONE. The RNA incubated with 250 nM $p65^{\Delta C}$ showed much higher cleavage efficiency, despite the fact that this $p65^{\Delta C}$ concentration does not support an appreciable amount of protein-RNA complex formation (see Supplemental Fig.S5). This effect is not due to a contaminating nuclease, as incubation with as high as 4 µM $p65^{\Delta C}$ in the absence of RNase demonstrates no cleavage of the RNA. It is unclear what the basis of this phenomenon is, we hypothesize it may be a crowding effect or could be due to transient p65-TER interactions somehow stimulating RNase ONE activity. No additional stimulation of RNase ONE cleavage efficiency was observed by increasing p65 concentration above 250 nM (see Figure 4), thus for RNase ONE footprinting experiments single-hit conditions were determined in the presence of 125-500 nM of the desired p65 construct. This "low protein" condition was used as the baseline for RNase ONE protection experiments in lieu of a no protein condition.

Supplemental Fig. S5



Supplemental Fig. S5. Gel shift analysis of p65 construct protein concentrations used in RNase ONE protection experiments. 5 ng/µl stem I-IV RNA was incubated with $p65^{FL}$, $p65^{\Delta N}$, $p65^{\Delta C}$, and $p65^{CTD}$ at the indicated concentrations. The native gels reveal that at the lowest protein concentrations used in the RNase ONE protections experiments, little or no protein was bound to stem I-IV RNA compared to the higher protein concentrations which demonstrate an appreciable amount of complex formation.

Supplemental Fig. S6



Supplemental Fig. S6. RNase ONE digestion of full-length TER in the presence of p65^{FL} recapitulates results obtained with stem I-IV RNA. A) Full-length telomerase RNA was incubated with increasing concentrations of p65^{FL} and digested with RNase ONE. An RNase T1 ladder indicates the positions of various bands in the gel. **B)** The radioactive intensity profile was quantified in lanes representing RNase ONE digestion in the absence of p65^{FL} (black line), in the presence of 300 nM p65^{FL} (red line), or in the presence of 600 nM p65^{FL} (blue line). The positions of various structural features in the RNA are labeled. The profile demonstrates that p65^{FL} protects residues in stem IV and in the linker between stem I and stem IV. Interestingly, some protection is also observed in stem-loop III residues (the unpaired residues in the telomerase RNA pseudoknot). Furthermore, hypersensitivity to cleavage in the presence of p65^{FL} is observed in stem-loop IV. These results are in agreement with the results obtained with the stem I-IV RNA construct.

References

1. Palm, W. & de Lange, T. How shelterin protects mammalian telomeres. *Annu Rev Genet* **42**, 301-34 (2008).

2. Gilson, E. & Geli, V. How telomeres are replicated. *Nat Rev Mol Cell Biol* **8**, 825-38 (2007).

3. Vulliamy, T.J. & Dokal, I. Dyskeratosis congenita: the diverse clinical presentation of mutations in the telomerase complex. *Biochimie* **90**, 122-30 (2008).

4. Kim, N.W. et al. Specific association of human telomerase activity with immortal cells and cancer. *Science* **266**, 2011-5 (1994).

5. 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).

6. Lin, J. et al. A universal telomerase RNA core structure includes structured motifs required for binding the telomerase reverse transcriptase protein. *Proc Natl Acad Sci U S A* **101**, 14713-8 (2004).

7. Blackburn, E.H. & Collins, K. Telomerase: an RNP enzyme synthesizes DNA. *Cold Spring Harb Perspect Biol* **3**(2010).

8. Mitchell, J.R., Wood, E. & Collins, K. A telomerase component is defective in the human disease dyskeratosis congenita. *Nature* **402**, 551-5 (1999).

9. Seto, A.G., Zaug, A.J., Sobel, S.G., Wolin, S.L. & Cech, T.R. Saccharomyces cerevisiae telomerase is an Sm small nuclear ribonucleoprotein particle. *Nature* **401**, 177-80 (1999).

10. Aigner, S., Postberg, J., Lipps, H.J. & Cech, T.R. The Euplotes La motif protein p43 has properties of a telomerase-specific subunit. *Biochemistry* **42**, 5736-47 (2003).

11. Witkin, K.L. & Collins, K. Holoenzyme proteins required for the physiological assembly and activity of telomerase. *Genes Dev* **18**, 1107-18 (2004).

12. Prathapam, R., Witkin, K.L., O'Connor, C.M. & Collins, K. A telomerase holoenzyme protein enhances telomerase RNA assembly with telomerase reverse transcriptase. *Nat Struct Mol Biol* **12**, 252-7 (2005).

13. O'Connor, C.M. & Collins, K. A novel RNA binding domain in tetrahymena telomerase p65 initiates hierarchical assembly of telomerase holoenzyme. *Mol Cell Biol* **26**, 2029-36 (2006).

14. Berman, A.J., Gooding, A.R. & Cech, T.R. Tetrahymena telomerase protein p65 induces conformational changes throughout telomerase RNA (TER) and rescues telomerase reverse transcriptase and TER assembly mutants. *Mol Cell Biol* **30**, 4965-76 (2010).

15. Stone, M.D. et al. Stepwise protein-mediated RNA folding directs assembly of telomerase ribonucleoprotein. *Nature* **446**, 458-61 (2007).

16. Richards, R.J. et al. Structural study of elements of Tetrahymena telomerase RNA stem-loop IV domain important for function. *RNA* **12**, 1475-85 (2006).

17. Lai, C.K., Miller, M.C. & Collins, K. Roles for RNA in telomerase nucleotide and repeat addition processivity. *Mol Cell* **11**, 1673-83 (2003).

18. Bayfield, M.A., Yang, R. & Maraia, R.J. Conserved and divergent features of the structure and function of La and La-related proteins (LARPs). *Biochim Biophys Acta* **1799**, 365-78 (2010).

19. Axelrod, D., Thompson, N.L. & Burghardt, T.P. Total internal inflection fluorescent microscopy. *J Microsc* **129**, 19-28 (1983).

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

21. Ye, A.J. & Romero, D.P. Phylogenetic relationships amongst tetrahymenine ciliates inferred by a comparison of telomerase RNAs. *Int J Syst Evol Microbiol* **52**, 2297-302 (2002).

22. Chen, Y. et al. Structure of stem-loop IV of Tetrahymena telomerase RNA. *EMBO J* **25**, 3156-66 (2006).

23. Robart, A.R., O'Connor, C.M. & Collins, K. Ciliate telomerase RNA loop IV nucleotides promote hierarchical RNP assembly and holoenzyme stability. *RNA* **16**, 563-71 (2010).

24. O'Connor, C.M., Lai, C.K. & Collins, K. Two purified domains of telomerase reverse transcriptase reconstitute sequence-specific interactions with RNA. *J Biol Chem* **280**, 17533-9 (2005).

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**, 344-54 (2005).

CHAPTER III: A conserved motif in *Tetrahymena thermophila* telomerase reverse transcriptase defines the template boundary by promoting telomerase RNA binding

(originally published in the *Journal of Biological Chemistry*)

Abstract

The ends of linear chromosomes are extended by telomerase, a ribonucleoprotein complex minimally consisting of a protein subunit called telomerase reverse transcriptase (TERT) and the telomerase RNA (TER). TERT functions by reverse transcribing a short template region of TER into telomeric DNA. Proper assembly of TERT and TER is essential for telomerase activity; however a detailed understanding of how TERT interacts with TER is lacking. Previous studies have identified an RNA binding domain (RBD) within TERT, which includes three evolutionarily-conserved sequence motifs: CP2, CP, and T. Here, we used site-directed hydroxyl radical probing to directly identify sites of interaction between the TERT RBD and TER, revealing that the CP2 motif is in close proximity to a conserved region of TER known as the template boundary element (TBE). Gel shift assays on CP2 mutants confirmed that the CP2 motif is an RNA binding determinant. Our results explain previous work which established that mutations to the CP2 motif of TERT and to the TBE of TER both permit mis-incorporation of nucleotides into the growing DNA strand beyond the canonical template. Taken together, these results

suggest a model in which the CP2 motif binds the TBE to strictly define which TER nucleotides can be reverse transcribed.

Introduction

Eukaryotic cells must distinguish the natural ends of chromosomes from double-stranded DNA breaks that are a mark of DNA damage. In order to do so, chromosome ends are capped by chromatin structures called telomeres, comprised of repetitive DNA sequences that recruit specific DNA-binding proteins [1]. Telomeres are shortened with each round of cell division, and so require the enzyme telomerase to maintain telomere length in actively dividing cells. Telomerase dysfunction is associated with diseases that affect proliferative tissues, such as dyskeratosis congenita and aplastic anemia [2]. On the other hand, aberrant telomerase over-expression can help to confer proliferative potential to cells, and is associated with 90% of human cancer cell lines, making telomerase an attractive target for cancer therapies [3].

Telomerase is a protein-RNA complex which includes both the protein telomerase reverse transcriptase (TERT) and telomerase RNA (TER). TER contains a region complementary to the telomeric DNA sequence, known as the template. The template region of TER is repetitively reverse transcribed by TERT to extend telomere DNA [4].

TERT contains four domains, an N-terminal domain (TEN), an RNA binding domain (RBD), a reverse transcriptase domain (RT), and a C-terminal extension

(CTE) [5]. TERs also have several conserved structural motifs required for the function of the telomerase holoenzyme. TERs from ciliates, yeasts, and mammals all contain an essential RNA pseudoknot proximal to the template [6]. In addition, TERs universally possess a template boundary element (TBE) that defines the region of TER to be reverse transcribed by TERT [7,8]. The high fidelity of template definition established by RBD-TBE interactions is a central feature of the telomerase catalytic cycle, since incorporation of a single non-template nucleotide will prevent synthesis of subsequent telomere DNA repeats.

Recent studies in the *Tetrahymena thermophila* model system have shown that the RBD domain of TERT is primarily responsible for interactions with TER, is sufficient to bind TER with high affinity, and includes several conserved protein motifs involved in RNA binding [9-13]. However, it is still not known which portions of the RBD interact with which conserved RNA motifs. An X-ray crystal structure of a large C-terminal fragment of the *T. thermophila* RBD has been solved [14], and comparison with the *Tribolium castaneum* TERT structure suggests that the conserved T-motif forms a beta-strand hairpin near the position of the template RNA [15]. Furthermore, the conserved CP motif was shown to be adjacent to the T-motif in an electropositive groove, perhaps positioned to bind the TBE [14]. However, a co-crystal with the RBD bound to RNA was not attained, and so it could not be definitively shown which protein motifs interact with which RNA domains. A third conserved domain, the CP2 motif, was not included in the RBD construct that was solved by X-ray crystallography.

In the present work, we employed site-directed Fe(II)-EDTA hydroxyl radical probing to map protein-RNA interactions in telomerase, as done previously for the ribosome [16,17]. Our results demonstrated the CP2 motif and the TBE of *Tetrahymena* TER are in close proximity. To further explore to role of the CP2 motif in TER binding, we generated CP2 point mutants and measured their affinity for TER. Electrophoretic mobility shift assays (EMSAs) indicated that deletion of the entire CP2 motif severely compromised TER binding, and many single point mutations in the CP2 motif result in a reduced affinity for TER, consistent with the CP2 motif being an important determinant of RBD binding. Among the single amino acid mutants analyzed, the strongest binding defect was observed with a mutation to R237, a residue previously shown to play a role in telomerase activity and template definition [8,10]. Quantitative EMSAs demonstrated that a single amino acid CP2 point mutant (R237A) showed an approximate 7-fold reduction in TER affinity compared to WT RBD.

Finally, we verified that telomerase harboring an R237A mutation is severely knocked-down in telomerase activity. Interestingly, the assembly protein p65 can partially rescue telomerase activity in R237A TERT, however p65 does not rescue the previously-reported R237A template boundary defect. Taken together, our results demonstrate that the CP2 motif is essential for a functional interaction between TERT and TER, and is a critical protein determinant of template definition.

Results

Tethered hydroxyl radical probing demonstrates the CP2 motif is proximal to the TBE

We initially characterized the structural basis of RBD-TER interactions using site-directed hydroxyl radical probing. In this technique, a single reactive cysteine in the protein of interest is labeled with an Fe(II)-EDTA moiety. The functionalized protein is incubated with RNA in a buffer containing ascorbate and H_2O_2, generating hydroxyl radicals by the Fenton reaction [18]. These hydroxyl radicals cleave any RNA nearby, but can only travel ~20Å before quenching. Sites of RNA cleavage can be identified by polyacrylamide gel electrophoresis, giving a readout of regions of RNA-protein interaction. To facilitate site-specific Fe(II)-EDTA tethering, six endogenous cysteines were mutated out of a plasmid coding for the TERT RBD, generating a Cys-lite RBD construct (Fig. 1A). Single cysteines were then introduced back into the protein, generating a series of constructs containing only a single cysteine at specific regions of interest in the TERT RBD. The proteins were expressed in *E. coli*, affinity purified, and labeled with Fe(III)-1-(p-bromoacetamidobenzyl)-EDTA (Fe-BABE). Fe-BABE-labeled proteins were incubated with ^{32}P-labeled TER (Fig. 1B), and ascorbate and hydrogen peroxide were added to the protein-RNA complexes to initiate the formation of hydroxyl radicals.

The majority of amino acid labeling sites showed no significant difference between Fe-BABE-labeled protein and unlabeled protein (Supplemental Table 1). However, RBD labeled at residue Cys232 showed a considerable increase in cleavage

in TER nucleotides 17-20 and 37-41 (Fig. 1C). Titrating increasing concentrations of 232C-Fe-BABE RBD resulted in a corresponding increase in the intensity of the cleavage products (Fig. 1C, lanes 3-7), whereas incubation of the RNA with mock-labeled cys-lite RBD, unlabeled 232C RBD, or 232C-Fe-BABE RBD in the absence of hydrogen peroxide and ascorbate resulted in no cleavage at this site (Fig. 1C, lanes 8-10, respectively). A chromatogram of the lane intensity profiles of a lane with 232C-FeBABE (lane 7) and a mock-labeled control (lane 8) clearly demonstrates the extent of the increase in cleavage (Fig. 1C, right).

To demonstrate the reproducibility of the result, the experiment was repeated three times and the intensities of the cleavage products were quantified using the gel quantification program SAFA [19]. The intensity of the cleavage product was measured at each individual residue and compared between RNA incubated with 1000 nM 232C-FeBABE cys-lite RBD and RNA incubated with 1000 nM mock-labeled cys-lite RBD (Fig. 1D). The experiment demonstrates a highly-reproducible cleavage pattern, with the most hydroxyl radical-induced cleavage observed at the base of stem II, most notably at residue 17. Plotting the mean cleavage intensity against the secondary structure of the RNA demonstrates the extent to which the pattern of cleavage mirrors the RNA secondary structure, peaking at the base of stem II and diminishing along the basepaired residues (Fig. 1D, inset).

The observed sites of cleavage at the base of TER stem II are part of a previously-described conserved RNA motif known as the template boundary element (TBE). Mutations to the RNA in this region show errors in template definition,

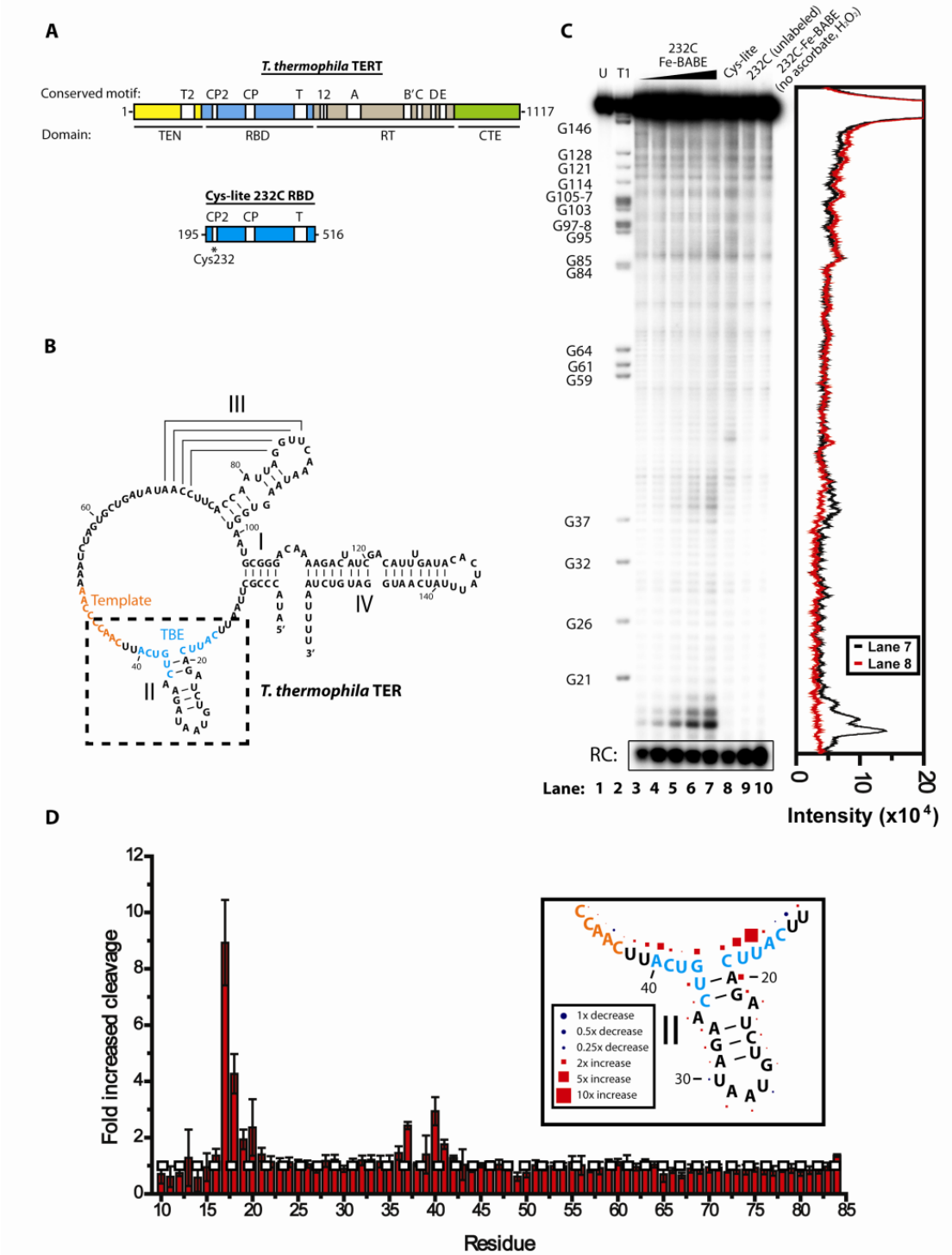


FIGURE 1. Hydroxyl radical probing of RBD-TER complexes. (Caption on next page)

FIGURE 1. Hydroxyl radical probing of RBD-TER complexes. *(A)* Primary structure map of *T. thermophila* TERT. For hydroxyl radical probing studies, an RBD construct was expressed with all endogenous cysteines mutated to alanine or serine, and a single cysteine re-introduced at residue 232 in the CP2 motif. *(B)* Secondary structure of *T. thermophila* TER. The template RNA is highlighted in orange and the template boundary element (TBE) is highlighted in blue. Box indicates region of interest highlighted in Fig. 1D. *(C)* Site-directed hydroxyl radical probing results for Cys-lite RBD 232C. Lane 1 – TER alone. Lane 2 – RNase T1 digestion ladder of TER indentifying position of guanine residues. Lanes 3-7 – Hydroxyl radical probing performed in the presence of 62.5 nM, 125 nM, 250 nM, 500 nM, and 1000 nM of 232C-Fe-BABE Cys-lite RBD, respectively. Lane 8 – 1000 nM Cys-lite RBD was mock-labeled with Fe-BABE and hydroxyl radical probing was performed. Lane 9 – 1000 nM unlabeled 232C TERT RBD hydroxyl radical probing. Lane 10 – 1000 nM Fe-BABE-labeled Cys-lite RBD 232C was incubated with RNA without adding ascorbate and hydrogen peroxide. The chromatogram on the right demonstrates the intensity of cleavage products, comparing lane 7 (black) and lane 8 (red). *(D)* Hydroxyl radical probing was performed using 1000 nM 232C-Fe-BABE Cys-lite RBD and 1000 nM mock-labeled Cys-lite RBD. The experiment was repeated in triplicate and the fold increase in cleavage was calculated at each residue using the gel quantifying program SAFA [19]. Values represent the mean ± s.d., with values above 1 indicating increased cleavage due to Fe-BABE-induced hydroxyl radical probing at residue 232 and values below 1 indicating decreased cleavage. The white horizontal dashed line indicates no change in cleavage compared to the mock-labeled Cys-lite RBD construct. The inset represents the mean values in Fig. 1D plotted against the secondary structure, with the size of the shapes scaled to represent the degree of cleavage or protection.

allowing the non-template residue U42 to be aberrantly reverse transcribed by TERT in the presence of dATP [8,12]. Furthermore, this region was shown to be required for high affinity TERT interaction, leading to the model that the TBE is bound tightly by the RBD, and that this interaction prevents the TERT RT domain from reverse transcribing non-template TER nucleotides [8,12].

The site of protein labeling, Cys232, is part of an evolutionarily-conserved motif, known as the CP2 motif. The CP2 motif is found only in ciliate TERT and

previous mutational studies demonstrated that mutations to many of the residues in the CP2 motif have a significant defect in telomerase activity [10]. Interestingly, one CP2 mutant showed the same defect in template definition as observed with a TBE mutant, a result that is consistent with the observed physical proximity of these two domains in our hydroxyl radical probing experiments.

EMSAs demonstrate the CP2 motif contributes to TER binding

While our hydroxyl radical probing results indicate that the CP2 motif is physically proximal to the TBE in the protein-RNA complex, they do not directly demonstrate that CP2 plays an essential role in RNA-binding. To further explore the role of the CP2 motif in mediating protein-RNA interactions, we first expressed an RBD construct that lacks the CP2 motif entirely (ΔCP2) (Fig. 2A) and compared its RNA binding activity to the full-length RBD protein by EMSA. The ΔCP2 RBD construct was readily expressed and purified (Fig. 2B) but demonstrated considerably diminished RNA binding activity compared to WT protein (Fig. 2C). The high ΔCP2 RBD protein concentrations required to bind TER in our EMSA experiments, together with the propensity to form aggregates at protein concentrations above 320 nM, precluded our ability to accurately determine a dissociation constant. However, these results strongly suggest the evolutionarily conserved CP2 motif is necessary for efficient and specific binding to TER.

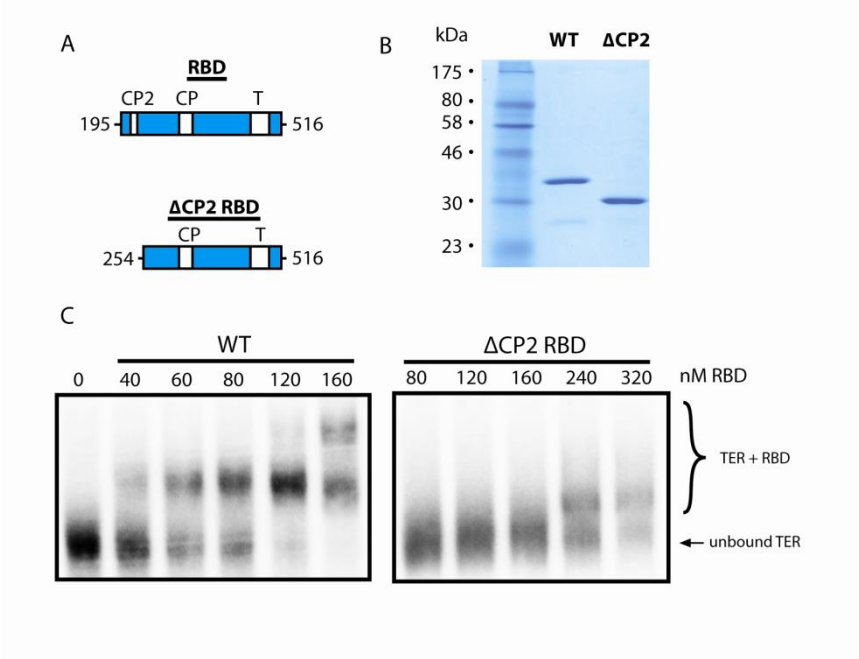


FIGURE 2. An RBD construct lacking the CP2 motif has a highly reduced affinity for TER. *(A)* Primary structure map of *T. thermophila* TERT wild-type and ΔCP2 RBD. *(B)* SDS-PAGE gel of purified wild-type RBD and the ΔCP2 RBD mutant. *(C)* EMSAs were performed using full-length TER and increasing concentrations of wither wild-type RBD (left) or an RBD construct lacking the CP2 motif (ΔCP2, right). The lower band represents unbound TER and the upper bands represent TER-RBD complexes.

The CP2 motif is a 12 amino acid sequence conserved across various ciliate species, containing many hydrophobic and positively charged residues potentially involved in RNA binding (Fig. 3A). To further dissect the contribution of the CP2 motif to TER binding we next performed an alanine scan through the CP2 region of TERT. In order to simplify the purification procedure for a semi-quantitative initial screen, we only purified the proteins with one round of nickel-affinity chromatography. Wild-type RBD and proteins encoding mutants Y231A, C232A, H234A, and R237A showed robust expression, and purified to high levels of purity as

determined by SDS-PAGE (Fig. 3B, left panel). Mutations R226A, I229A, and F230A showed a modest decrease in protein expression (Fig. 3B, right panel) and consequently co-purified with increased levels of contaminating protein, confounding accurate determination of protein concentrations by Bradford assays. Instead, the protein concentrations of these constructs were estimated by comparing Coomassie blue staining of only the RBD band on SDS-PAGE gels (Fig. 3B, arrow).

This initial screen with partially-purified protein indicated an approximate dissociation constant, or the concentration of protein where 50% of the RNA is bound, for the WT protein of ~60 nM (Fig. 3C). The CP2 point mutations F230A and Y231A appeared to have no effect on RNA binding, as gels for these constructs were virtually identical to the WT protein. Mutants I229A, C232A, and H234A appeared to have a modest defect in RNA binding, as evidenced by the slightly increased concentrations of protein required to bind the RNA (Fig. 3C). However, the largest and most-obvious defect was observed with the mutants of the flanking arginines (R226A and R237A). In addition to the reduced affinity, the higher molecular weight complexes observed with the R237A mutant appeared as diffuse smear on the gel, likely due to partial RBD binding or weakened complex formation, indicative of perhaps a further defect (Fig. 3C).

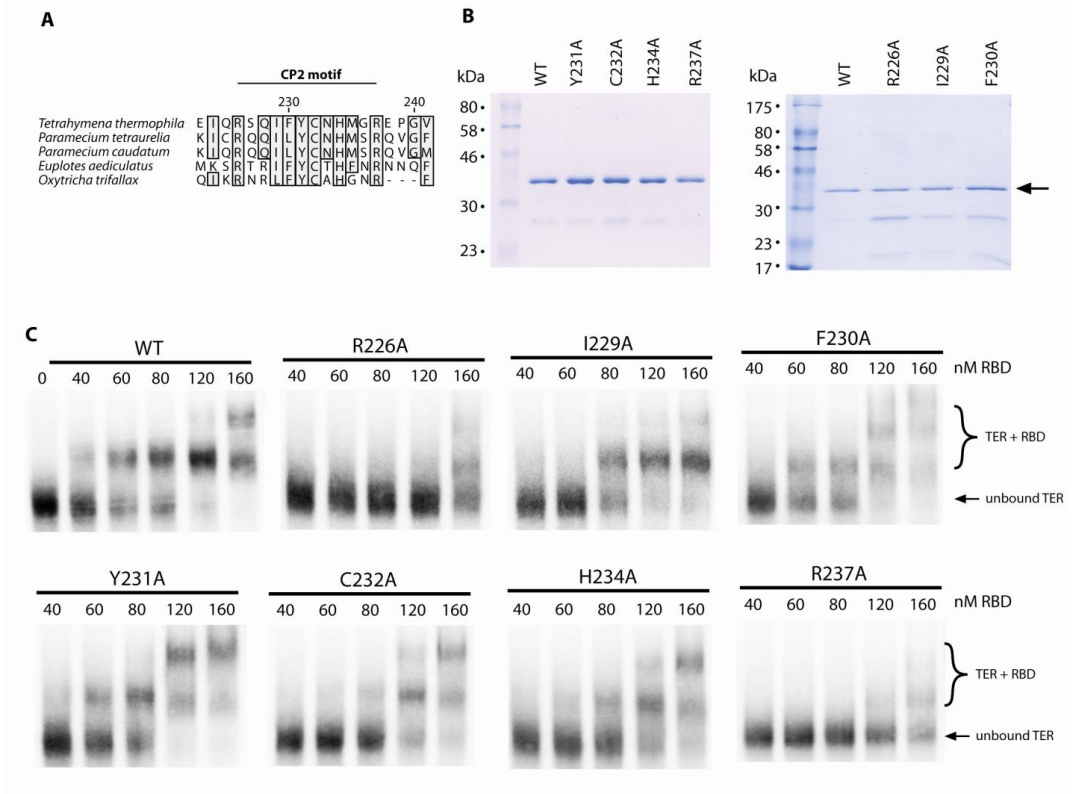


FIGURE 3. CP2 motif mutants demonstrate weaker TER affinity by EMSA. *(A)* Sequence conservation map of the CP2 motif, comparing TERT sequences of five ciliate species. Numbering is for the residue numbers in the *T. thermophila* TERT sequence. *(B, left panel)* SDS-PAGE gel of WT RBD as well as equal quantities of point mutants Y231A, C232A, H234A, and R237A. *(right panel)* SDS-PAGE gel of purified WT RBD as well as R226A, I229A, and F230A protein. These constructs showed greater contamination, therefore the protein concentration was estimated by comparing the Coomassie blue intensity at the RBD band (arrow) against the known WT RBD concentration. *(C)* EMSAs for WT RBD (as shown in Fig. 2C) as well as all mutagenesis constructs. The lower band represents unbound TER and the upper bands represent TER-RBD complexes.

In order to further interrogate the contribution of the CP2 motif to RNA binding, we followed our initial screen with quantitative EMSAs using fully-purified protein constructs. For these measurements we focused on a comparison of WT RBD to one point mutant that showed a significant defect in our screen (R237A). These

protein constructs were purified using nickel affinity chromatography, ion exchange chromatography, and sizing chromatography before use in EMSAs.

An SDS-PAGE gel indicated that both WT and R237A proteins were purified to homogeneity (Fig. 4A). In addition, analytical sizing chromatography demonstrated that both proteins eluted at the same elution volume, consistent with a monomeric protein of ~40 kDa (Fig. 4B), demonstrating that the mutation did not predispose the R237A mutant to aggregation. The quantitative EMSAs on these protein constructs confirmed the results of the initial screen. Under the conditions of our assay, the highly-purified WT RBD demonstrated a K_d of 24 ± 5 nM, a value that is slightly better than observed in our initial EMSA screens. The binding defect observed for the R237A mutant persisted, exhibiting a considerable decrease to 183 ± 26 nM when compared to WT RBD (Fig. 4C, D). The effect of the mutation is not due to a general folding defect, as both proteins showed similar specific activities in stoichiometric binding assays (see Experimental Procedures).

Under higher protein concentrations, EMSAs tended to reveal higher molecular weight complexes, likely corresponding to multiple RBD constructs binding to a single RNA (Fig. 4C, asterisks). A recent electron microscopy structure of *Tetrahymena* telomerase confirms that *T. thermophila* TERT and TER bind in a 1:1 ratio [20]. Therefore, the higher order RBD-TER complexes we observe are likely artifacts of using purified protein constructs at high relative protein concentrations, and do not reflect a physiologically relevant complex. These higher molecular weight complexes were nevertheless included in the quantification as it is likely that they

represent naturally-bound complexes augmented with an additional protein interaction. While this may subtly shift the equilibrium between bound and unbound forms, this analysis is sufficient to make a meaningful comparison between the affinities of the WT and R237A protein constructs.

Previous work has suggested multiple sites of interaction between TERT and TER [12,21]. This is likely a large source of the residual binding activity in the ΔCP2

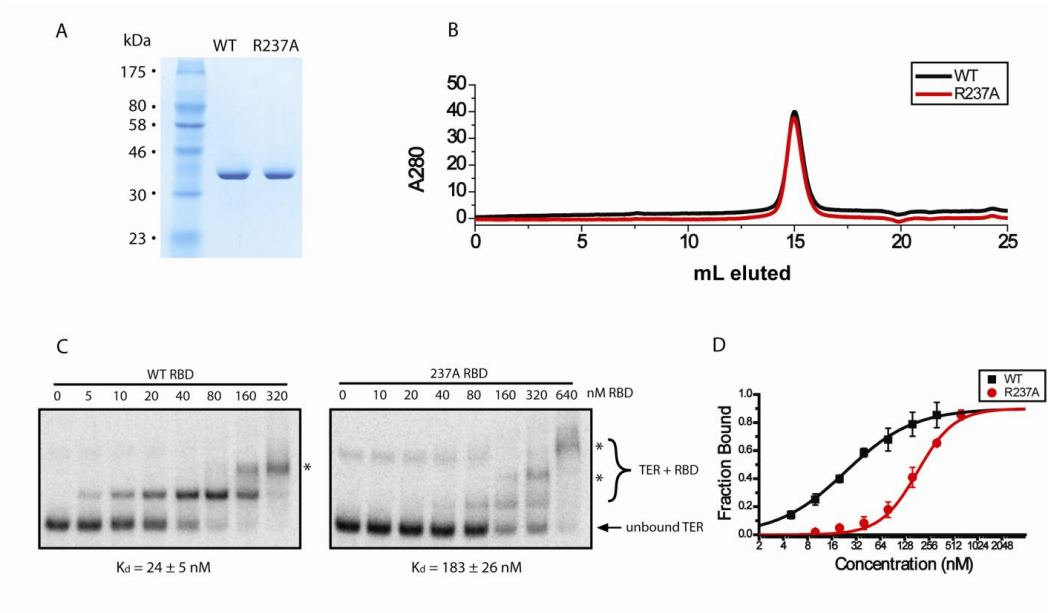


FIGURE 4. Quantitative EMSAs on highly-purified protein demonstrate the CP2 motif is critical for RNA binding. *(A)* SDS-PAGE gel of equal quantities of WT and R237A RBD protein. *(B)* Analytical sizing chromatography demonstrates that both proteins elute identically, consistent with a monomeric protein of ~40 kDa. *(C)* EMSAs for WT and R237A RBD. The dissociation constant (K_d) was determined for all protein variants as well as the standard deviation from triplicate measurements, and is indicated below the gel as $K_d \pm$ S.D. *(D)* K_d quantification for WT and R237A RBD. The K_d was determined by determining the fraction bound at each concentration and fitting the data to the Hill equation. The Hill co-efficients obtained were n=1.0, n=1.9 for WT and R237A constructs, respectively. The higher Hill co-efficient for R237A protein is likely the result of the increased contribution of higher molecular weight species (asterisks) skewing the fitting.

mutant. It is likely that different parts of the RBD maintain separate binding sites on TER, increasing the overall affinity through an avidity effect [21-24]. Our results are consistent with the CP2 motif constituting one of these points of RNA interaction and specifically implicate R226 and R237 as being critical residues in mediating TER binding.

Mutations to the CP2 motif reduce telomerase activity and are partially rescued by the assembly protein p65

The R237A mutation that showed the most dramatic defect in RBD affinity has been previously identified as a mutation that causes a template definition defect in full-length telomerase [8,10]. A previous study determined that the *Tetrahymena thermophila* assembly protein p65 can rescue deleterious TERT mutations, presumably by suppressing defects that arise from weakened TERT-TER interactions [25]. We set out to test the effect of p65 on R237A mutant telomerase activity. Full-length TERT was reconstituted with TER in the absence or presence of purified p65 in rabbit reticulocyte lysate (RRL) and the catalytic activity of the telomerase complexes was assessed using a direct telomere DNA primer extension assay. As expected, wild-type telomerase shows robust extension activity both in the presence and absence of p65 (Fig. 5). However, R237A telomerase displayed a dramatic decrease in telomerase activity, with virtually no telomerase activity observed in the absence of p65 (Fig. 5). We note that a previous study of the R237A TERT in the absence of p65 reported detectable telomerase activity, albeit at a significantly

reduced level compared to wild type TERT [8]. We expect the inability of our present experiments to detect activity with the R237A mutant in the absence of p65 likely reflects small differences in the protocols used to reconstitute and purify the telomerase complex. Nevertheless, our results are qualitatively consistent with the previous report in that the R237A mutation results in a marked reduction in the efficiency of telomerase reconstitution.

The presence of p65 improved R237A telomerase activity, though there was still considerably less extension than wild-type telomerase. Previous studies have shown that the R237A mutation inhibits TERT-TER assembly [8]. Since p65 is an assembly co-factor, we conclude that p65 likely rescues telomerase activity in the CP2 mutant by stabilizing the interaction between R237A TERT and TER.

We also tested R237A telomerase for template boundary defects. In the presence of a template boundary defect, residue U42 of TER aberrantly enters the TERT active site. This results in dATP incorporation into the telomeric DNA primer, which is observed as an additional band one nucleotide above the canonical repeat addition band. In our assay, the repeat addition band is at the +1 site, therefore template boundary defects can be observed by an increase in intensity of the +2 band when telomerase activity reactions are supplemented with dATP. Wild-type telomerase showed no increase at the +2 position in the presence of dATP, indicative of robust wild-type template definition (Fig. 5). Due to the undetectable level of catalytic activity in the absence of p65, R237A mutants could not be assessed for template boundary defects under these conditions. However in the presence of p65,

there was a considerable increase in template read-through when dATP was present (Fig. 5, asterisks). Taken together, our results indicate that the CP2 motif is important for TERT-TER assembly and telomerase activity, and that the assembly protein p65 can partially rescue telomerase activity in a CP2 mutant. Furthermore, the CP2 motif is essential for proper template definition as reported previously [8,10], and template definition defects persist in a CP2 mutant even in the presence of p65.

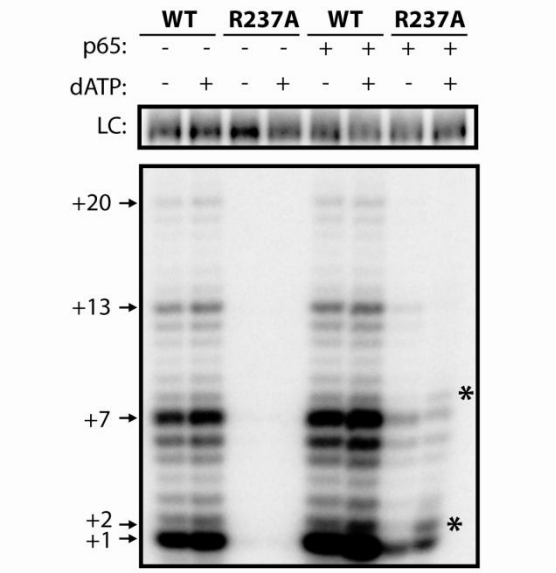


FIGURE 5. A CP2 mutant demonstrates telomerase activity defects. Telomerase activity assays performed from RRL-reconstituted telomerase, using either WT TERT or TERT harboring a CP2 mutation (R237A). Assays were performed in the presence or absence of the assembly protein p65 and in the presence or absence of dATP as indicated. Numbers (+1,+2,+7, etc.) indicate the number of nucleotides incorporated by telomerase at the indicated band. R237A TERT displayed a template boundary defect as indicated by an increase in the +2 product upon addition of dATP (asterisks). L.C. = loading control.

Discussion

The CP2 motif binds proximal to the TBE and contributes to template definition

Our site-directed hydroxyl radical probing results reveal that the CP2 motif is adjacent to the TBE in the TERT-TER complex and gel shift assays verify that the CP2 motif is essential for proper RNA binding. Furthermore, the residues most strongly implicated in CP2-TER interaction (R226 and R237) are basic residues that may contact the RNA backbone through an electrostatic interaction. Finally, telomerase activity assays on a CP2 mutant confirm that the motif is critical for telomerase activity, showing both a severe assembly defect and a template boundary defect. Taken together, our experiments favor a model in which the CP2 motif binds the TBE to establish the template boundary. Nevertheless, we note that it remains a formal possibility that the CP2 motif is proximal to the TBE and influences RNA binding through an indirect interaction mediated by other protein factors.

A model for TBE interaction with the TERT RBD

Comparison of x-ray crystal structures of the *T. castaneum* TERT in complex with its DNA substrate and the *T. thermophila* RBD show significant structural homology [14,15,26]. Alignment of the two structures reveals the likely position of the *T. thermophila* RBD with respect to the reverse transcriptase domains and the RNA template (Fig. 6A). In this structure, the 5' end of the template is positioned adjacent to the T motif. Since the TBE begins three nucleotides upstream of this position, this puts a constraint on the position of the TBE within the structure. Based on this

alignment, the CP motif is optimally positioned to interact with the TBE in an electropositive groove adjacent to the template [14]. The evolutionary conservation of the CP motif, previous biochemical experiments [13], and its position within the RBD structure suggest that the CP motif plays some part in binding the TBE. We therefore propose a structural model based on the solved crystal structures of the *T. thermophila* RBD, the *T. castaneum* full-length TERT, and NMR structures of stem II of *T. thermophila* TER which places the base of stem II adjacent to the CP motif [14,15,27](Fig 6B).

Currently, there is no structural information on the CP2 motif. Nevertheless, we can use our hydroxyl radical probing results to place the CP2 motif adjacent to the TBE in our structural model (Fig. 6B). In this model, the CP and CP2 domains co-operate to bind the base of stem II and maintain the template boundary. This conclusion is strongly supported by our biochemical data as well as telomerase activity assays which previously demonstrated that mutants to both the CP and CP2 domains affect template definition [8].

Comparison of the Tetrahymena CP2 motif to vertebrate TERT

The CP2 motif is an area of conservation found only in ciliates. The lack of conservation of the CP2 motif across species is not entirely unexpected, as it is known that there is also a divergence in the RNA sequences of the TBE [7]. This raises the possibility that this region of the protein may co-evolve with TER TBEs, accounting for the divergence of this region of the protein across phyla. Interestingly,

in an analogous position in vertebrate TERT there is another region of conservation, found only in vertebrates, known as the vertebrate-specific region (VSR) [28]. Our results suggest that the VSR may function in vertebrates in an analogous manner to the CP2 motif in ciliates, promoting TBE interactions and establishing the template boundary.

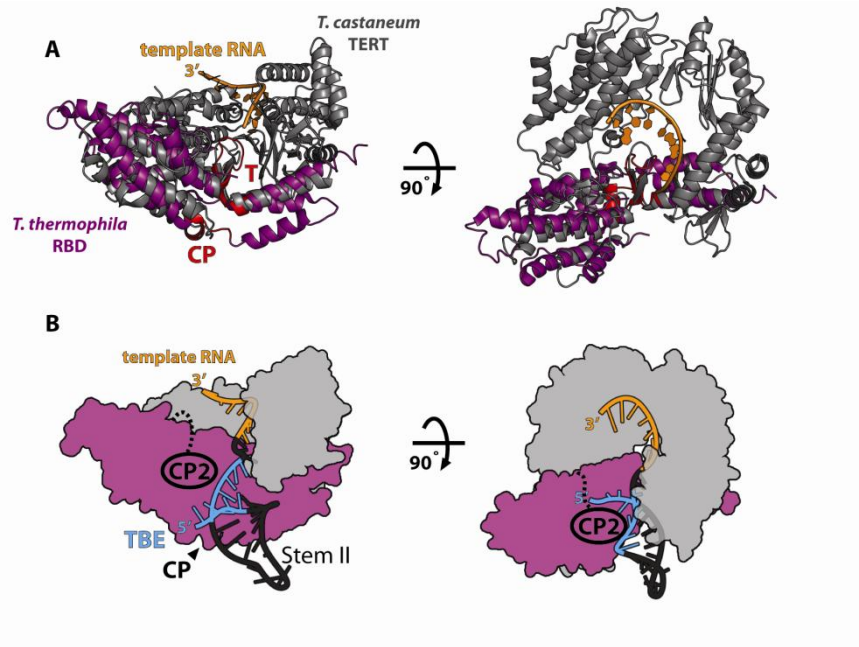


FIGURE 6. Model of TERT-TER interactions. *(A)* Structural alignment of *T. castaneum* FL TERT (grey, PDB ID 3KYL) and *T. thermophila* RBD (purple, PDB ID 2R4G). The position of the conserved T and CP motifs are highlighted in red. The position of the 5' end of the template (TER residue 43) near the T motif places a constraint on the position of the TBE (residues 15-40) likely placing the TBE near the CP motif in the RBD structure. *(B)* Model of telomerase structure. The RT and CTE motifs from the *T. castaneum* structure are modeled in grey and the *T. thermophila* RBD is in purple as in Fig. 5A. Based on the structural alignment in Fig. 5A, the CP motif likely binds the base of stem II RNA, placing stem II (PDB ID 2FRL) within the existing structural model. Our results definitively implicate the CP2 motif in stem II binding, suggesting that the CP2 motif extends towards stem II to cooperatively bind stem II with the CP motif.

EXPERIMENTAL PROCEDURES

PCR mutagenesis

Plasmids containing the *T. thermophila* RBD fused to an N-terminal his-tag were PCR mutagenized using custom DNA primers. Linear plasmid amplicons were treated with T4 polynucledotide kinase (NEB) and T4 DNA ligase to generate circular plasmids, which were used to transform *E. coli* 10β cells. Clones were sequenced to confirm the presence of the desired mutation. The cys-lite RBD had the following mutations: C232A, C300S, C331A, C359A, C387A, C424A.

Protein expression and purification

TERT RBD was expressed in *E. coli* BL21 (DE3) cells induced at 21°C with 0.8 mM IPTG for 4 hours. Cells were harvested by centrifugation and lysed by sonication in buffer containing 20 mM Tris pH 8.0, 500 mM NaCl, 1 mM $MgCl_2$, 1 mM PMSF, 10% glycerol, and 5 mM β-mercaptoethanol (BME). Cell lysate was centrifuged to remove precipitates and cell debris, and supernatant protein was purified by nickel affinity chromatography. Purified protein was eluted into a buffer containing 20 mM Tris pH 8.0, 200 mM NaCl, 1 mM $MgCl_2$, 500 mM imidazole, and 5 mM β-mercaptoethanol and flash frozen with liquid nitrogen for later use.

For quantitative EMSAs, proteins underwent additional rounds of purification. Following nickel affinity chromatography, the eluate was diluted to a buffer containing 20 mM Tris pH 7.0, 50 mM NaCl, 1mM $MgCl_2$, 10% glycerol, and 5 mM BME and purified on an ion exchange source-S column (GE lifesciences) using a

gradient from 50 mM to 1M NaCl. The protein eluted at around 300 mM NaCl and was concentrated in a centrifugal concentrator (Millipore) with a 30 kDa molecular weight cut-off. Protein was flash frozen with liquid nitrogen and later purified on a Sephadex-200 (GE lifesciences) sizing column in 20 mM Tris pH 8.0, 200 mM NaCl, 1 mM $MgCl_2$, 10% glycerol, and 5 mM BME. Protein was concentrated off of the sizing column and flash frozen for later use. All protein constructs eluted as a single monomeric peak off of the sizing column, consistent with a ~40 kDa protein. The percent activity of WT and R237A protein constructs were determined using stoichiometric gel shifts. 2.5 µM cold TER was incubated with 4 nM end-labeled 32P-TER and 2-16 µM RBD constructs in 20 mM Tris pH 8.0, 100 mM NaCl, 1 mM MgCl2, 10% glycerol, and 5 mM BME. WT and R237A RBD demonstrated approximate percent binding activities of ~30% and ~27%, respectively (data not shown).

Fe-BABE labeling

Protein constructs containing only a single cysteine at the desired labeling site were dialyzed into a buffer lacking reducing agent (20 mM Tris pH 8.0, 200 mM NaCl, 1 mM $MgCl_2$, 10% glycerol) overnight, then switched to fresh dialysis buffer and dialyzed for an additional 2 hours. The concentration of Tris in the buffer was then raised to 80 mM, and the dialyzed protein was incubated with a 4-fold molar excess of Fe-BABE for 3.5 hours at room temperature in the dark. Next, the protein was dialyzed overnight against fresh dialysis buffer to remove excess Fe-BABE, and again for an additional 2 hours. Fe-BABE-labeled RBD was quantified by Bradford

assay and flash frozen in liquid nitrogen for later use. Mock-labeled cys-lite RBD was treated identically to labeled protein, however the protein construct lacked any cysteines to interact with the Fe-BABE moiety.

Hydroxyl radical probing

Protein constructs were incubated with 125 ng *in vitro* transcribed TER end-labeled with ^{32}P (Perkin-Elmer). Binding was performed in a buffer containing 20 mM Tris pH 8.0, 100 mM NaCl, and 1 mM $MgCl_2$, 0.1 mg/ml tRNA, 80 units RNasin, and 450 nM p65for 10 minutes at room-temperature in a final volume of 50 µL. 1 µl 250 mM sodium ascorbate and 1 µl 1.25% H_2O_2 were added to the side of each tube and mixed instantaneously by a centrifuge pulse. Reactions were incubated for 10 minutes on ice, then quenched with 10 µl 20 mM thiourea and a radiolabeled DNA recovery control was added. The reactions were then phenol:chloroform extracted and ethanol precipitated. The RNA was run on a 7% sequencing PAGE gel containing 8M urea and the gel was dried and imaged using a phosphor screen (GE lifesciences) and a typhoon scanner (GE lifesciences). The T1 ladder was generated using full-length TER and RNase T1 (Ambion).

Hydroxyl radical probing quantification

Hydroxyl radical probing gels were quantified with SAFA [19] as previously described [22]. The quantified band intensities were compared between 1000 nM Fe-BABE-labeled protein and 1000 nM mock-labeled protein by dividing the intensity of

the band in the labeled protein lane over the intensity of the band in the mock-labeled protein lane.

EMSAs

EMSAs were performed as previously described [22]. For the quantitative EMSAs in Fig. 4, 0.4 nM end-labeled RNA was used instead of body-labeled RNA. Band intensities were quantified using Imagequant and data were plotted and fit in Origin to determine K_d values. Percent bound complexes were plotted in Origin and K_d values were determined by fitting to the binding equation $F = [(F_{max})(c^n)]/[(K_d)^n + (c^n)]$, where F represents the fraction bound, c represents the concentration, K_d is the dissociation constant (the concentration at which 50% of the RNA is bound), F_{max} represents the maximal value of F, and n represents the Hill coefficient.

Telomerase Reconstitution

Pre-binding reactions were generated with 62.5 ng TER with or without 5 pmol p65 in 20 mM Tris pH 8.0, 100 mM NaCl, 1 mM $MgCl_2$, and 1 mM dithiothreitol in a final volume of 2.5 µl. Pre-binding reactions were incubated for 10 minutes at room temperature and telomerase was reconstituted in a 50 µl scale rabbit reticulocyte lysate reaction as previously described [29].

Telomerase activity assays

Telomerase activity assays were performed as previously described [30]. Some reactions were supplemented with 100 μM dATP as indicated.

Supplemental Table 1

Location	Labeling site amino acid number	Strength of hydroxyl radical cleavage
TEN	61	-
TEN	77	-
TEN	129	-
TEN - T2 motif	134	-
TEN - T2 motif	173	-
TEN	187	-
RBD - CP2 Motif	232	++
RBD	260	-
RBD	283	-
RBD - CP Motif	331	-
RBD	342	-
RBD	443	-
RBD - T motif	486	-
RBD	500	-

SUPPLEMENTAL TABLE 1. Most protein labeling sites show no site-specific hydroxyl radical cleavage. Table displaying the domain location, the motif, and amino acid number of cysteine labeling sites used in site-directed hydroxyl radical probing experiments. Cleavage products of site-directed hydroxyl radical probing experiments were compared to those obtained from unlabeled protein and scored based on the observed increase in the intensity of cleavage products (- indicates no observed increase in intensity, ++ indicates moderate to strong cleavage products). Labeling sites in the TEN domain were obtained with a construct containing residues 1-516 of TERT. We note that a negative probing result does not necessarily preclude the possibility of RNA interactions at this site. A specific labeling site may suffer from poor accessibility to Fe-BABE labeling, or the label itself could prevent the RNA interaction it was intended to measure.

References

1. Palm, W. & de Lange, T. How shelterin protects mammalian telomeres. *Annu Rev Genet* **42**, 301-34 (2008).

2. Vulliamy, T.J. & Dokal, I. Dyskeratosis congenita: the diverse clinical presentation of mutations in the telomerase complex. *Biochimie* **90**, 122-30 (2008).

3. Kim, N.W. et al. Specific association of human telomerase activity with immortal cells and cancer. *Science* **266**, 2011-5 (1994).

4. 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).

5. Hengesbach, M., Akiyama, B.M. & Stone, M.D. Single-molecule analysis of telomerase structure and function. *Curr Opin Chem Biol* **15**, 845-52 (2011).

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

7. Lin, J. et al. A universal telomerase RNA core structure includes structured motifs required for binding the telomerase reverse transcriptase protein. *Proc Natl Acad Sci U S A* **101**, 14713-8 (2004).

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

9. Lai, C.K., Mitchell, J.R. & Collins, K. RNA binding domain of telomerase reverse transcriptase. *Mol Cell Biol* **21**, 990-1000 (2001).

10. Miller, M.C., Liu, J.K. & Collins, K. Template definition by Tetrahymena telomerase reverse transcriptase. *EMBO J* **19**, 4412-22 (2000).

11. Bosoy, D., Peng, Y., Mian, I.S. & Lue, N.F. Conserved N-terminal motifs of telomerase reverse transcriptase required for ribonucleoprotein assembly in vivo. *J Biol Chem* **278**, 3882-90 (2003).

12. O'Connor, C.M., Lai, C.K. & Collins, K. Two purified domains of telomerase reverse transcriptase reconstitute sequence-specific interactions with RNA. *J Biol Chem* **280**, 17533-9 (2005).

13. Bryan, T.M., Goodrich, K.J. & Cech, T.R. Telomerase RNA bound by protein motifs specific to telomerase reverse transcriptase. *Mol Cell* **6**, 493-9 (2000).

14. Rouda, S. & Skordalakes, E. Structure of the RNA-binding domain of telomerase: implications for RNA recognition and binding. *Structure* **15**, 1403-12 (2007).

15. Mitchell, M., Gillis, A., Futahashi, M., Fujiwara, H. & Skordalakes, E. Structural basis for telomerase catalytic subunit TERT binding to RNA template and telomeric DNA. *Nat Struct Mol Biol* **17**, 513-8 (2010).

16. Heilek, G.M., Marusak, R., Meares, C.F. & Noller, H.F. Directed hydroxyl radical probing of 16S rRNA using Fe(II) tethered to ribosomal protein S4. *Proc Natl Acad Sci U S A* **92**, 1113-6 (1995).

17. Lancaster, L., Kiel, M.C., Kaji, A. & Noller, H.F. Orientation of ribosome recycling factor in the ribosome from directed hydroxyl radical probing. *Cell* **111**, 129-40 (2002).

18. Culver, G.M. & Noller, H.F. Directed hydroxyl radical probing of RNA from iron(II) tethered to proteins in ribonucleoprotein complexes. *Methods Enzymol* **318**, 461-75 (2000).

19. 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).

20. Jiang, J. et al. The architecture of Tetrahymena telomerase holoenzyme. *Nature* **496**, 187-92 (2013).

21. Robart, A.R., O'Connor, C.M. & Collins, K. Ciliate telomerase RNA loop IV nucleotides promote hierarchical RNP assembly and holoenzyme stability. *RNA* **16**, 563-71 (2010).

22. Akiyama, B.M., Loper, J., Najarro, K. & Stone, M.D. The C-terminal domain of Tetrahymena thermophila telomerase holoenzyme protein p65 induces multiple structural changes in telomerase RNA. *RNA* **18**, 653-60 (2012).

23. Stone, M.D. et al. Stepwise protein-mediated RNA folding directs assembly of telomerase ribonucleoprotein. *Nature* **446**, 458-61 (2007).

24. Sperger, J.M. & Cech, T.R. A stem-loop of Tetrahymena telomerase RNA distant from the template potentiates RNA folding and telomerase activity. *Biochemistry* **40**, 7005-16 (2001).

25. Berman, A.J., Gooding, A.R. & Cech, T.R. Tetrahymena telomerase protein p65 induces conformational changes throughout telomerase RNA (TER) and rescues telomerase reverse transcriptase and TER assembly mutants. *Mol Cell Biol* **30**, 4965-76 (2010).

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

27. Richards, R.J., Theimer, C.A., Finger, L.D. & Feigon, J. Structure of the Tetrahymena thermophila telomerase RNA helix II template boundary element. *Nucleic Acids Res* **34**, 816-25 (2006).

28. 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. *Mol Cell Biol* **22**, 1253-65 (2002).

29. 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).

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

CHAPTER IV: The RNA accordion model for template positioning by telomerase RNA during telomeric DNA synthesis

(Originally published in *Nature Structural and Molecular Biology*)

Abstract

Telomerase is a ribonucleoprotein (RNP) enzyme that maintains the ends of linear eukaryotic chromosomes and whose activation is a hallmark of 90% of all cancers. This RNP minimally contains a reverse transcriptase protein subunit (TERT) that catalyzes telomeric DNA synthesis[1] and an RNA subunit (TER) that has templating, architectural and protein-scaffolding roles[2-8]. Telomerase is unique among polymerases in that it synthesizes multiple copies of the template on the 3' end of a primer following a single binding event, a process known as repeat addition processivity (RAP). Using biochemical assays and single-molecule Förster resonance energy transfer (smFRET) experiments on *Tetrahymena thermophila* telomerase, we now directly demonstrate that TER contributes to template positioning within the active site and to the template translocation required for RAP. We propose the single-stranded RNA elements flanking the template act as a molecular accordion, undergoing reciprocal extension and compaction during telomerase translocation.

Introduction

The ribonucleoprotein particle telomerase catalyzes the addition of telomeric DNA repeats to the ends of linear chromosomes by using the reverse transcriptase motifs within telomerase reverse transcriptase (TERT) and the template contained within telomerase RNA (TER)[1,9]. Telomerase can use its internal template RNA sequence to template several consecutive rounds of telomeric DNA repeat addition, a phenomenon known as repeat addition processivity (RAP); this feature of telomerase is unique among polymerases. Because inappropriate activation of telomerase contributes to malignant transformation, understanding the molecular mechanism of the enzyme's unique ability to recycle an internal template is central to understanding telomerase function on basic and clinical levels.

While TER sequences vary widely, their RNA secondary and tertiary structural elements are largely conserved[10-12]. Most TERs contain a template sequence flanked on the 5' side by a template boundary element (TBE) and on the 3' side by a stretch of single-stranded RNA followed by a pseudoknot[13,14], as exemplified in the exceptionally streamlined *Tetrahymena thermophila* TER (Fig. 1a). Experiments with circularly permuted RNAs showed that covalent connectivity between the template and the sequences flanking it, the TBE and the template recognition element (TRE), is required for efficient translocation and may be involved in template positioning in *Tetrahymena*[15,16]. In addition, nuclease protection experiments demonstrated the TRE and template are susceptible to both single- and

double-stranded RNA nucleases, suggesting this region may alternate between a helical or stacked conformation and a flexible unstacked conformation[17].

To explain these previous observations, we developed a detailed model for the contribution of TER to template progression through the TERT active site and the subsequent translocation required for RAP (Fig. 1b, left). Telomerase has a single active site for nucleotide addition[1], so addition of nucleotides to the 3' end of the primer must necessarily be accompanied by movement of the template through the active site. Given evidence for a site of strong protein-RNA interaction 5' of the template (the TBE)[18] and the existence of a conserved RNA structure 3' of the template (the pseudoknot)[5,12], it is likely that RNA elements on both sides of the template occupy fixed positions relative to the active site (distances *d* and *d'* in Fig. 1b). We hypothesize the single-stranded regions of RNA on either side of the template undergo reciprocal compression and expansion in an accordion-like manner to accommodate movement of the template RNA during the catalytic cycle. Here we provide evidence for the RNA positioning predicted by the accordion model, which may also have implications for the energetics of template translocation.

Results

Deleting TRE nucleotides reduces RAP in a predictable manner

One prediction of the accordion model is that deletions in the TRE could prevent the 3' template nucleotides from reaching the active site, so that primers paired in the alignment region (nucleotides 49-51) would give no reaction (No

reaction, Fig. 1b, right). Primers with sequences that aligned within the template might be able to reach the active site, especially with smaller TRE deletions, but after one round of extension they would pair in the alignment region and undergo no further reaction. In general, the activities of different primers should reveal which template nucleotides are able to reach the active site in each mutant telomerase.

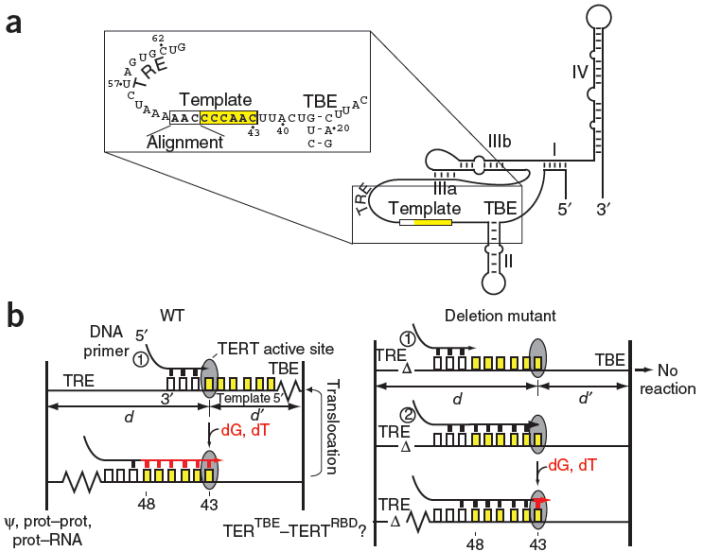


Figure 1. The RNA accordion model for primer positioning and translocation. (**a**) Secondary structure of *T. thermophila* TER, with the inset showing the sequence of the template, the TBE and the TRE. Template is boxed; the alignment region is white and the sequence that is reverse transcribed is yellow. (**b**) Accordion model for template progression and translocation of telomerase. Left, in wild-type (WT) TER, primer extension is accompanied by compaction of the TRE and extension of the TBE. Right, deletion of nucleotides in the TRE (delta symbol) prevents bound primer 1 from reaching the active site, giving No reaction. Primer 2, which pairs near the 5' end of the template, is able to undergo one round of extension. The active site of TERT is represented as a grey circle. Distances *d* and *d'* represent our assumption that the position of the active site remains constant relative to an anchored TRE and TBE, respectively. The template sequence is represented as boxes, colored as in panel **a**. Primers are in black, newly added nucleotides in red. A maximum of seven base pairs forms between the primer and template[34,35].

To test this hypothesis, we generated a series of TER molecules containing deletions in the TRE. Each mutant TER was assembled into a telomerase RNP and tested in telomerase activity assays with primers representing each of the six permutations of telomeric sequence (Fig. 2a); wild-type telomerase utilizes all primers efficienty (Fig. 2b). Single-nucleotide TRE deletions were tolerated by telomerase (Fig. 2c and Supplemental Fig. 1a). For all double-nucleotide deletion mutants, RAP was reduced compared to the activity of wild-type (WT) telomerase (Supplemental Fig. 1b). Deleting three sequential nucleotides had a dramatic effect, allowing robust addition of one repeat but very little RAP (Fig. 2d and Supplemental Fig. 1c). Primer B, which anneals to the alignment region of the template, was essentially inactive even for the first round of extension, and primer C showed reduced activity. This primer usage profile was independent of the location of the deletion within the TRE, suggesting that sequence is not as important as length. A similar observation was made for mutants with four sequential nucleotides deleted, but now primer C was also inactive (Fig. 2e), suggesting that in all four-nucleotide deletion mutants, template nucleotide 47 can no longer reach the active site. Deletion of six nucleotides (Δ54-59) continued the trend, preventing TER nucleotide 46 from being used as template. The largest deletion tested (Δ52-59) was only able to efficiently use primer A, which base pairs with the penultimate nucleotide of the template (Fig. 3). This eight-nucleotide deletion presumably acts as a molecular ruler, measuring the minimum distance between the pseudoknot and C43 in the active site.

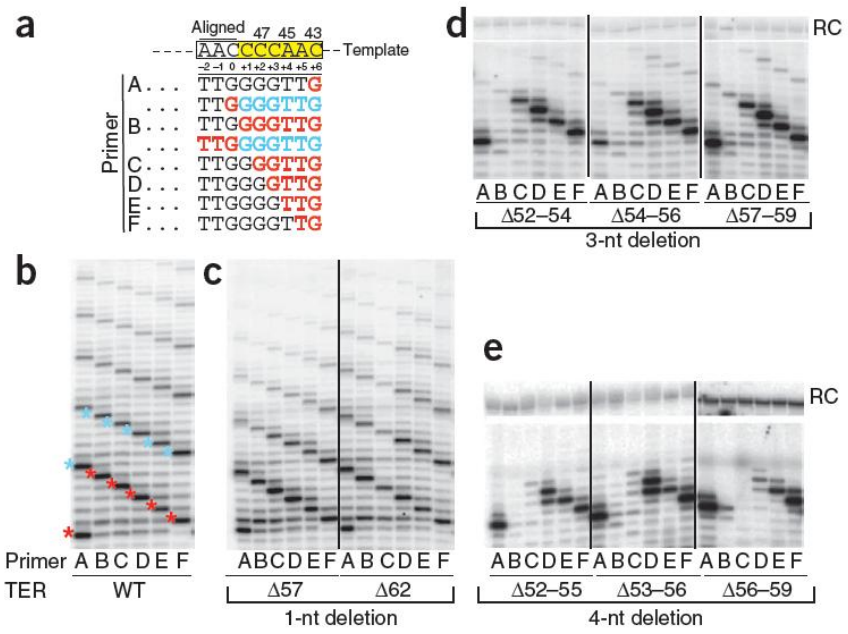


Figure 2. Deletion mutants in the TRE have predictable primer usage profiles. (**a**) The 3' ends (black) of the six primers of permuted telomeric sequence aligned with the template. All primers were 18 nucleotides (three repeats of permuted telomeric sequence. Nucleotides added during the first round of synthesis are red. Nucleotides added during the second round are shown for primers A and B in blue. Templating position is indicated below the template sequence, and is defined by the first nucleotide that can be copied[3]. (**b-e**) Telomerase activity assays for WT (**b**) and a selection of single (**c**), triple (**d**) and quadruple (**e**) deletion mutants within the TRE of TER with each of the indicated primers. Red and blue asterisks in the WT panel indicate the first and second rounds of RAP, respectively. RC, recovery control.

Insertions in the TBE reinforce template misalignment

A second prediction of the accordion model is that flexibility in the single-stranded RNA portion of the TBE, which is on the 5' side of the template, is also required for template progression. Insertions of nucleotides adjacent to the TBE are predicted to reinforce the misalignment of the template relative to the active site in the context of TRE deletion mutants. We modified two TRE deletion mutants (Δ56-

59, a four-nucleotide deletion, and Δ54-59, a six-nucleotide deletion) by inserting two uridines between nucleotides 40 and 41 of the TBE, just 5' of the template (Fig. 3). While the four-nucleotide TRE deletion had only a modest ability to use primer D, the combination of the insertion of 2Us in the TBE and the four-nucleotide TRE deletion exacerbated this effect, resulting in a primer usage profile similar to that of the six-nucleotide TRE deletion mutant. Similarly, the six-nucleotide deletion mutant containing an insertion of two uridines 5' of the template had a weakened ability to use primer E, matching the primer usage profile of the eight-nucleotide deletion mutant.

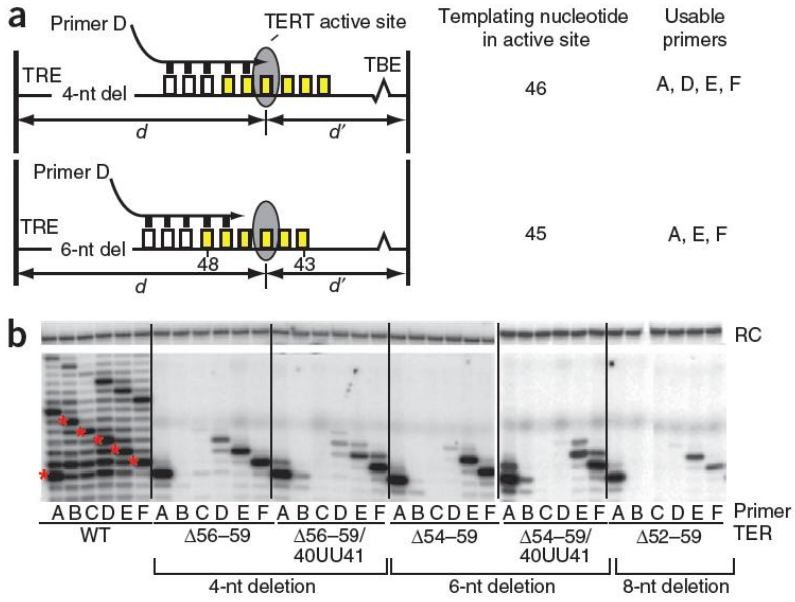


Figure 3. The accordion model predicts primer usage for TRE deletion mutants. (**a**) Prediction of the interaction of primer D ($[GTTGGG]_3$) with a four-nucleotide deletion mutant (top) and a six-nucleotide deletion mutant (bottom). (**b)** Telomerase activity assays on deletion and deletion-insertion combination mutants. Asterisks within the WT lanes indicate the first repeat for each primer. The exact nucleotides deleted and inserted are indicated for each panel. RC, recovery control.

Discontinuities in the RNA near the template eliminate RAP

Previous experiments with circularly permuted (cp) telomerase RNAs demonstrated that discontinuities on either side of the template decreased RAP[16]. We interpret these results as showing that coordinated structural rearrangements on both sides of the template contribute to template movement. The previous experiments were performed in the absence of p65, the telomerase-specific RNA folding protein[19] that is essential *in vivo*[20] and has the ability to rescue many severely destabilizing TERT and TER mutations *in vitro*[17]. If nicked RNAs were defective simply because of structural destabilization, they might likewise be suppressed in telomerase containing p65.

We therefore tested the activities of cpRNAs assembled into telomerase RNPs in both the presence and absence of p65 (Supplemental Fig. 2). Because the TBE is necessary for TERT binding[18,21] and proper telomerase function[3], we tested four circularly permuted RNAs (cpRNA) within this region, and one in the TRE. Telomerase RNPs containing the cpRNAs had essentially full first-round activity, but thereafter RAP was severely reduced, with nicks in the TBE affecting the reaction similarly to the nick at position 59 in the TRE. This limited RAP even in the presence of p65 provides further support for the requirement for direct connectivity of the RNA on both sides of the template.

Single-molecule FRET observes predicted RNA movement

Taken together, these biochemical experiments demonstrate that specific deletions and insertions, as well as the introduction of nicks, in the RNA elements flanking the template evoke catalytic defects predicted by the accordion model. However, these biochemical assays cannot directly probe RNA structure; therefore single-molecule Förster resonance energy transfer (smFRET) was employed to monitor RNA conformational changes within individual telomerase-DNA primer complexes at successive steps in the telomerase catalytic cycle (Fig. 4a and Supplemental Fig. 3)[19]. Ideally, smFRET experiments performed in the presence of dTTP and dGTP could permit real time observation of TER structural rearrangements throughout the telomerase catalytic cycle. Thus far, however, this approach has not produced interpretable results. Therefore, to simplify interpretation of our data, independent smFRET experiments were performed with primers successively extended on the 3' end in the absence of dGTP and dTTP. In this way, the primers used in separate experiments emulate the different conformational states that would be sampled during telomerase-catalyzed DNA synthesis. To ensure telomerase enzymes bound the RNA template in a single register, primers used in our smFRET experiments consisted of a 5'-$(TG)_8$ stretch followed by a single copy of the various telomere repeat permutations (Supplemental Table 1, Supplemental Fig. 4)[22]. Since FRET efficiency is highly sensitive to the distance between the dyes, FRET provides a direct readout of the intervening RNA structure with high FRET and low FRET corresponding to compressed and extended RNA, respectively.

Surface-immobilized telomeric DNA primers coupled to a donor dye (Cy3) were monitored in the presence of purified telomerase enzymes harboring an acceptor dye (Cy5) at U63 in the TRE (Fig. 4a, top panel) or at U36 in the TBE (Fig. 4a, bottom panel). Importantly, these modified DNA primers and telomerase RNAs have been shown previously to support near wild-type levels of telomerase activity[23]. In this experiment, the binding of a Cy3-labeled DNA primer within the active site of a Cy5-labeled telomerase enzyme is observed in real time as the onset of energy transfer between the two FRET probes (Fig. 4b). For each experiment, histograms corresponding to the frequency of observed FRET values were compiled and fit with Gaussian functions to determine the center of the distribution (Fig. 4c and Supplemental Fig. 4).

Measurements from the bound primer to the TBE show a dramatic decrease in FRET as the template-complementary sequence at the primer 3' end is extended (Fig. 4d). Measurements of the distance from the primer to the TRE show a smaller but appreciable rise in FRET with the same series of primers. These reciprocal changes are just as predicted by the accordion model. The larger FRET change measured across the TBE region than across the TRE has several possible explanations. First, the compression in the TBE may be greater in magnitude than the expansion in the TRE. A second possibility is that the U63 labeling position may not be fixed in space throughout the catalytic cycle and further expansion may occur 3' of the TRE. Lastly, the observed asymmetry could be explained if the donor dye movement is out of plane with respect to the position of the acceptor dye. In other words, in three-

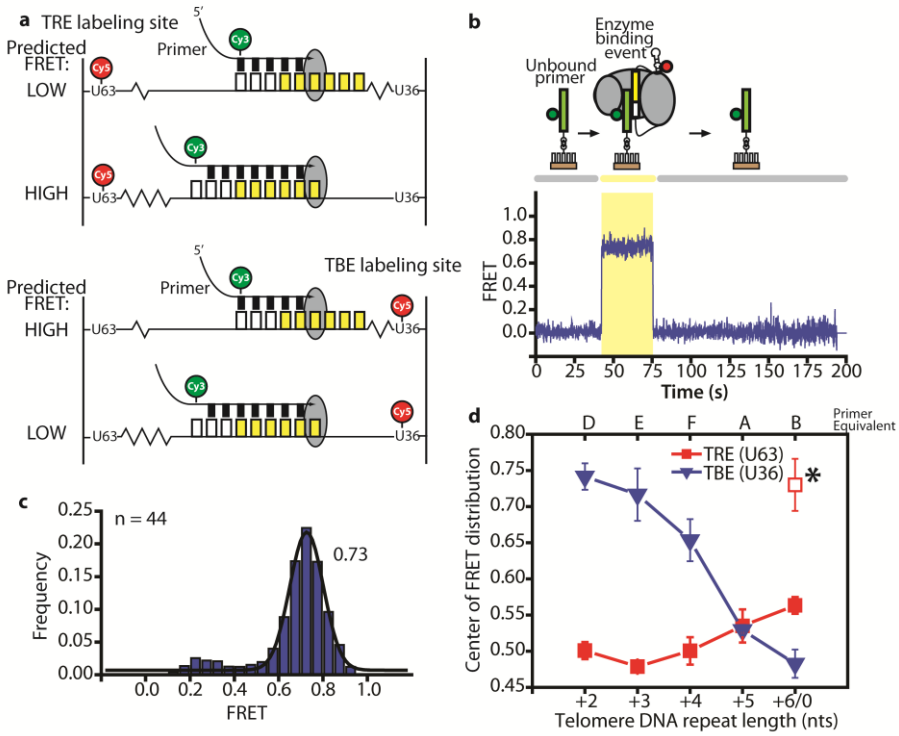


Figure 4 Single-molecule FRET measurements of RNA compression and extension at successive steps of telomere synthesis. (**a**) The accordion model predicts RNA extension (low FRET) and compression (high FRET) in the TRE for short and long telomere repeat primers, respectively (top panel). The opposite trend is predicted for the TBE region (bottom panel). In addition to the TBE, it is possible part of the template itself compacts. (**b**) Representative single molecule trace of a primer-telomerase binding event, detected by the onset of FRET (shaded area). (**c**) Representative FRET histogram of 44 binding events of a GT-rich primer with the same 3' terminal sequence as primer D (Figure 2 and Supplemental Fig. 4) bound by TBE (U36)-labeled telomerase. A Gaussian fit determines the center of the distribution to be 0.73 FRET. (**d**) Mean values for the center of each FRET distribution plotted as a function of the length of the 3' extension of the primer. Complete primer sequences shown in Supplemental Fig. 4. TBE labeling site (U36) data, blue triangles. TRE labeling site (U63) data, red squares. Error bars are the standard deviation of triplicate experiments (number of binding events in each experiment varied from n = 23 to 106). Asterisk represents an alternative FRET state. Primer B with U63-labeled enzyme demonstrated a bimodal distribution, centered at FRET=0.57 and 0.73, as shown in Supplemental Figs. 4 and 6. The points are plotted separately and the line is drawn to the 0.57 FRET state on the basis of continuity with previous primers.

dimensional space the relatively small FRET change observed for the TRE may in fact correspond to a significant rearrangement of the RNA.

Many of the smFRET distributions generated in our experiments are broader than would be expected for a single FRET distribution. In particular, we observed a gradual broadening of the FRET distributions as the length of the DNA primer was increased (Supplemental Fig. 4, primers F, A, and B). Analysis of smFRET trajectories revealed the broadening to be derived from anti-correlated donor and acceptor fluctuations in individual traces over a narrow range of FRET values (~0.15 FRET) (Supplemental Fig. 5 and 6). Since the physical basis for these FRET states is not known at present, we opted to treat the majority of the data by fitting to a single Gaussian function to extract the average FRET value for each primer.

One exception to this approach was for data obtained with the TRE (U63) label and the primer corresponding to a complete telomeric repeat (primer B). In this experiment, we observed a substantial number of binding events exhibiting a long-lived 0.73 FRET state (Figure 4d, see asterisk, Supplemental Fig. 6, Primer B) which is likely due to an alternative primer conformation. These data were therefore fit with two Gaussian functions (centered at 0.57 and 0.73) in order to avoid the artificially large apparent increase in FRET resulting from a single- Gaussian fit. Notably, transient binding events corresponding to a similar high FRET state were also observed for the shorter primers (D, E, F, A) (Supplemental Fig. 7, TRE (U63)-labeled enzyme), but their short durations resulted in a negligible contribution to the FRET histograms. Given that primer B is the only primer that shows this stable

alternative high FRET state and that primer B represents the substrate for translocation during RAP, it is possible that the high FRET state represents a post-translocated state of the enzyme that is preferentially stabilized for telomeric primers terminating in a complete telomere repeat sequence. Importantly, our conclusion that there is increased FRET between the TRE (U63) and the most extended primer holds if we consider either of the two individual FRET states (0.57 and 0.73) or the average of the two.

Discussion

Because TERT is a reverse transcriptase related to retroviral and retrotransposon RTs, it is thought to use a similar metal-ion-catalyzed mechanism for nucleotide addition[1,24,25]. In contrast, the mechanism by which it achieves its unique activities -- template translocation and repeat addition processivity -- remains unknown. Here, the predictable effects of nucleotide deletions on primer use and measurements of sequential primer position using single-molecule FRET both support an RNA accordion model for RAP. We conclude that the RNA elements flanking the template serve to position the template for nucleotide addition and for RAP.

We have considered alternative models that could explain our data. The requirement for connectivity of the RNA on the 5' side of the template might be a consequence of its interaction with the RNA-binding domain of TERT[18], which might be disrupted by nicking. This is difficult to address in the absence of a detailed

structure of the telomerase RNP. In addition, a small RNA stem-loop might form within the TRE during telomere synthesis, and this stem-loop could melt for translocation. This alternative model has been tested with deletion and point mutations predicted to stabilize or destabilize the putative hairpin. If hairpin formation and melting contributed to translocation, one would expect a large effect on RAP in the context of these mutants, similar to the effect on RAP observed with nicked RNAs or with RNAs containing deletions of three or more nucleotides in the TRE. However, in the presence of mutants that either stabilized or destabilized the putative hairpin, only a modest impact on RAP was observed; RAP decreased by less than fifty percent as compared to WT for all mutants tested (Supplemental Fig. 8).

While our data do not directly address the energetics of translocation, the disruption of RAP by RNA nicking observed by Collins[16] and in our Supplemental Fig. 8 suggests that the RNA has an active role in template movement. More specifically, upon synthesis of a complete telomere DNA repeat, the TRE region of the RNA may exist in a high-energy state. For example, this single-stranded region of RNA could become compressed, or energetically unfavorable TER-TERT interactions could be produced. The relaxation of this high-energy state would then drive translocation.

A precedent exists for polymerase translocation powered by energy stored within nucleic acid. Polymerases move along their templates without external energy; the energy provided by nucleotide incorporation and pyrophosphate release can be converted directly into directional movement or it can be stored within

scrunched nucleic acid that can later be manifested as translocation. In RNA polymerases, for example, the DNA template bunches near the active site during RNA synthesis initiation, resulting in a higher energy state that likely provides the energy for promoter clearance[26-28]. The conformational changes observed within telomerase TER are reminiscent of this mechanism. Here, however, rather than the primer DNA storing energy, perhaps the RNA downstream of the template bunches up as the template is copied, potentially storing energy during repeat synthesis, and releasing it to drive template recycling. The molecular basis for compaction could also involve base stacking[17] and RNA-TERT protein interactions.

In addition, it is possible that a ratchet might be involved in translocation; the pawl for this ratchet would eliminate backtracking of the template until the end is reached[29]. A ratchet mechanism is compatible with RNA conformational changes being necessary for translocation, and thus could represent an elaboration of the accordion model. Further understanding of the translocation mechanism awaits detailed structural information concerning the RNA-protein interactions in the vicinity of the active site and their dynamics during processive extension of telomeric DNA.

Materials and Methods

RNA construct design and preparation

RNA point, deletion and substitution mutants were cloned using the site-specific mutagenesis kit QuikChange (Agilent) with ptetTELO as the template[30]. The

cpRNA constructs were generated using two rounds of PCR using ptetTELO as the template and then inserted into pUC19; the sequence of the linker joining the natural 5' and 3' ends of TER is TTTTGGATCC[16]. RNA was synthesized using T7 RNA polymerase for run-off transcription of EarI-digested plasmids containing the template for TER, TER mutants or cpRNAs. Samples were transcribed overnight at 37 °C prior to hammerhead cleavage with 25 mM $MgCl_2$ at room temperature for an hour. Samples were then ethanol precipitated prior to denaturing gel purification (6% v/v polyacrylamide, 7 M urea and 1X TBE). Purified RNAs were eluted and concentrated in 5 mM sodium cacodylate, pH 6.5.

Construction of dye-labeled telomerase RNA

Cy5-labeled telomerase RNA was constructed as described[31]. Briefly, synthesized RNA fragments harboring a site-specific reactive amine group at position U36 or U63 were ordered from Dharmacon. The fragment containing the modification was conjugated to amine-reactive Cy5 (GE Healthcare) and purified by reverse phase HPLC on a C8 column. Full-length telomerase RNAs were generated in a three-part DNA-splinted RNA ligation with T4 DNA ligase (NEB).

Telomerase reconstitution and immunoprecipitation

Telomerase was reconstituted *in vitro* in RRLs as previously described[17,32]. The p65 protein used in these translations was purified as previously described[17]. The final concentrations of RNA and p65 in the translation reactions were 400 nM and 2 μM, respectively. Assembled telomerase was immunoprecipitated with T7-tag antibody-conjugated agarose (Novagen). ^{35}S-met (Perkin Elmer) was quantitated

using a scintillation counter, and all samples were diluted with 1x tTB (50 mM Tris-HCl, pH 8.3, 1.25 mM $MgCl_2$, 5 mM DTT)+ 30% v/v glycerol to normalize for the amount of ^{35}S-TERT.

Full-length Cy5-labeled RNAs were PAGE purified, pre-bound with recombinantly expressed and purified p65, and reconstituted with FLAG-labeled TERT using the TnT Quick Coupled Transcription-Translation System (Promega). Reconstituted RNAs with dyes at either position were assayed for activity, and showed comparable activity to wild-type RNA as previously described. The reconstituted TERT was purified by immunoprecipitation with anti-FLAG resin (Sigma) and eluted with FLAG peptide (Sigma) as described.

Direct telomerase activity assays

Telomerase activity assays were performed as previously described[17,32]. All reactions were incubated at 30 °C for 1 h and then stopped with stop buffer containing 3.6 M sodium acetate, pH 5.2, 1 mg glycogen and 2,000-6,000 counts of a recovery control (a 5'-end labeled and gel-purified 63-mer DNA oligonucleotide of random sequence), with the exception of the assays shown in Supplemental Fig. 2b, which were incubated at 30 °C for 1.5 h to maximize signal. All reactions conducted at the same time received the same number of radioactive counts of recovery control. Radioactive nucleotides were purchased from Perkin Elmer.

Real time detection of telomerase binding to DNA primers by single molecule FRET

Single molecule FRET-based telomerase binding assays were conducted as previously described[23]. Biotinylated primers labeled at the n-2 alignment residue with Cy3 corresponding to different overhang lengths were immobilized on a quartz slide through a biotin-streptavidin linkage (biotinylated polyethylene glycol from Laysan Bio, strepatvidin from Invitrogen, monoreactive Cy3 and Cy5 from GE Lifesciences). Immunopurified telomerase complexes were flowed over immobilized primers in telomerase binding and imaging buffer (20 mM Tris pH 8.0, 5 mM $MgCl_2$, 10% w/v glucose, 10% v/v glycerol, 2 mM trolox, 1µg ml^{-1} catalase, and 1.5 mg ml^{-1} glucose oxidase). Binding events were identified by anti-correlated donor and acceptor dye fluorescence detected using a prism-type TIRF microscope and an Andor Ixon CCD camera using 100 millisecond integration time. 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. γ is a correction factor that adjusts for differences in dye quantum yields, and can be useful in correcting FRET efficiency when there is a boost in Cy3 quantum yield upon protein binding as was observed in a subset of our traces. Because in our system we cannot distinguish between acceptor bleaching events and enzyme dissociation from the primer, we could not determine γ by the standard method of detecting the change in donor and acceptor intensities upon acceptor bleaching. 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 $I_{D2} + I_{A2}$ represent the sum of the donor and acceptor intensities after binding. γ was determined individually for each trace and was

consistent with previously reported values observed due to protein binding Cy3-labeled DNA[33].

Each experiment consisted of traces obtained from 40 200-second movies, which generated between 23-106 separate binding events. Binding event traces were analyzed in Matlab and FRET values from binding events from each experiment were used to generate histograms (see Fig. **4c** and Supplemental Fig. 5). Each histogram consisted of all of the FRET values observed at each time point over the course of all of the binding events. A Gaussian distribution was fit to the histogram to determine the center of the distribution. All plotting and fitting of the single-molecule FRET data was performed in Origin. The experiment was repeated in triplicate for each primer permutation with both RNA labeling sites.

SUPPLEMENTAL TABLE 1. Oligonucleotides used in single molecule FRET Experiments

TER 1-31 Fragment

AUA CCC GCU UAA UUC AUU CAG AUC UGU AAU A

U36 TER 32-95 fragment

5'-P-GAA C[5-N-U]G UCA UUC AAC CCC AAA AAU CUA GUG CUG AUA UAA CCU UCA CCA AUU AGG UUC AAA UAA G

U63 TER 32-95 fragment

5'-P-GAA CUG UCA UUC AAC CCC AAA AAU CUA GUG C[5-N-U]G AUA UAA CCU UCA CCA AUU AGG UUC AAA UAA G

TER 96-159 fragment

5'-P-UGG UAA UGC GGG ACA AAA GAC UAU CGA CAU UUG AUA CAC UAU UUA UCA AUG GAU GUC UUA UUU U

Stem III DNA Splint

5' TGT CCC GCA TTA CCA CTT ATT TGA ACC TAA 3'

Stem II DNA Splint

5’ GTT GAA TGA CAG TTC TAT TAC AGA TCT GAA 3’

Telomere DNA Primers

Primer D : Biotin – TGTGTGTGTGTGTGTG [5-Amino-C6-dT] TGGG

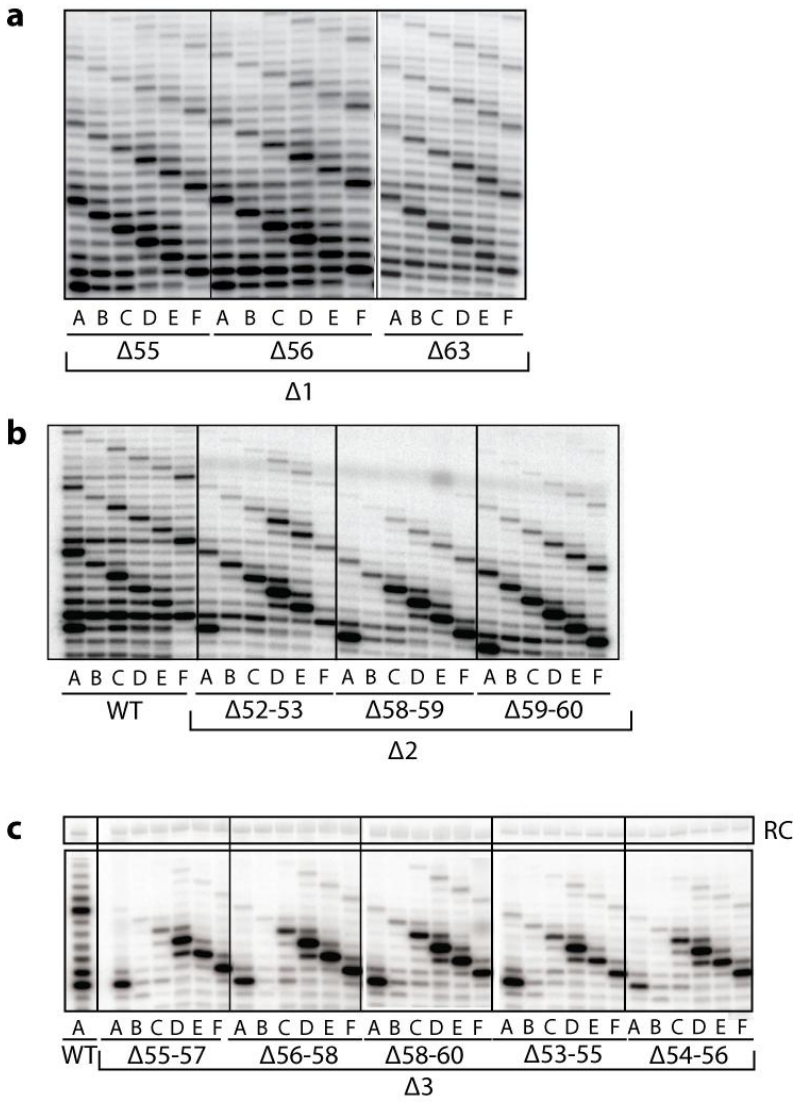
Primer E : Biotin – TGTGTGTGTGTGTGTG [5-Amino-C6-dT] TGGGG

Primer F : Biotin – TGTGTGTGTGTGTGTG [5-Amino-C6-dT] TGGGGT

Primer A : Biotin – TGTGTGTGTGTGTGTG [5-Amino-C6-dT] TGGGGTT

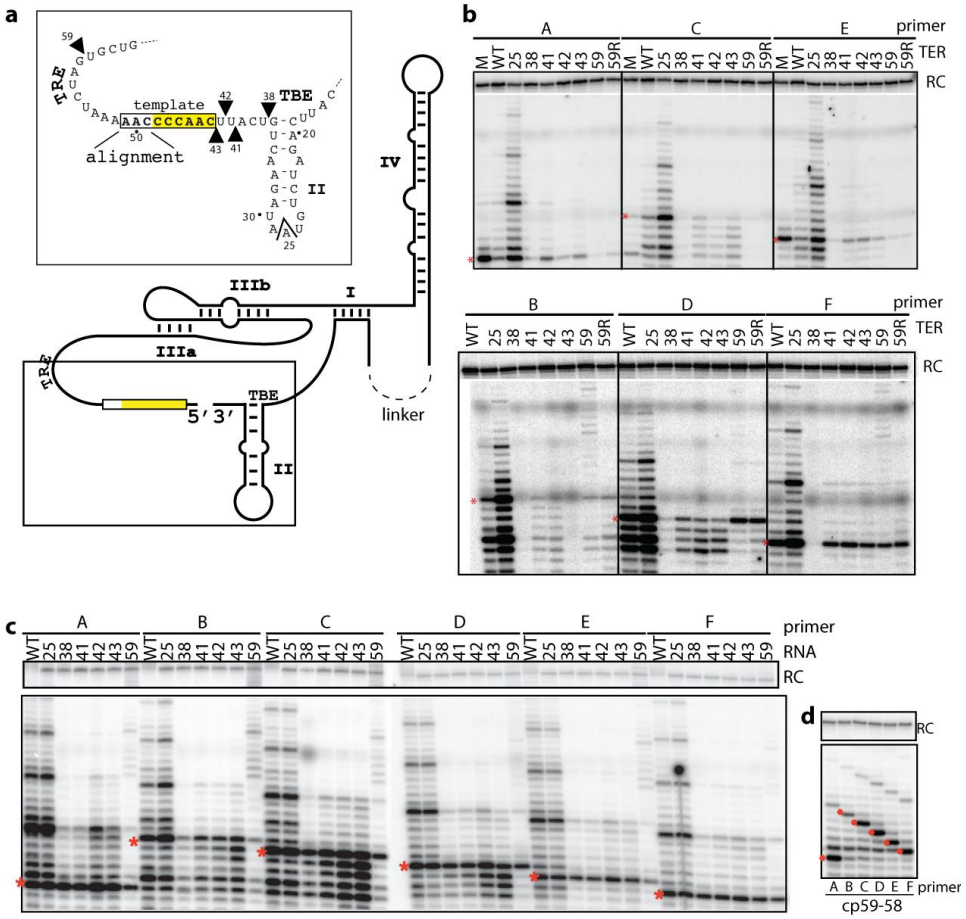
Primer B : Biotin – TGTGTGTGTGTGTGTG [5-Amino-C6-dT] TGGGGTTG

SUPPLEMENTAL FIGURE 1



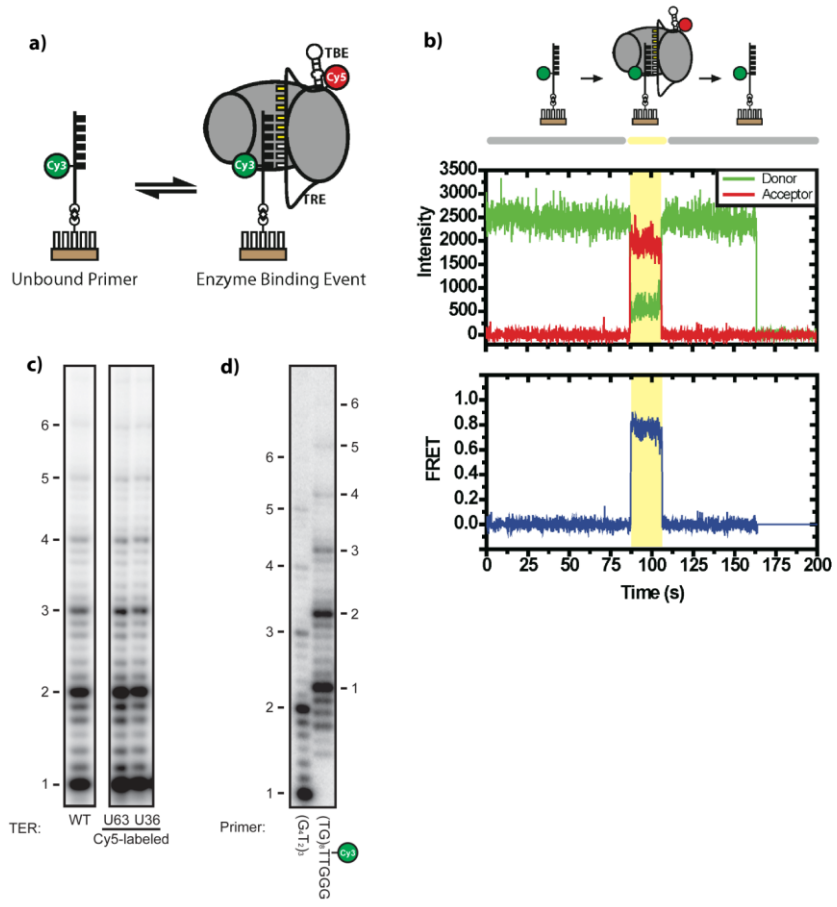
Supplemental Fig. 1. Primer profiles of deletion mutants. Telomerase activity assays with the indicated primers for **a**, single nucleotide deletions, **b,** double nucleotide deletions and **c,** triple nucleotide deletions. RC, recovery control.

SUPPLEMENTAL FIGURE 2



Supplemental Fig. 2. Telomerase molecules containing nicks in the TRE or the TBE exhibit minimal telomerase RAP. a, Construction of cpRNAs. The natural 5‘ and 3’ ends were connected by a linker[16] (dashed line). Inset, TER sequence spanning the template in the cpRNAs. The closed triangles indicate where each circular permutation was generated, with the 5’ end of the cpRNA labeled. For example, 42 is a circular permutation with the 5’ end at nucleotide 42 and the 3’ end at nucleotide 41. cp25 begins at nucleotide 31 and ends at nucleotide 25; three GC base pairs terminate this stem for stabilization[16]. **b,** Telomerase activity assays of indicated cpRNAs and WT TER assembled in the (**b**) absence and (**c-d**) presence of p65. **d,** Telomerase activity assays on telomerase containing cp59-58 RNA that were treated during the last five minutes of the assay with RNase A. Primers are indicated below each lane. Asterisks, first round of repeat synthesis for each primer. RC, recovery control.

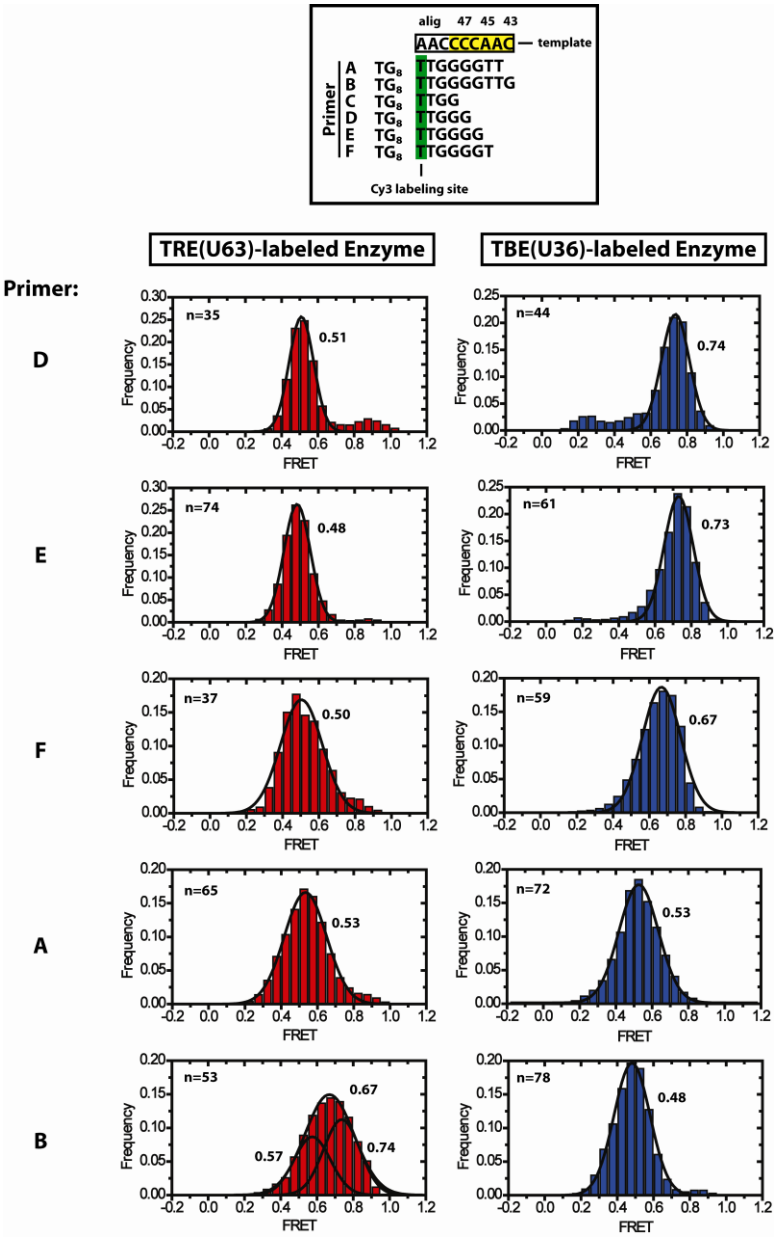
SUPPLEMENTAL FIGURE 3



Supplemental Fig. 3. Real-time detection of telomerase-DNA interactions by single molecule FRET. A) Surface immobilized telomere DNA primers labeled at the n-2 alignment position with a FRET donor dye (Cy3) are incubated in the presence of telomerase enzymes labeled with an acceptor dye (Cy5) within the telomerase RNA. **B)** Binding of the Cy5-labeled telomerase to the surface immobilized Cy3-labeled DNA primer gives rise to FRET, characterized by the anti-correlated change in dye donor (green) and acceptor (red) dye intensities (top panel). The sudden drop in donor intensity at the end of the single molecule trace is due to irreversible photobleaching of the dye. FRET was calculated as described in the Methods section. **.** C**)** Telomerase activity assays were performed on telomerase enzymes reconstituted with unmodified *in vitro* transcribed TER (WT) or with TER labeled at either U63 or U36 with Cy5. The telomerase activity assays demonstrate that fluorophore modification at these two sites has no effect on telomerase processivity. Numbers indicate how many telomeric repeats have been added at each

band. **D)** Telomerase activity assays were performed either in the presence of a telomeric 18-nt DNA primer $(G_4T_2)_3$ or in the presence of a $(TG)_8$ repeat primer labeled with Cy3 at the n-2 position. The activity assays reveal that Cy3-labeled $(TG)_8$ primers support a similar level of telomerase activity as the unmodified native telomeric sequence primers. The $(TG)_8$ primer appears higher on the gel due to the combined effect of being three nucleotides longer and containing the Cy3 modification.

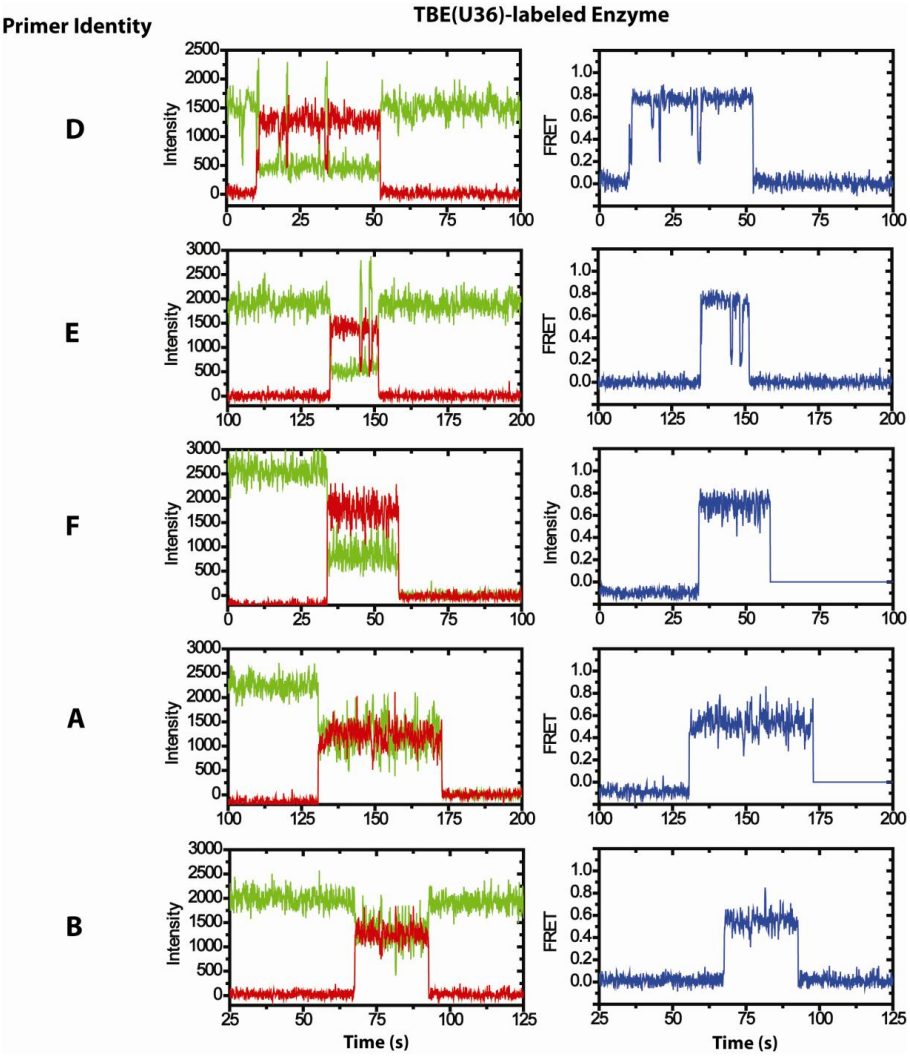
SUPPLEMENTAL FIGURE 4



Supplemental Fig. 4. Representative smFRET histograms for each telomere DNA primer permutation. smFRET experiments were conducted using DNA primers with a 5' $(TG)_8$ stretch followed by a single copy of various permutations of telomere repeat sequence corresponding to intermediates in the telomere DNA repeat synthesis reaction as shown in the top panel. Representative histograms for TRE (U63)- and TBE (U36)-labeled enzymes binding to the indicated telomere DNA primers were generated by compiling all observed FRET values for the duration of the indicated number of binding events. Histograms were fit to Gaussian functions to determine the center of the FRET distribution.

Notably, the FRET distributions for both the TRE (U63)- and TBE (U36)-labeled enzyme become gradually broader as the primer length is increased (primers F, A, B), and are not consistent with a single stable FRET distribution. As discussed further in Supplemental figures 5 and 6, the histogram broadening is the result of transient FRET fluctuations for these primers. However, since the physical basis for these FRET dynamics is not understood at present, we chose to fit the data with a single Gaussian to obtain the average FRET behavior for all experiments except the experiment with TRE (U63)-labeled enzyme and primer B, which showed a distinct bimodal distribution of FRET values (see Supplemental Fig. 7, primer B). In this case, the broad FRET distribution was fit with two Gaussian functions rather than one in order to suppress an artificially large increase in the apparent FRET value, which would be observed if we treated the data as a unimodal distribution. Experiments performed with the shortest primer, primer C, yielded very few binding events. In addition they tended to be very short-lived and unstable in their FRET values. As a result, primer C was not used in this study.

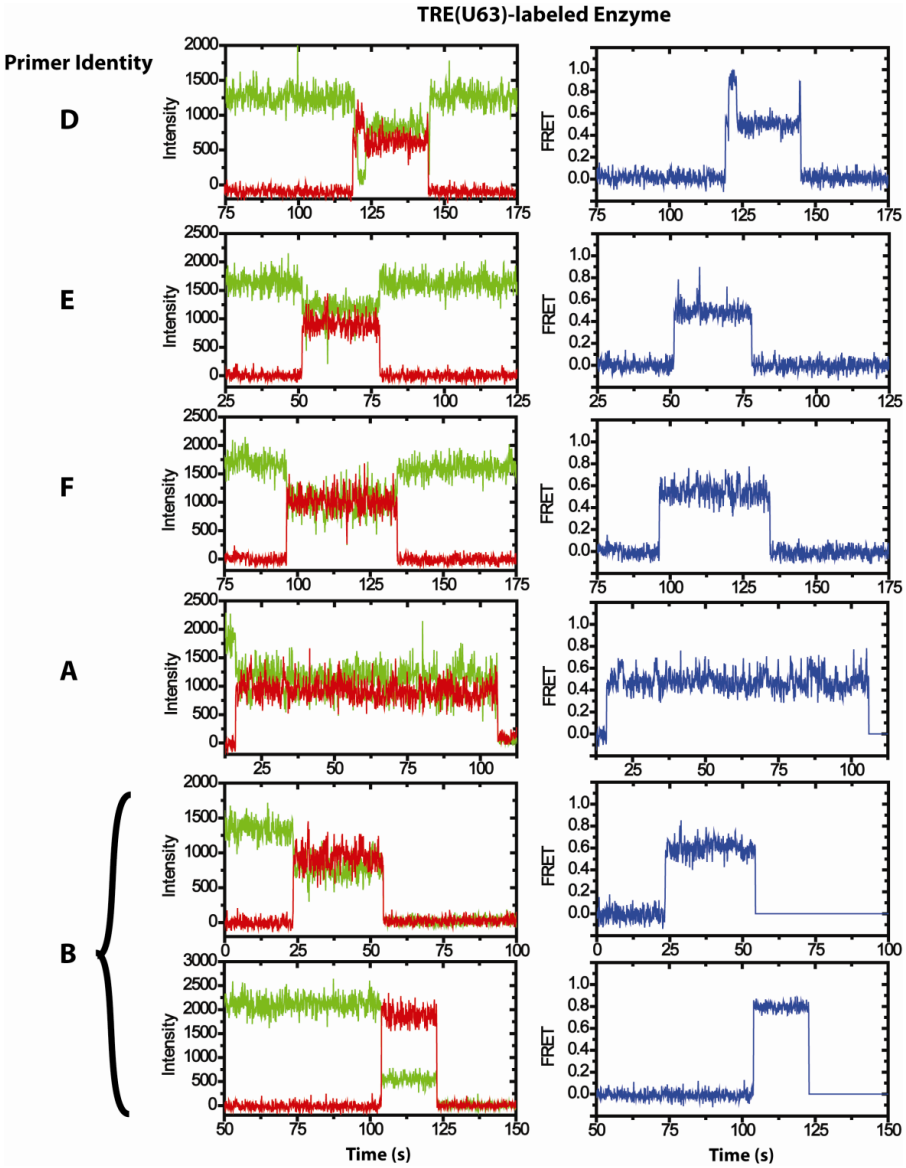
Supplemental Figure 5



Supplemental Fig. 5. Representative single molecule FRET trajectories for TBE (U36)-labeled telomerase. Examples of single molecule traces obtained with each of the indicated primer types are shown. The left column shows the donor (green) and acceptor (red) dye intensities and the right column is the calculated FRET ratio after normalization and correction for Cy3 intensity enhancement due to the presence of the protein (see Methods for more details on γ correction). In the case of primer D, transient fluctuations to a lower FRET state (~0.2) were observed, giving rise to a small additional peak in the FRET histogram (see Supplemental Fig. 4, primer D). The origin of these FRET dynamics is not known at present. One possibility is that this low FRET state represents a complex in which the DNA primer is only bound at

the previously described anchor site, which comprises an interaction surface that is independent of the enzyme active site. As the primer length is increased, we observe oscillations in individual FRET traces consistent with a broadening observed in the histograms for these primers (see Supplemental Fig. 4). In some cases, these FRET fluctuations could be clearly resolved as anti-correlated donor and acceptor dye intensities, suggesting the increase in FRET dynamics likely represents a transient rearrangement of the FRET probes. These fluctuations may represent functionally important structural dynamics of the telomere DNA primer, telomerase RNA, or both; however, a precise description of the physical basis for the observed FRET dynamics awaits further investigation.

Supplemental Figure 6

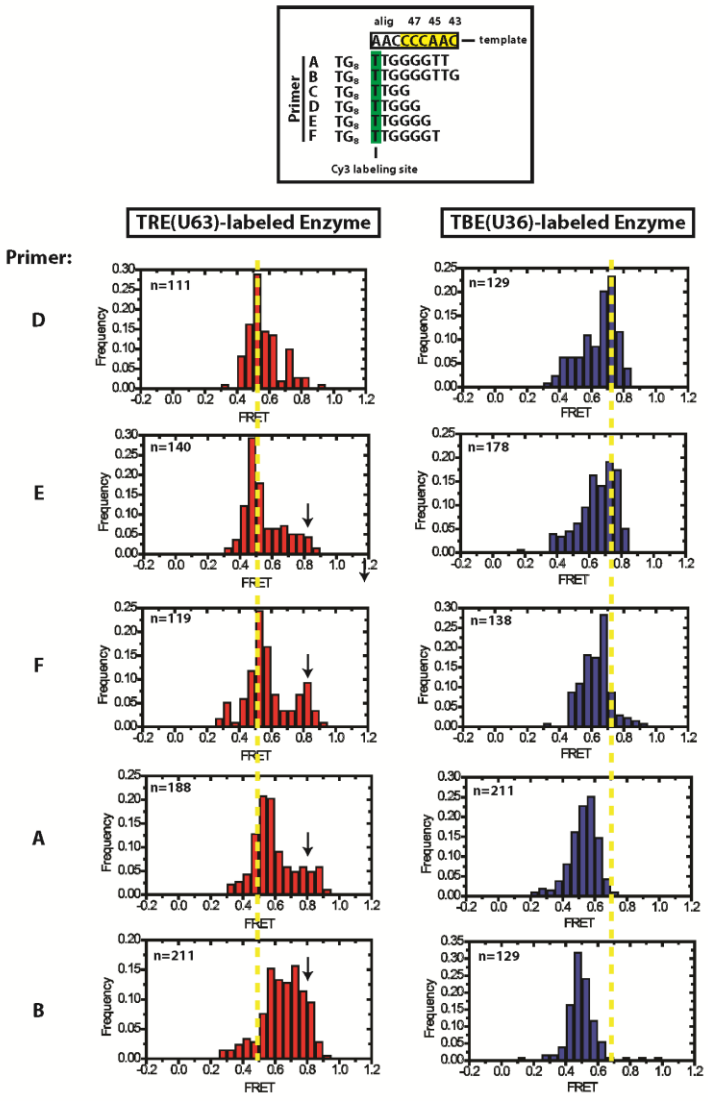


Supplemental Fig. 6. Representative single molecule FRET trajectories for TRE (U63)-labeled telomerase. Examples of single molecule traces obtained with each of the indicated primer types are shown. The left column shows the donor (green) and acceptor (red) dye intensities and the right column is the calculated FRET ratio after normalization and correction for Cy3 intensity enhancement due to the presence of the protein (see Methods for more details on γ correction). As described in Supplemental Fig. 6, experiments performed with primer D showed a predominantly stable FRET level interrupted by transient FRET transitions. In the case of the TRE (U63) enzyme, we observed transient excursions to a higher FRET value, giving rise

to a small peak at FRET ~0.9 in the FRET histograms (see Supplemental Fig. 4, primer D). Interestingly, the direction of these transient FRET dynamics is anti-symmetric for the TRE (U63)- and TBE (U36)-labeled enzymes, suggesting the signal is derived from a rearrangement of the DNA primer that brings the Cy3 label closer to the TRE and further from the TBE, consistent with a primer binding mode in which only the DNA anchor site (located near the TRE) is occupied. As was seen for the TBE (U36)-labeled enzyme, we again observed a gradual increase in the amplitude of the FRET fluctuations with longer DNA primers (primers F, A, and B). In a subset of the traces, these fluctuations were clearly resolved as anti-correlated donor and acceptor dye intensities, suggesting they are the result of a dynamic reorientation of the FRET probes, rather than a photo-physical artifact.

Notably, in the case of the TRE (U63) enzyme and DNA primer B (corresponding to a complete telomeric DNA repeat sequence), in addition to the trend of increased FRET dynamics around 0.5-0.6 FRET, we also observed substantial number of single molecule traces showing a very stable high FRET behavior (FRET~0.75). The observation of these two sub-populations could be explained by the fact that primer B, which corresponds to a complete telomere DNA repeat may bind in two distinct registers, one corresponding to the pre-translocated state and the other to the post-translocated state of the enzyme with respect to the primer. Importantly, as stated in the main text, the trend we observe for structural changes in the TRE holds whether we consider either of the two individual FRET states or the average of the two. Curiously, we only observe a single distribution for the TBE (U36) labeling site with primer B (see Supplemental Fig. 6). There are two possible explanations for this observation. One possibility is that the increase in FRET is due to movement of the U63 labeling site closer to a primer that remains in a fixed position in the enzyme. The other possibility is that it is the primer labeling site that moves relative to fixed U63 and U36 labeling sites, however the movement is such that the primer moves closer to the U63 labeling site without changing its net distance to the U36 labeling site.

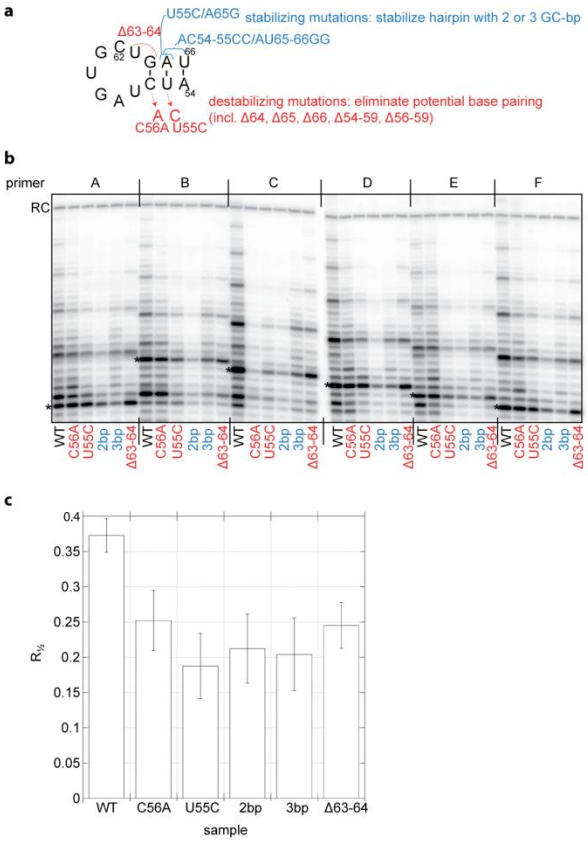
Supplemental Figure 7



Supplemental Fig. 7. Mean FRET histograms for each telomere DNA primer permutation. As an alternative way to represent the FRET data, we took the mean FRET value for each individual binding event and binned these values into mean FRET histograms. Since this approach yields fewer data points, we combined all binding events for each enzyme-primer combination from each of the experiments performed in triplicate to generate the mean FRET histograms. Importantly, the trends in the FRET values for the dominant populations observed in our experiments are similar in the mean FRET histograms and the total FRET histograms (see Supplemental Fig. 4). It is evident from these histograms that for both the TRE

(U63)- and TRE (U36)-labeled enzymes, there are short-lived binding events that are not represented in the total FRET histograms (black arrows). In the case of the TRE (U63)-labeled enzyme and primer B, the number and duration of the high FRET binding events is substantially increased, consistent with the observation of a large number of traces exhibiting a stable high FRET behavior (see Supplemental Fig. 7). Importantly, we observe the same trends in the mean FRET histograms as we observe in the histograms of total events (see Supplemental Fig. 4), as indicated by the positions of the distributions relative to the yellow dashed lines.

Supplemental Figure 8



Supplemental Fig. 8. Alternative hairpin model for template progression through the active site. a, Diagram of potential hairpin that might form in the TRE. Mutations were generated that would be predicted to either stabilize (blue) or destabilize (red) the putative hairpin. **b,** Telomerase activity assays for telomerase RNPs containing each indicated mutant. Asterisks indicate the first nucleotide added to each primer. RC, recovery control. **c,** Quantitation of the gel in panel **b.** Error bars are the standard deviation over six primers (n=6). Processivity is represented as $R_{1/2}$, a number that summarizes the percent of primers that have been extended by the telomerase RNP past a given repeat[34].

References

1. Lingner, J. et al. Reverse transcriptase motifs in the catalytic subunit of telomerase. *Science* **276**, 561-7 (1997).

2. Licht, J.D. & Collins, K. Telomerase RNA function in recombinant Tetrahymena telomerase. *Genes & development* **13**, 1116-25 (1999).

3. Autexier, C. & Greider, C.W. Boundary elements of the Tetrahymena telomerase RNA template and alignment domains. *Genes & development* **9**, 2227-39 (1995).

4. Gilley, D. & Blackburn, E.H. Specific RNA residue interactions required for enzymatic functions of Tetrahymena telomerase. *Molecular and cellular biology* **16**, 66-75 (1996).

5. Tzfati, Y., Knight, Z., Roy, J. & Blackburn, E.H. A novel pseudoknot element is essential for the action of a yeast telomerase. *Genes & development* **17**, 1779-88 (2003).

6. Theimer, C.A., Blois, C.A. & Feigon, J. Structure of the human telomerase RNA pseudoknot reveals conserved tertiary interactions essential for function. *Molecular cell* **17**, 671-82 (2005).

7. Zappulla, D.C. & Cech, T.R. Yeast telomerase RNA: a flexible scaffold for protein subunits. *Proceedings of the National Academy of Sciences of the United States of America* **101**, 10024-9 (2004).

8. Shefer, K. et al. A triple helix within a pseudoknot is a conserved and essential element of telomerase RNA. *Molecular and cellular biology* **27**, 2130-43 (2007).

9. 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).

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

11. Chen, J.L., Blasco, M.A. & Greider, C.W. Secondary structure of vertebrate telomerase RNA. *Cell* **100**, 503-14 (2000).

12. ten Dam, E., van Belkum, A. & Pleij, K. A conserved pseudoknot in telomerase RNA. *Nucleic acids research* **19**, 6951 (1991).

13. Chen, J.L. & Greider, C.W. An emerging consensus for telomerase RNA structure. *Proceedings of the National Academy of Sciences of the United States of America* **101**, 14683-4 (2004).

14. Tzfati, Y., Fulton, T.B., Roy, J. & Blackburn, E.H. Template boundary in a yeast telomerase specified by RNA structure. *Science* **288**, 863-7 (2000).

15. Mason, D.X., Goneska, E. & Greider, C.W. Stem-loop IV of tetrahymena telomerase RNA stimulates processivity in trans. *Molecular and cellular biology* **23**, 5606-13 (2003).

16. 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**, 6585-90 (2002).

17. Berman, A.J., Gooding, A.R. & Cech, T.R. Tetrahymena telomerase protein p65 induces conformational changes throughout telomerase RNA (TER) and rescues telomerase reverse transcriptase and TER assembly mutants. *Molecular and cellular biology* **30**, 4965-76 (2010).

18. Lai, C.K., Miller, M.C. & Collins, K. Template boundary definition in Tetrahymena telomerase. *Genes & development* **16**, 415-20 (2002).

19. Stone, M.D. et al. Stepwise protein-mediated RNA folding directs assembly of telomerase ribonucleoprotein. *Nature* **446**, 458-61 (2007).

20. Witkin, K.L. & Collins, K. Holoenzyme proteins required for the physiological assembly and activity of telomerase. *Genes & development* **18**, 1107-18 (2004).

21. O'Connor, C.M., Lai, C.K. & Collins, K. Two purified domains of telomerase reverse transcriptase reconstitute sequence-specific interactions with RNA. *J Biol Chem* **280**, 17533-9 (2005).

22. 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**, 8485-90 (1998).

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

24. Kaushik, N. et al. Biochemical analysis of catalytically crucial aspartate mutants of human immunodeficiency virus type 1 reverse transcriptase. *Biochemistry* **35**, 11536-46 (1996).

25. Patel, P.H. et al. Insights into DNA polymerization mechanisms from structure and function analysis of HIV-1 reverse transcriptase. *Biochemistry* **34**, 5351-63 (1995).

26. Cheetham, G.M., Jeruzalmi, D. & Steitz, T.A. Transcription regulation, initiation, and "DNA scrunching" by T7 RNA polymerase. *Cold Spring Harbor symposia on quantitative biology* **63**, 263-7 (1998).

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

28. Brieba, L.G. & Sousa, R. T7 promoter release mediated by DNA scrunching. *The EMBO journal* **20**, 6826-35 (2001).

29. Zaug, A.J., Podell, E.R. & Cech, T.R. Mutation in TERT separates processivity from anchor-site function. *Nature structural & molecular biology* **15**, 870-2 (2008).

30. Zaug, A.J. & Cech, T.R. Analysis of the structure of Tetrahymena nuclear RNAs in vivo: telomerase RNA, the self-splicing rRNA intron, and U2 snRNA. *RNA* **1**, 363-74 (1995).

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

32. Bryan, T.M., Goodrich, K.J. & Cech, T.R. A mutant of Tetrahymena telomerase reverse transcriptase with increased processivity. *The Journal of biological chemistry* **275**, 24199-207 (2000).

33. 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**, 893-8 (1999).

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

CHAPTER V: Direct observation of DNA dynamics in the *Tetrahymena thermophila* telomerase holoenzyme identifies the function of a conserved telomerase reverse transcriptase domain

Abstract

Telomerase is a ribonucleoprotein enzyme that adds repetitive DNA sequences to the ends of chromosomes and has been implicated in both aging and cancer. Telomerase consists of two main components, a protein called telomerase reverse transcriptase (TERT) and an associated telomerase RNA (TER). The telomerase essential N-terminal (TEN) domain is a conserved component of TERT, which previous studies indicate interacts with the telomerase DNA substrate. Mutations to the TEN domain have been previously shown to severely affect the processivity of the telomerase enzyme, however it was not determined the mechanism by which the TEN domain influences processivity. We have previously developed a smFRET binding assay to directly observe DNA dynamics within the telomerase holoenzyme and we adapted the assay to study the effects of TEN domain mutations on telomerase-DNA dynamics. Our results demonstrate that the DNA bound to telomerase is in a conformational equilibrium between two states: an active conformation and a newly-discovered alternative conformation. The relative

stabilities of the active and alternative states are governed by the degree of potential basepair formation between the DNA substrate and TER, with more basepairs favoring the active state. The active state is further stabilized by the TEN domain itself, as mutations to the TEN domain destabilize the active conformation as measured in our smFRET assay. Using our smFRET results, we have developed a model in which the TEN domain stabilizes the formation of short DNA-RNA duplexes in the active site of the enzyme, altering the equilibrium between the alternative state and the active state in order to improve enzymatic processivity. Predictions of the model were confirmed by *in vitro* telomerase extension assays on DNA primer permutants and a thorough dwell-time analysis of our smFRET results. Our assay demonstrates the power of single-molecule techniques, only by directly observing DNA dynamics in single enzymes were we able to identify this alternative conformation of the enzyme and determine the mechanism by which the TEN domain alters DNA dynamics.

Introduction

Telomerase is a ribonucleoprotein enzyme that maintains the ends of eukaryotic chromosomes, adding a repetitive DNA sequence to the ends of the chromosome to generate protective structures known as telomeres[5]. Telomerase counteracts the loss of telomeric DNA due to the DNA replication mechanism's inability to fully replicate DNA ends, therefore telomerase is required for many rounds of DNA replication and cell division[6]. Mutations to telomerase components

can result in genetic disorders characterized by deterioration of proliferative tissues, such as the heritable diseases dyskeratosis congenita and aplastic anemia[7]. On the other hand, inappropriate telomerase activation helps to confer the ability for cells to divide rapidly and is associated with ~90% of human cancers, making telomerase a promising target for potential cancer therapies[8].

Telomerase consists of two main components, a protein telomerase reverse transcriptase (TERT) and a telomerase RNA (TER). TERT is tightly associated with TER, and functions by repetitively reverse transcribing a short template region of TER into telomeric DNA[9]. This template region basepairs with the DNA primer, forming a DNA-RNA duplex which is recognized by the active site of the enzyme for extension. Telomerase undergoes a catalytic cycle where the DNA is progressively extended, concomitantly extending the length of the duplex DNA-RNA hybrid, undergoing a process known as nucleotide addition processivity (NAP). When the end of the template is reached, the enzyme must denature the DNA-RNA duplex and reposition the telomerase RNA for another round of DNA extension, a process known as repeat addition processivity (RAP)[9].

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)[10]. Previous studies in *Oxytricha aediculatus* have shown that TERT cross-links to a region of the telomeric DNA ~20 nts upstream of the nascent telomeric 3' end[11]. 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 that this is a functional interaction[12]. Recently, a crystal structure of the TEN domain from *Tetrahymena thermophila* was reported[3]. The structure identified a protein domain with a novel fold. Though a co-crystal structure with DNA was not obtained, a series of surface-exposed residues were implicated in both cross-linking activity and telomerase extension activity using mutagenesis experiments. A particularly interesting mutant in this study was to residue Q168, which is highly conserved amino acid in the TEN domain, and a component of the conserved T2 (also known as the GQ) motif of TERT[13]. A mutation to this amino acid had a severe effect on both the cross-linking activity of the enzyme and the telomerase activity of reconstituted telomerase[3]. Subsequently, another cross-linking study identified another cross-link between the DNA and the TEN domain, mapping a contact between residue W187 in the TEN domain and the telomeric DNA primer in a region directly adjacent to the 3' end of the DNA in the active site of the enzyme[14], suggesting multiple points of contact between the TEN domain and the telomeric primer.

These experiments resulted in a model of the TEN domain as the site of an "anchor site" interaction in the enzyme[3,11,12,14]. This model suggests 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 loss of the primer during this step. Following the formulation of this model, a new mutant was identified in the TEN domain which had a severe effect on RAP without affecting the anchor site

interaction in the protein [15]. This mutation to residue L14 permitted NAP to the first repeat addition band, but could not extend beyond that band[15]. Interestingly, when the L14A mutant was tested in the context of endogenous *Tetrahymena thermophila* telomerase processivity factors, the enzyme retained the ability to undergo repeat addition processivity, however at a severely reduced rate[16]. This suggested that the mutation to L14 did not completely abolish RAP, but significantly reduced the rate of RAP such that it was not observed in the absence of telomeric DNA binding co-factors.

We have developed a single molecule Förster resonance energy binding (smFRET) assay in order to study dynamics in the telomerase enzyme[17,18]. Previously we established the system to study changes in RNA conformation that accompany DNA extension in wild-type *Tetrahymena thermophila* telomerase[18]. In the current study, we used the assay to study TEN domain mutants to establish how this crucial domain influences DNA dynamics within the enzyme. We demonstrate that wild-type telomerase exhibits a conformational change between two distinct states. This conformational change is governed by the extent of duplex formation between the primer and the RNA template, such that primers with 5 basepairs of duplex formation demonstrate clear dynamics between at least two distinct FRET states, however primers with the potential to form 8 basepairs of duplex do not.

We went on to demonstrate that mutations to TEN domain residues L14, Q168, or F178 significantly alter the equilibrium between the two-previously observed conformations. This conformational change was only observed in primers

with shorter DNA-RNA duplexes and no appreciable difference between WT and mutant enzyme was observed with longer primers. Taken together our results indicate that there are two major conformations of the enzyme: an active conformation where the DNA is hybridized to the template RNA in the active site and an inactive conformation in which the DNA remains bound to the enzyme but is positioned away from the active site. Furthermore, our results demonstrate that the TEN domain is responsible for stabilizing the active conformation of the enzyme for short primers where base pairing stability is insufficient. Both *in vitro* telomerase extension assays and hidden-Markov modeling-based dwell time analyses confirmed predictions based on the details of our proposed model. We hypothesize that this alternative conformation represents an important intermediate in the telomerase catalytic cycle during RAP, permitting the enzyme to remain associated with the primer during DNA-RNA duplex denaturation and that the enzyme tightly regulates the relative affinities of the alternative and active states.

Results

In order to understand the mechanism by which the TEN domain functions in the telomerase holoenzyme, we turned to a smFRET binding assay in order to monitor DNA dynamics within the enzyme[17,18]. Specifically we were interested in how known TEN domain mutants might alter the movement of DNA within the holoenyzme in such a way that could explain the effects of those mutants on telomerase activity. In order to monitor structural changes in telomerase, donor and

acceptor dyes were site specifically incorporated into a telomeric DNA primer and telomerase RNA, respectively (Figure 1B,C). TER was labeled with Cy5 at residue U36, a component of the conserved RNA structure element known as the template boundary element. This motif is thought to be fixed in space by protein contacts within the telomerase holoenzyme[18-20]. The labeled TER was then reconstituted with TERT and an assembly protein p65 in rabbit reticulocyte lysate and assembled telomerase was purified by means of a FLAG-tag engineered into the N-terminus of TERT.

The telomeric DNA primer consisted of a $(TG)_8$ repeat followed by a telomeric DNA sequence labeled at the most 5' alignment residue with Cy3 (Figure 1C). $(TG)_8$ repeat containing primers were used in order to prevent the formation of an alternative DNA structure, known as a G-quadruplex, which would prevent the purification of labeled DNA. Importantly, a $(TG)_8$ repeat primer containing the identical Cy3 modification has been used in previous studies and it was demonstrated that these modifications had no effect on the rate of telomerase activity[18,21]. We generated 5 separate DNA primers for our experiments, each extended by one additional nucleotide on their 3' end, representing separate intermediates in the telomerase catalytic cycle. Each primer was capable of forming 1 additional basepair with the telomerase RNA template than the primer before it. The primers were identified by a letter designation, consistent with the nomenclature of previous experiments on primer permutants[18], and a number designation consisting of the maximum number of potential baspairs the primer can form with template telomerase

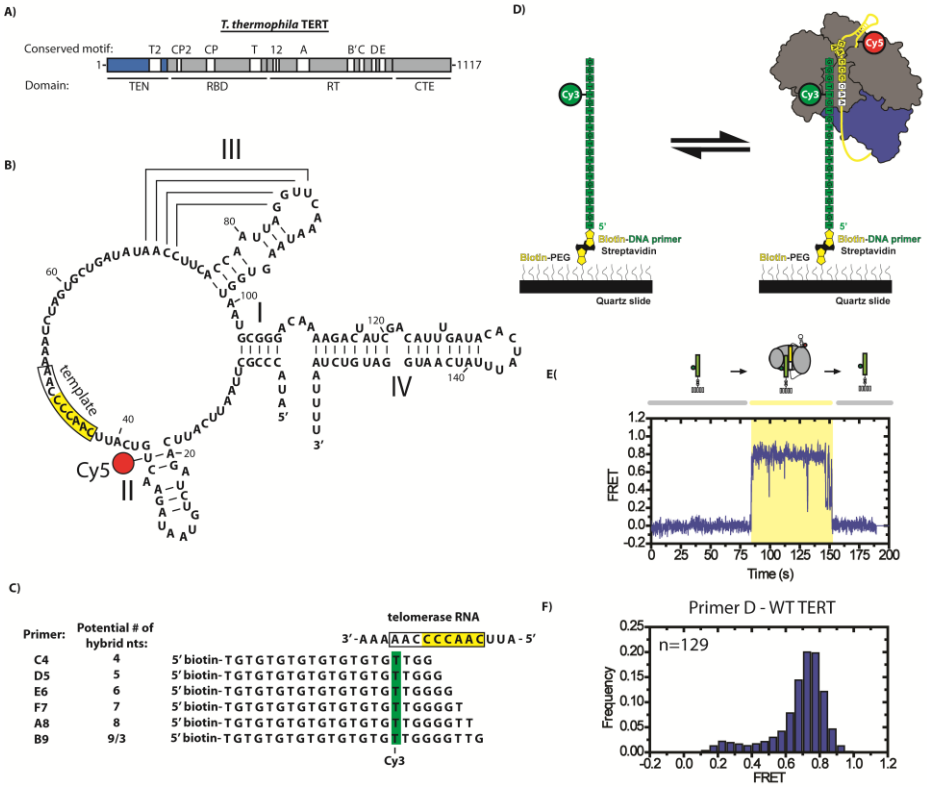


Figure 1. A) Domain map organization of telomerase reverse transcriptase with conserved motifs listed above. The conserved TEN domain is highlighted in blue. **B)** The secondary structure of telomerase RNA. The template is highlighted in yellow with the template alignment residues in the white box. The position of the Cy5 modification on residue U36 is marked by the red circle. **C)** Sequences of primers used in the smFRET assay. Primers contained a 5' biotin modification followed by a $(TG)_8$ sequence and were terminated in a telomeric DNA sequence of varying length. Each primer was named by a letter designation consistent with the nomenclature from previous studies and a number designation based on the number of potential basepairs it can form with the template RNA. The location of the Cy3 modification on the six DNA primers is indicated by the green box. **D)** Schematic diagram of smFRET telomerase binding assay. The DNA primers from Figure 1C were immobilized on the surface of a quartz microscope and telomerase containing a Cy5 modification on U36 of TER was flowed over the slide. FRET was measured on individual molecules for the duration of the binding events. **E)** Example single-molecule FRET traces of Primer D5 incubated with WT telomerase enzyme. Binding is measured as the onset of FRET (shaded area). All FRET values observed during the binding event are used to construct FRET histograms over several separate binding events. **F)** FRET histogram for primer D5 incubated with WT TERT.

RNA. Thus the primers were named D5, E6, F7, A8, and B9. The exact number of basepairs formed by primers A8 and B9 is not clear in the *Tetrahymena* system. Previous experiments from yeast and *Euplotes* telomerase suggest that for late catalytic cycle intermediates there are a maximum number of basepairs formed, such that as new baspairs are added to the 3' end, additional basepairs are released from the 5' end[22,23], however it is not clear if this is present in the *Tetrahymena* enzyme. Nevertheless, for simplicity's sake this nomenclature was used so as to easily identify the order of the primers in the catalytic cycle.

In the telomerase binding assays, Cy3-labeled DNA primers were immobilized on a microscope slide by means of a biotin-streptavidin linkage. Next, our immunopurified telomerase labeled with Cy5 in the TER subunit were flowed on to the slide. The Cy3 and Cy5 intensity of individual molecules was measured during binding events, in order to generate smFRET binding traces. FRET data was also plotted as a histrogram of all the FRET events observed over several different binding events. In our assay, telomerase was incubated with primers in the absence of nucleotide in order to visualize intermediates in the catalytic cycle. In order to test multiple intermediates, Cy5-labeled telomerase was incubated with several different Cy3-labeled primers in separate experiments, consisting of primers D5, E6, F7, A8, and B9 as previously described.

Primers representing intermediates that form 3 and 4 basepairs with the template residues of TER demonstrated very transient binding events in our assay and as a result could not be measured by FRET. Therefore the shortest primer observed in

our assay was primer D5. As previously observed, when this primer was incubated with Cy5-labeled TER, two clear FRET states were observed at ~0.75 FRET and ~0.25 FRET (Figure 2A)[18]. A third FRET state at ~0.5 FRET is also observed, though this state could represent a time-average of a rapid equilibrium between the previous two FRET states that was not resolved at our 100 ms time resolution.

We next tested telomerase harboring an L14A mutation in the TEN domain of TERT in our smFRET assay. This mutation was previously discovered as a mutation that severely perturbed the rate of RAP in *in* vitro telomerase extension assays[15,16]. When telomerase harboring an L14A mutation in the TEN domain of TERT was tested in an identical experiment, the equilibrium between the 0.75 FRET state and the 0.25 FRET state was significantly altered (Figure 2B). The FRET state at ~0.25 FRET now became the major population and the FRET state at ~0.75 FRET became the minor population. As the L14A mutant was shown to decrease telomerase activity and the primary effect of the L14A mutant in our assay was to destabilize the 0.75 FRET state, we conclude that the 0.75 FRET state likely represents the active state of the enzyme and the 0.25 FRET state represents an inactive alternative conformation. This is corroborated by the continuity that the 0.75 FRET state displays with subsequent intermediates in the catalytic cycle that have been observed previously and is consistent with prediction of the FRET changes that are expected as the primer is extended out of the enzyme active site[18].

We next tested the same primer in assays containing mutations Q168A and F178A in the TEN domain of TERT. These two mutations were previously

determined to reduce the rate of RAP in telomerase activity assays, although to a lesser extent than was observed with L14A mutants[3,15]. These mutations were also found to reduce the amount of cross-linking to the 5' end of the telomeric primer in primer binding assays[3]. When Q168A and F178A mutants were tested in our assay, they also demonstrated a destabilization of the 0.75 FRET state. However, the effect of these mutations was less pronounced than was observed with the L14A TERT mutant (Figure 2C&D). As Q168A and F178A mutants were shown to not have a severe defect in *in vitro* telomerase assays, it appears the degree of destabilization of the 0.75 FRET state correlates with the extent of the decrease in activity observed in telomerase activity assays.

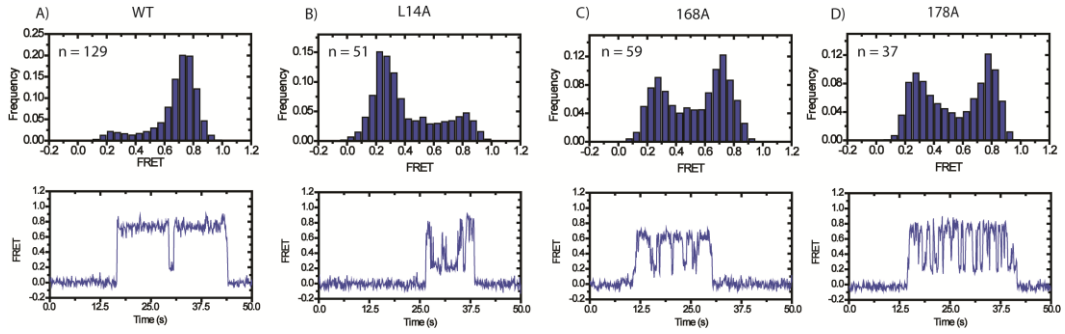


Figure 2: A) Primer D5 (Fig. 1C) was measured using the smFRET telomerase assay with WT telomerase. (Top panel) A bulk histogram of all FRET values observed during all binding events. (Bottom panel) A representative smFRET trace, demonstrating dynamics between a predominant 0.75 FRET state and an alternative 0.25 FRET state. **B)** Bulk histogram and representative trace of primer D5 with L14A TERT. The L14A mutant shows a large change in the distribution of the two states, with the alternative 0.25 FRET state now as the more occupied state. **C)** Bulk histogram and representative trace of primer D5 with Q168A TERT. This mutant demonstrates an intermediate shift in distributions, consistent with its less severe activity defect. **D)** Bulk histogram and representative trace of primer D5 with Q178A TERT. This mutant appears similar to the Q168A mutant.

Effect of primer length on telomerase binding events

When WT telomerase is incubated with primers of increasing length, we observe a gradual decrease in the major FRET population from an ~0.75 FRET state to an ~0.50 FRET state, as previously observed[18], representing an expansion of telomerase RNA 5' of the template as an increasing length of telomerase RNA is extruded from the active site of the enzyme. We tested L14A mutant telomerase in order to determine if the observed defect with primers that form a 5 bp DNA-RNA duplex was also observed with primers of increasing length. Interestingly, L14A mutant primers showed only a modest defect in formation of the active FRET state when the primer was only a single nucleotide longer (Figure 3). As the primer was further extended, the L14A mutant increasingly resembled the WT distribution, with no detectable difference in FRET distributions observed in primers A8 and B9. The occupancy of the low FRET state in Q168A and F178A mutants also demonstrated a bias towards shorter primers (Supplemental Figure 1). Our assays demonstrate that L14A, Q168A, and F178A mutants exhibit detectable defects in early catalytic intermediates that contain short DNA-RNA duplexes, however they show no detectable differences to WT distributions with late catalytic intermediates with long DNA-RNA duplexes.

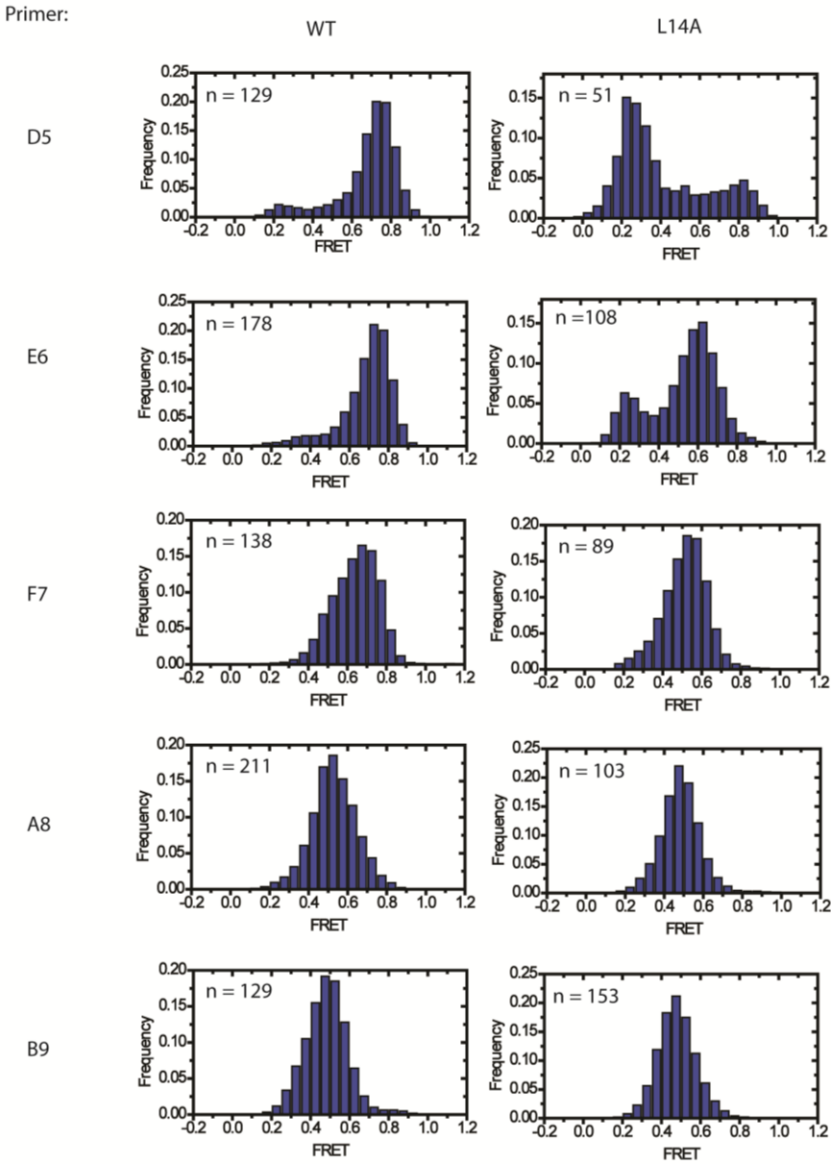


Figure 3. Effect of primer-template hybrid formation on FRET distributions. Primers capable of forming 5-9 basepairs with template RNA were tested in smFRET telomerase binding assay. As in Figure 2, primers D5 and E6(5-6 bp duplexes) showed a marked difference in FRET distributions between WT and L14A enzymes. However, primers F7, A8, and B9 (7-9 bp duplexes) showed very little difference between WT and L14A enzymes. This suggests that L14 participates in stabilizing DNA-RNA duplex formation in the case of short duplexes, but is less important in the case of long duplexes where intrinsic basepairing stability plays a larger role. This was also observed in Q168A and F178A enzymes (Supplemental Figure 1).

A model of TEN domain function

We have developed a model to explain the smFRET data we observed (Figure 4). In this model, DNA bound to telomerase is in a conformational equilibrium between two states. In the active state, the 3' end of the DNA is held in the active site by basepairing contacts with the template RNA and these contacts are reinforced by contacts within the TEN domain. The alternative conformation occurs when the DNA-RNA duplex is transiently broken and the 3' end of the DNA primer moves out of the active site. The rate constants of the transitions between the two states are governed by the duplex stability between the DNA and RNA (the number of basepairs that can be formed) and by the sequence of the TEN domain (the WT sequence stabilizes the active state). For this reason, even WT TERT displays an equilibrium between the active and alternative states, however this is only with shorter primers with less duplex stability. When the TEN domain mutants are present, the active state is destabilized, especially with shorter primers.

When extremely short DNA-RNA duplexes are present, the TEN domain mutants remain in the inactive alternative state longer than the off-rate of the enzyme, such that they dissociate before primer extension can occur. This explains how telomerase accessory factors which bind DNA and reduce the off-rate of the enzyme can partially restore the ability of L14A TERT to undergo RAP[16]. In this context, the L14A telomerase complexes still heavily favor the alternative state in their equilibrium, however they now remain bound to primers longer, permitting some primers to reach the active state and extend before the enzyme falls of the primer.

It is important to note that the distributions between alternative and active conformations of the enzyme were not measured for the two shortest primers, which can form 3 and 4 basepairs with telomerase RNA. Extrapolating from the effects of primer length on the alternative FRET state for 5 and 6 basepair primers suggests that the 3 and 4 basepair forming primers should spend an even greater proportion of their time in the alternative state, even in the context of WT enzyme. Therefore it is likely that the 3 baspair primer that is the immediate result of translocation during RAP may actually spend the majority of its time in the alternative state, even in the absence of TEN domain mutants, forming a barrier that slows down the rate of RAP. The

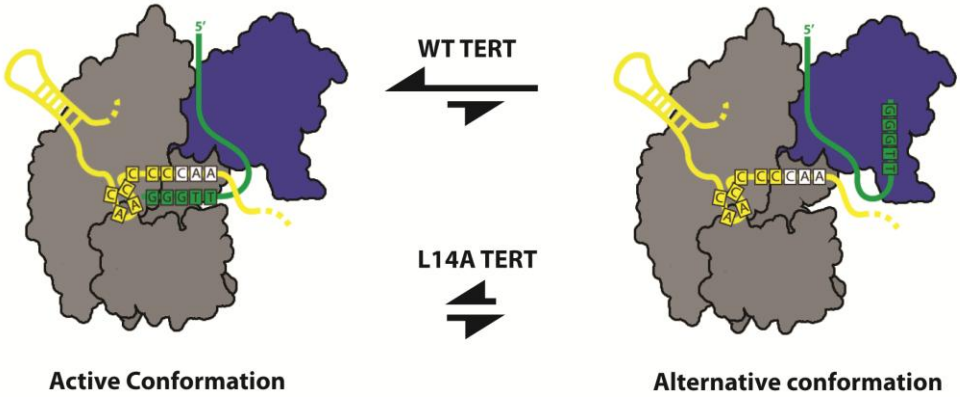


Figure 4. Model of TEN domain function. Telomere DNA bound to telomerase exists in two states: an active conformation (left) and an alternative conformation (right). In the active state the DNA (green) is basepaired to the template RNA (yellow) and positioned in the active site for catalysis. In the alternative conformation the DNA is not basepaired to the template. For primers with short DNA-RNA hybrids, the TEN domain (blue) stabilizes the hybridized conformation. In the case of long DNA-RNA hybrids the intrinsic stability of the duplex is enough to stabilize the hybrid state and mutations to conserved TEN domain residues have less of an effect. The cartoon model is loosely based on a computational model of the human TERT subunit[1] performed with the aid of high-resolution crystal structures from the *T. thermophila* TEN domain and the *T. castaneum* full-length TERT[2,3].

occupancy of this primer in the alternative state would be even further exaggerated by the presence of TEN domain mutants, explaining how these primers could have exchange rates between the active and alternative state slower than the off-rate of the enzyme.

in vitro telomerase extension asssays confirm TEN domain mutants are unable to extend short primers

A prediction of our model is that TEN domain mutants do not affect the translocation of telomerase RNA during RAP. Instead, the mutants perturb the ability of the DNA to bind telomerase in the active conformation for short primers with low DNA-RNA duplex stability. As a result, a prediction of our model would be that TEN domain mutants would have defects in NAP for shorter primers, but would be fully capable of extending longer primers. To test this prediction, we assayed the telomerase extension activity of WT and TEN mutant telomerase on primers with varying DNA-RNA duplex lengths. The primers were length matched so that all primers contained exactly 3 telomeric repeats consisting of 18 total nucleotides, but were staggered so that they formed different lengths of duplex with telomerase RNA (Figure 5A). These particular primers have been used previously in other experiments in telomerase extension assays and given letter designations for each primer (Primers A-F, matching the letter designations in Figure 1). When WT telomerase was incubated with primers A-F, WT enzyme efficiently extended all 6 primers to the first

repeat addition processivity band (asterisks) and was capable of multiple rounds of repeat addition.

As predicted by our model, primers harboring L14A mutations were severely perturbed in NAP for primers with short DNA-RNA duplexes (Figure 5B). Primers B, C, and D were inefficiently extended to even the first repeat addition band (asterisks). However, primers with successively longer potential duplex formation (primers E, F, and A) extended to the first repeat addition processivity band with increasing efficiency, but failed to extend beyond the first repeat, as at this point translocation of the telomerase RNA would generate another short DNA-RNA duplex. Q168A and F178A mutants also displayed a bias in their extension activity due to their DNA-RNA duplex length. While the defect is not as severe in Q168A and F178A mutants as L14A mutants, there is a trend in the deficiency of the ability to reach the first repeat addition band across the primers.

In order to quantify this defect, we compared the intensity of the first repeat addition band between WT and mutant telomerase for each primer. Thus, for primer A the intensity of the first repeat addition band for WT enzyme was compared against the first repeat addition band intensity of L14A, Q168A, and F178A mutants also incubated with primer A. Each mutant's activity was quantified as a percentage of the WT activity observed for the same primer. The results demonstrate a clear trend across primers, with short DNA-RNA duplexes displaying a stronger defect than longer DNA-RNA duplex primers for the three mutants tested. For L14A, telomerase nucleotide addition activity began at 13% of WT for the shortest DNA-RNA duplex

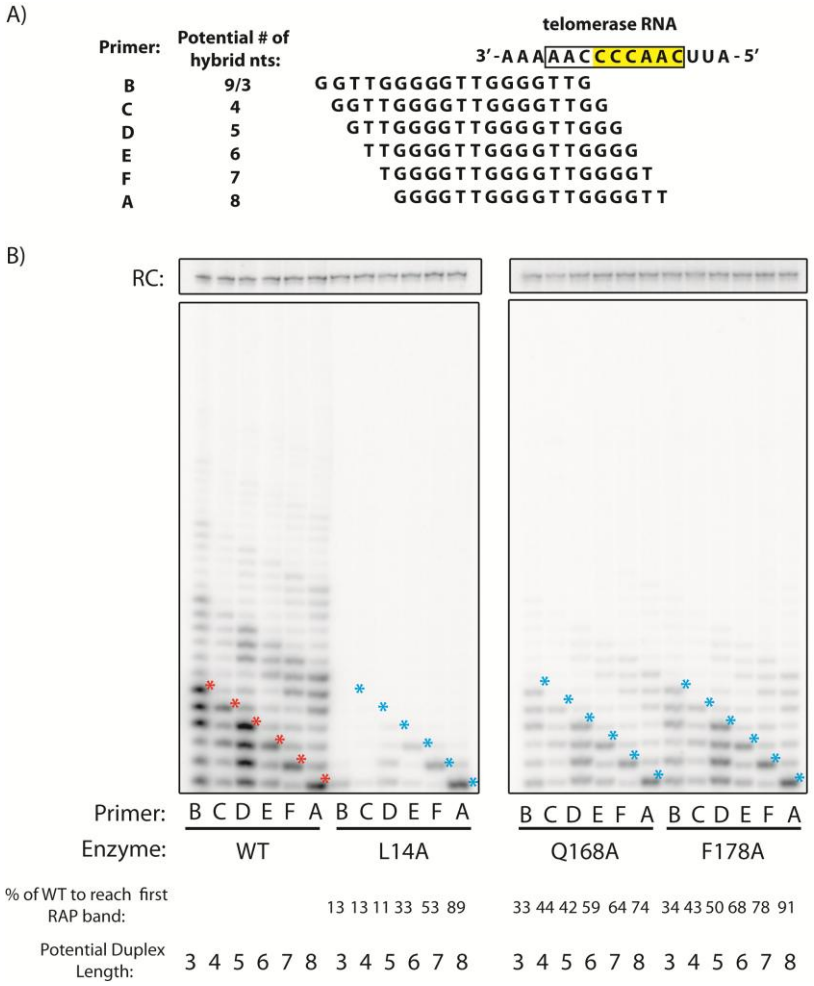


Figure 5. Telomerase activity assays demonstrate TEN domain mutations affect nucleotide addition processivity of short DNA primers. A) Primer permutants used in *in* vitro extension assays. Primers were length-matched at 3 telomeric repeats (18 nts), but staggered such that they formed different initial potential DNA-RNA duplex lengths with template RNA. **B)** Telomerase was reconstituted in rabbit reticulocyte lysate and telomerase activity was assayed on six DNA primers corresponding to six potential DNA-RNA hybrid lengths. WT enzyme was compared against enzyme harboring L14A, Q168A, and F178A mutations. NAP defects for L14A, Q168A, and F178A were measured as the percent of radiolabel incorporation at the first RAP band (blue asterisks) compared against the RAP band for the same primer with WT enzyme (red asterisks). The percent of NAP activity with TEN mutant enzyme shows a clear trend, with shorter DNA-RNA duplex primers demonstrating a severe NAP defect, whereas longer DNA-RNA duplex primers demonstrate very little NAP defect.

primer (primer B) and increased to 89% with the longest DNA-RNA duplex primer (primer A).

Hidden Markov modeling analysis of TEN mutant DNA dynamics reveals that the TEN domain stabilizes the hybridized form of the enzyme

In order to more precisely determine the manner in which TEN domain mutants affect the kinetics of telomerase-DNA motions, we used HaMMy, a program that uses hidden-Markov modeling to identify transitions between FRET states in order to determine dwell times between states[4]. For WT and TEN mutant telomerase incubated with 5-bp hybrid primer D, single molecule traces were analyzed by HaMMy to identify individual dwell times observed in both the active ~0.75 FRET state and the inactive ~0.25 FRET states (Figure 6A). The dwell times were compiled into histograms and the distributions were fit with an exponential function to find the average dwell time for each state (Figure 6B). Each distribution contained at least 100 indvidual dwell time measurements.

For WT telomerase incubated with primer D, HaMMy revealed that as expected the active 0.75 FRET state was considerably more stable than the inactive alternative 0.25 FRET state. The average time spent in the active conformation was 4.95 s while the alternative conformation dwell time was 0.75 s. When L14A mutant telomerase was incubated with primer D, the average dwell time for the active 0.75 FRET state dropped considerably to 0.46 s whereas the dwell time for the alternative 0.25 FRET state remained virtually unchanged at 0.85 s (Figure 6C & D). This

demonstrates that L14 acts primarily to stabilize the active 0.75 FRET conformation, as a mutation to L14 destabilizes the active conformation but has no effect on the alternative conformation. In this way, the kinetic information obtained through dwell time analysis confirms our mechanistic understanding of the role of the TEN domain in telomerase-DNA dynamics.

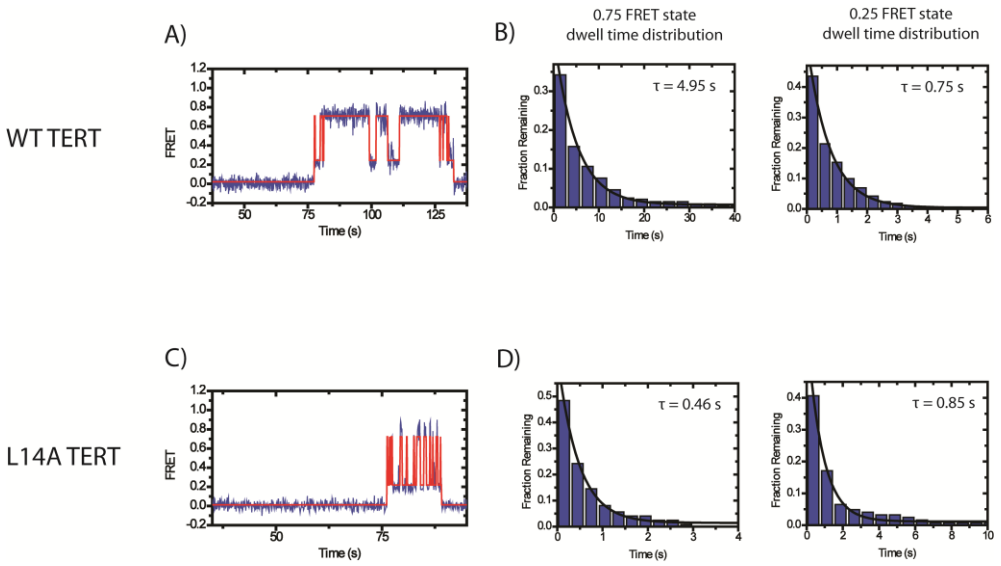


Figure 6. Dwell time analysis of WT and L14A mutants demonstrates that the TEN domain acts by stabilizing the active state of the enzyme. A) smFRET traces (blue) were analyzed by HaMMy[4] to generate idealized traces (red). These were used to determine the dwell time of the enzyme in each state. **B)** All of the dwell times for the 0.75 FRET state and the 0.25 FRET state for WT enzyme incubated with primer D5 were compiled into histograms. The histograms were fit to an exponential function to identify the average dwell time. WT TERT demonstrated a dwell time of τ = 4.95 s for the 0.75 FRET state and a dwell time of τ =0.75 s for the 0.25 FRET state. **C)** Representative smFRET trace and idealized HaMMy trace for L14A TERT telomerase incubated with primer D5. **D)** Compiled histograms for L14A enzyme. L14A TERT demonstrated a dwell time of τ = 0.46 s for the 0.75 FRET state and a dwell time of τ =0.85 s for the 0.25 FRET state. The L14A mutation decreases the dwell time in the active 0.75 FRET state by over ten-fold, whereas the mutation has very little effect on the dwell time in the alternative 0.25 FRET conformation. This indicates that the natural role of L14 in the enzyme must be to stabilize the active state.

Single-molecule FRET experiments using an alternative RNA labeling site corroborate the previous single-molecule results

Previous experiments have demonstrated that dye modifications to residue U63 of TER, a component of the functionally-imporatant template recognition element[24] are well-tolerated by the enzyme[17,18]. Furthermore, smFRET experiments using this RNA construct suggest that U63 is relatively far away from U36 in three dimensional space, on the opposite side of the RNA template[18]. As the DNA primer is extended during the first round of NAP, the primer becomes progressively farther from U36 and closer to U63 (Supplemental Figure 2A)[18]. When U63-labeled TER was incubated with Cy3-labeled primer D in the context of WT TERT, single-molecule traces appeared as in Supplemental Figure 2B. These traces demonstrated very similar kinetics to traces observed with U36-labeled RNA, however in this case the predominant FRET distribution was at ~0.5 FRET and the alternative FRET state was at ~0.9 FRET. Histograms of the smFRET distributions confirmed the relative abundance of the two populations closely mirror the abundance of the 0.75 FRET and 0.25 FRET states observed with U36-labeled enzyme (Supplemental Figure 2B). The fact that the alternative conformation shows ~0.9 FRET with the U63 labeling site and ~0.3 FRET with the U36 labeling site again confirms that U63 and U36 are far in space, such that as the Cy3 label on the primer move farther away from U36 it moves closer to U63 and vice versa.

When L14A mutant U63-labeled enzyme was incubated with primer D, the FRET distribution shifted considerably (Supplemental Figure 2C), again consistent

with the interpretation that the ~0.5 FRET state represents the hybridized state and the ~0.9 FRET state represents the alternative state. Furthermore, the relative populations of the two FRET states again shows a strict dependence on the potential for hybrid formation between the DNA and the RNA template, as observed for U36 labeled enzyme (Supplemental Figure 3). Short primer-RNA duplexes favored the alternative conformation for L14A mutant telomerase, however longer primers showed no obvious L14A defect. Finally, HaMMy analysis of U63-labeled WT and L14A mutant telomerase demonstrates a similar dwell time distribution between the active and alternative conformations observed with the U63-labeled and U36 labeled-enzyme (Supplemental Figure 4A-D). The HaMMy analysis demonstrates the L14A mutation disrupts the active conformation considerably, but has no effect on the alternative conformation, signifying that the L14A mutation works through disrupting the active hybridized conformation and not through stabilizing the alternative conformation.

Discussion

Previous experiments on the TEN domain of TERT had established the TEN domain as an important site of DNA interaction and identified TEN domain mutants that severely affect the rate of RAP[3,11,12,15,16]. After using a smFRET assay to demonstrate that DNA bound to telomerase is in a conformational equilibrium between two states—an active conformation and an inactive conformation—we began to study the effects of TEN domain mutants on this equilibrium. Using

mutational studies we showed that the TEN domain acts to stabilize the active conformation of the DNA in the enzyme. Furthermore, we demonstrated that the requirement for this TEN domain stabilization is governed by the length of the duplex formed between the DNA and the template RNA, suggesting that the role of the TEN domain in maintaining the active state of the enzyme is in fact to stabilize short DNA-RNA duplexes. We propose a model to this effect, where the TEN domains stabilizes short DNA-RNA duplexes in the active site of the enzyme, preventing the dissociation of the DNA-RNA duplex and the concomitant formation of the inactive alternative state (Figure 4). This model is confirmed by HaMMy analysis of dwell time distributions in each state, demonstrating that the TEN domain mutant L14A perturbs the dwell times in the active state while having a minimal effect on the alternative state. This model is also supported by *in vitro* telomerase extension assays, which demonstrate that TEN mutant telomerase has defects in extending short DNA-RNA duplex primers but has progressively less of a defect for primers with longer DNA-RNA duplexes.

Placing the alternative state into the context of the telomerase catalytic cycle

The fact that the alternative state occupies a discrete, well-defined FRET state suggests that it is the result of a separate DNA binding site in the enzyme, and is not the result of unstructured DNA becoming released from the active site of the enzyme. If the alternative state indeed is defined by a discrete DNA binding mode with specific enzymatic contacts, it is likely an important intermediate in enzymatic

function. One potential explanation for an alternative DNA binding site for the enzyme when it is not engaged in a DNA-RNA duplex is as follows. Telomerase must remain associated with primers with short DNA-RNA duplexes since their basepairing stability is insufficient to hold the primer in the enzyme. One way to maintain contact with the DNA primer is to possess an alternative DNA binding site. However, because the telomeric DNA sequence is highly repetitive, a site of strong DNA interaction would compete too strongly with the DNA-RNA duplex formation. Therefore telomerase could maintain its interaction with DNA by means of a weak DNA interaction site in the enzyme, whose binding affinity is specifically tuned so as to be strong enough to prevent loss of DNA upon DNA-RNA duplex dissociation but to not be too strong as to compete against DNA-RNA duplex formation. The enzyme could further tune this equilibrium by containing residues that directly stabilize the hybridized state of the enzyme, such as the TEN domain residues L14, Q168, and F178, the activites of which are characterized in this study.

In this model (Supplemental Figure 5), the alternative state represents an important intermediate in the telomerase catalytic cycle. Following RNA translocation, when the telomeric DNA can only form a 3 basepair interaction with the telomerase RNA template, the enzyme shifts into the alternative conformation in order to maintain contact with the DNA. This alternative conformation is in a dynamic equilibrium with the active state of the enzyme where the DNA-RNA duplex is formed. Once the duplex DNA-RNA hybrid engages stably in the active site,

nucleotide addition processivity can occur and elongate the telomeric DNA, stabilizing DNA-RNA hybrid formation.

Application to other telomerase systems

The conservation of the TEN domain across species—including between *Tetrahymena*, *S. cerevisae*, and human telomerases—suggests that the role of the TEN domain in stabilizing short DNA-RNA hybrids may be conserved across these organisms. A particularly strong argument for this point is the incredible degree of conservation observed between species for the residue Q168. This residue is found in a region of high conservation found in both yeast and human systems 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[25]. 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 DNA-RNA 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 similar repeat addition defect as the L14A mutation in *T. thermophila*[15]. 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 DNA-RNA 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.

The mechanism of TEN domain stabilization of DNA-RNA duplexes

While our smFRET results demonstrate a clear equilibrium between the active and alternative states, the mechanistic 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 DNA-RNA 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[3]. 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[15]. In the crystal structure, L14 is surface exposed and makes interactions with several other hydrophobic side chains near the surface of the domain[3]. L14 therefore may be important in protein-protein interactions, which may be important in order to organize an adjacent region of TEN within the domain, or L14 could interact with another domain of TERT and be responsible for the

positioning and/or dynamics of the TEN domain within the context of full-length TERT.

Another open mechanistic question with regards to the alternative state-active state equilibrium concerns the structure of the alternative state. There are two possible models by which the alternative state could be formed. In the first model, within the active state an anchor site interaction is formed at the 5' end of the primer and base-pairing interactions are formed at the 3' end of the primer in order to hold the 3' end in the active site. When this baspairing interaction is broken, the 5' anchor site interaction is maintained and the 3' end of the primer is released to form a new interaction with the enzyme, forming the alternative state. In a competing model, the alternative state could be formed by simultaneously breaking the DNA-RNA duplex interactions and the anchor site interactions and shifting the DNA one register over, such that the 3' end of the DNA now occupies the anchor site contact. At the current moment, our smFRET data cannot distinguish between these two possibilities. A further possibility is that large-scale TEN domain motions may accompany either RNA translocation or formation of the alternative state. It has been previously theorized that the TEN domain is a mobile element in the telomerase holoenzyme[15] due to its connection to the rest of TERT by an unstructured linker.

Using our smFRET assay, we have conclusively demonstrated the existence of two states in the telomerase holoenzyme—an active state and an alternative state—and have demonstrated that the TEN domain influences the equilibrium between these two states. We believe strongly that future experiments investigating the

structural and mechanistic details of these two states will be valuable to our understanding of how this important enzyme functions.

Materials and Methnods

Dye-labeling of synthetic oligonucleotides

Dye-labeling of synthetic DNA and RNA fragments was performed as previously described[18,26]. 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[18]. 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[27]. 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 analyzes 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[28]. Instead, we approximated γ as $(I_{D1} + I_{A1})/(_{ID2} + I_{A2})$, 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. The factor γ was determined individually for each trace and was consistent with previously reported values of protein-induced Cy3 enhancement

Telomerase extension assays

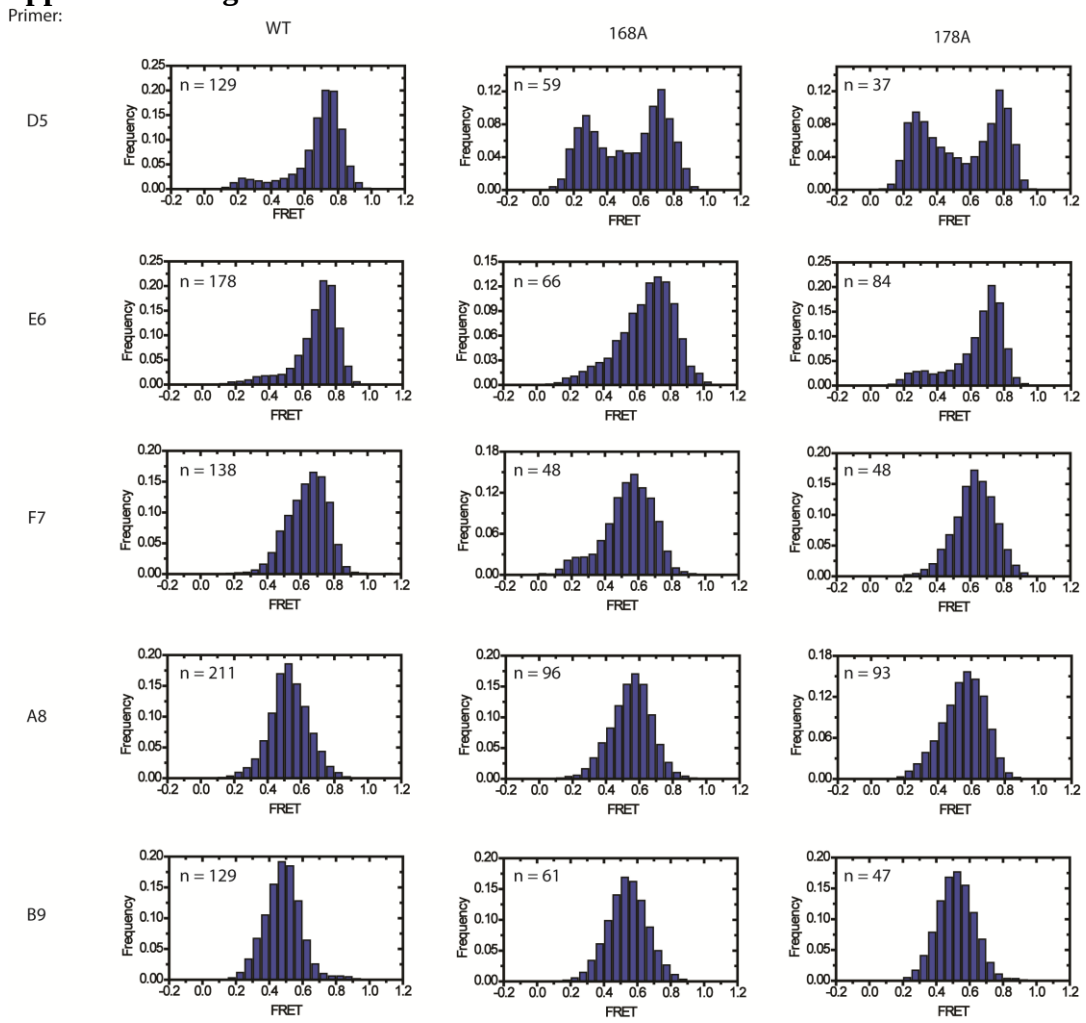
Telomerase for *in vitro* telomerase extension assays was prepared in RRL as described above, however instead of dye-labeled telomerase RNA, *in vitro* transcribed telomerase RNA was used instead. RRL reactions were not immunopurified, but were used directly in telomerase extension assays. 5 µL RRL reaction 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 consisted 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). Quantification of telomerase extension bands was performed using ImageQuant (GE lifesciences).

HaMMy analysis

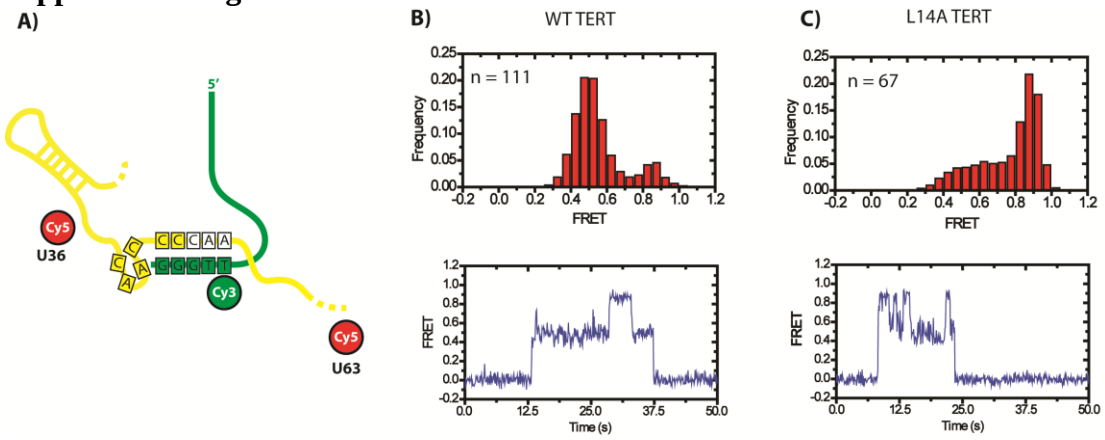
Individual single-molecule traces were analyzed by HaMMy[4]. HaMMy was instructed to identify 3 states for each trace, consisting of the 0.0 FRET unbound state, the 0.75 FRET active state, and the 0.25 FRET alternative state for U36-labeled enzyme and the 0.0 FRET state, the 0.50 FRET state, and the 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).

Supplemental Figure 1



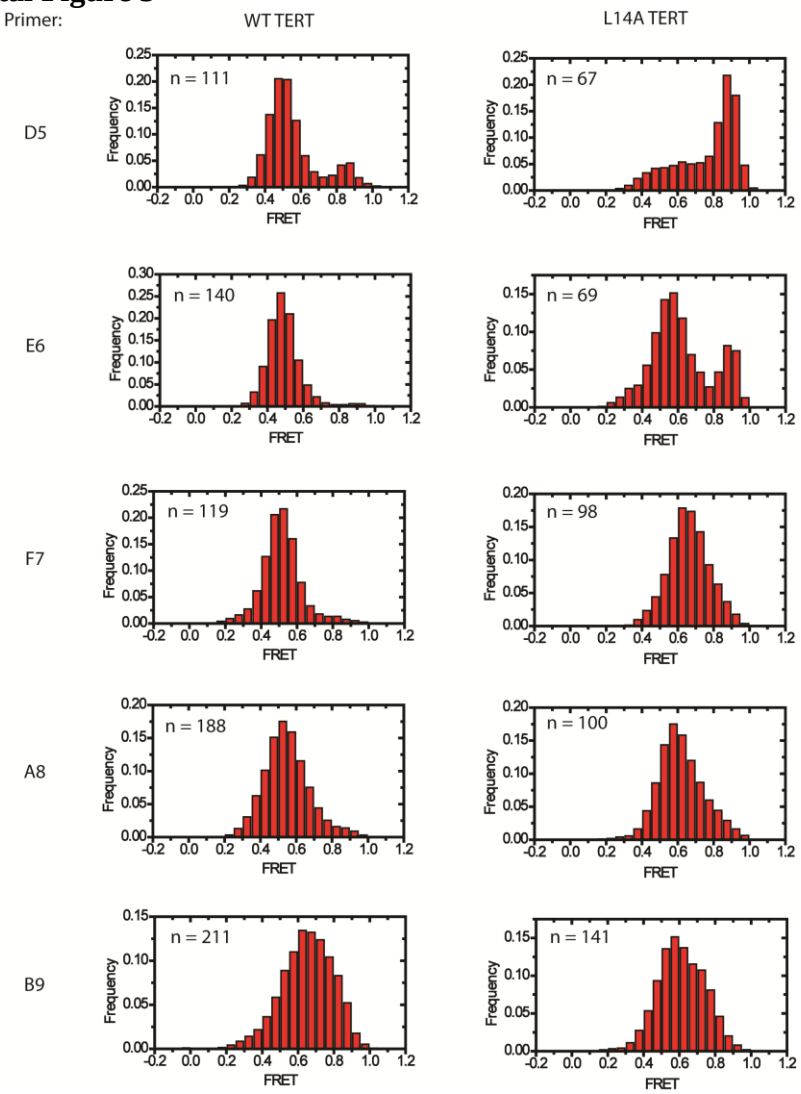
Supplemenatal Figure 1. Effect of primer-template hybrid formation on FRET distributions of Q168A and F178A mutants. Primers capable of forming 5-9 basepairs with template RNA were tested in the smFRET telomerase binding assay. Primer D5 (5 bp duplex) shows a marked difference in FRET distributions between WT, Q168A, and F178A enzymes (top row). Primer E6 (6 bp duplex) demonstrates only a slight shift in distribution between WT and mutant enzymes. Primers F7, A8, and B9 (7-9 bp duplexes) showed very little difference between WT and mutant enzymes. Considered together with Figure 3, our data indicates that TEN domain mutants have the greatest effect on primers that are capable of forming shorter DNA-RNA duplexes.

Supplemental Figure 2



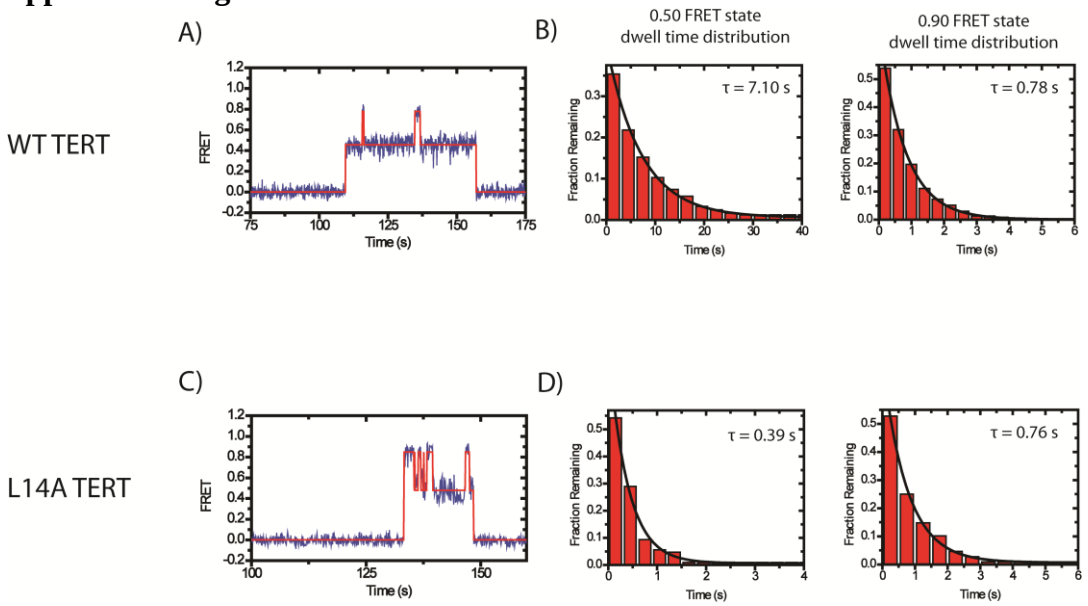
Supplemental Figure 2. Experiments with U63-labeled telomerase RNA confirm the results obtained with U36-labeled TER. A) Schematic diagram demonstrating the relative positions of U36 and U63 within the telomerase holoenzyme on opposite sides of the template. The RNA is in yellow and the DNA is in green. **B)** smFRET histogram (top) and representative trace (bottom) of WT TERT and U63-labeled TER in experiments with primer D5. When U63 is labeled with Cy5, the active state shows a distribution at ~0.5 FRET and the alternative state shows a distribution at ~ 0.9 FRET. **C)** smFRET histogram (top) and representative trace (bottom) of L14A TERT and U63-labeled TER in experiments with primer D5. As with U36-labeled enzyme, the presence of the L14A mutation completely shifts the distribution of primer D5, favoring the alternative state.

Supplemental Figure 3



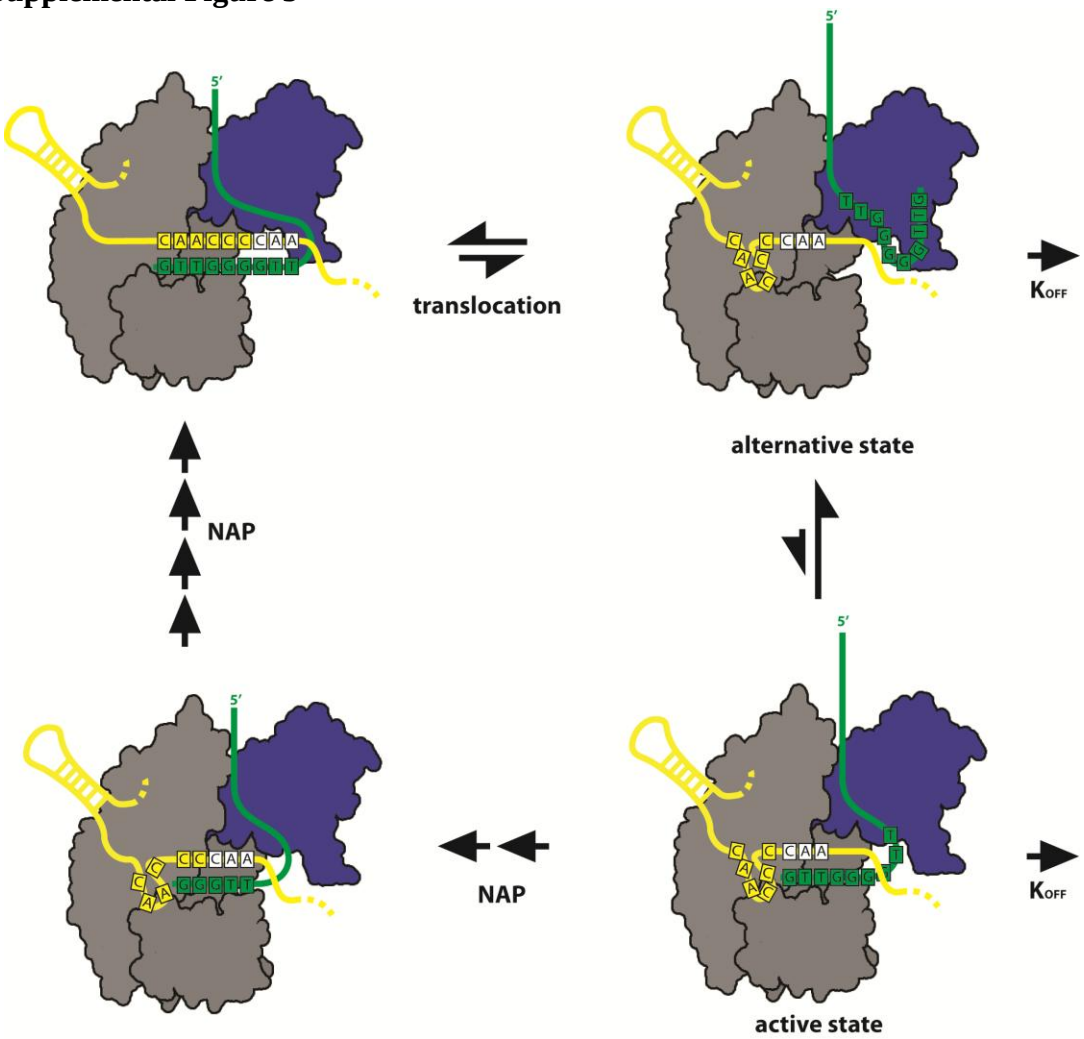
Supplemental Figure 3. Effect of primer-template hybrid formation on FRET distribution of U63-labeled enzyme. Primers capable of forming 5-9 basepairs with template RNA were tested in the smFRET telomerase binding assay with U63-Cy5-labeled TER. Primers D5 and E6 (5-6 bp duplexes) showed a marked difference in FRET distributions between WT and L14A enzymes. These distributions are highly consistent with those obtained with U36-Cy5-labled TER. Primers F7, A8, and B9 (7-9 bp duplexes) showed similar distributions between WT and L14A enzymes. As previously observed[18], primer B9 shows a broad FRET distribution in both WT and L14A enzymes, consistent with two separate states, perhaps a pre-translocation and a post-translocation state of the enzyme. These two states are not evident in the U36-Cy5-labeled enzyme suggesting that these two distinct states may have overlapping FRET values when tested with the U36 labeling site.

Supplemental Figure 4



Supplemental Figure 4. Dwell time analysis of WT and L14A mutants using U63-Cy5-labeled enzyme. A) smFRET traces (blue) were analyzed by HaMMy[4] to generate idealized traces (red). These were used to determine the dwell time of the enzyme in each state. **B)** All of the dwell times for the 0.50 FRET state and the 0.90 FRET state for WT enzyme incubated with primer D5 were compiled into histograms. The histograms were fit to an exponential function to identify the average dwell time. WT TERT demonstrated a dwell time of τ = 7.10 s for the 0.50 FRET state and a dwell time of τ =0.78 s for the 0.90 FRET state. While the dwell time measured for the 0.50 FRET is slightly higher than that was reported for the active state of the U36-labeled enzyme, the numbers are reasonably consistent with the values of 4.95 s and 0.75 s that were observed for the active and alternative states respectively. **C)** Representative smFRET trace and idealized HaMMy trace for L14A TERT telomerase incubated with primer D5. **D)** Compiled histograms for L14A enzyme. L14A TERT demonstrated a dwell time of τ = 0.39 s for the 0.50 FRET state and a dwell time of τ =0.76 s for the 0.90 FRET state. These numbers are highly consistent with those observed with U36-labeled enzyme. As observed in Figure 6, the L14A mutation decreases the dwell time in the active 0.50 FRET state by over ten-fold, whereas the mutation has very little effect on the dwell time in the alternative 0.90 FRET conformation.

Supplemental Figure 5



Supplemental Figure 5. Revised model of the telomerase catalytic cycle. When telomerase reaches the end of the template RNA, the DNA-RNA duplex becomes denatured leading to translocation of the telomerase RNA. As suggested by recent studies in the human enzyme, the pre- and post-translocation states likely exchange in a dynamic equilibrium. The post-translocation state itself exists in an equilibrium between the active and alternative state, favoring the alternative state due to the low duplex stability of the post-translocated enzyme. When the enzyme enters the active state for a sufficient length of time, NAP can occur elongating the DNA-RNA duplex and stabilizing the active state of the enzyme until the end of the template is reached again and the cycle continues. Mutations to the TEN domain affect the equilibrium between the active and alternative state, such that the enzyme transitions to the active state slower than the off-rate of the enzyme.

References

1. Steczkiewicz, K. et al. Human telomerase model shows the role of the TEN domain in advancing the double helix for the next polymerization step. *Proc Natl Acad Sci U S A* **108**, 9443-8 (2011).

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

3. 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**, 218-25 (2006).

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

5. Palm, W. & de Lange, T. How shelterin protects mammalian telomeres. *Annu Rev Genet* **42**, 301-34 (2008).

6. Bodnar, A.G. et al. Extension of life-span by introduction of telomerase into normal human cells. *Science* **279**, 349-52 (1998).

7. Vulliamy, T.J. & Dokal, I. Dyskeratosis congenita: the diverse clinical presentation of mutations in the telomerase complex. *Biochimie* **90**, 122-30 (2008).

8. Kim, N.W. et al. Specific association of human telomerase activity with immortal cells and cancer. *Science* **266**, 2011-5 (1994).

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. Blackburn, E.H. & Collins, K. Telomerase: an RNP enzyme synthesizes DNA. *Cold Spring Harb Perspect Biol* **3**(2010).

11. 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).

12. Lue, N.F. A physical and functional constituent of telomerase anchor site. *J Biol Chem* **280**, 26586-91 (2005).

13. 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. *Mol Cell Biol* **22**, 1253-65 (2002).

14. 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).

15. Zaug, A.J., Podell, E.R. & Cech, T.R. Mutation in TERT separates processivity from anchor-site function. *Nat Struct Mol Biol* **15**, 870-2 (2008).

16. Eckert, B. & Collins, K. Roles of telomerase reverse transcriptase N-terminal domain in assembly and activity of Tetrahymena telomerase holoenzyme. *J Biol Chem* **287**, 12805-14 (2012).

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

18. 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).

19. Akiyama, B.M., Gomez, A. & Stone, M.D. A conserved motif in Tetrahymena thermophila telomerase reverse transcriptase is proximal to the RNA template and is essential for boundary definition. *J Biol Chem* **288**, 22141-9 (2013).

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

21. Collins, K. & Gandhi, L. The reverse transcriptase component of the Tetrahymena telomerase ribonucleoprotein complex. *Proc Natl Acad Sci U S A* **95**, 8485-90 (1998).

22. Forstemann, K. & Lingner, J. Telomerase limits the extent of base pairing between template RNA and telomeric DNA. *EMBO Rep* **6**, 361-6 (2005).

23. Hammond, P.W. & Cech, T.R. Euplotes telomerase: evidence for limited base-pairing during primer elongation and dGTP as an effector of translocation. *Biochemistry* **37**, 5162-72 (1998).

24. Miller, M.C. & Collins, K. Telomerase recognizes its template by using an adjacent RNA motif. *Proc Natl Acad Sci U S A* **99**, 6585-90 (2002).

25. 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).

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

27. Selvin, P.R. & Ha, T. *Single-molecule techniques : a laboratory manual*, vii, 507 p. (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 2008).

28. Ha, T. et al. Single-molecule fluorescence spectroscopy of enzyme conformational dynamics and cleavage mechanism. *Proc Natl Acad Sci U S A* **96**, 893-8 (1999).

CHAPTER VI. Ongoing single-molecule FRET studies probing the role of the telomerase essential N-terminal domain in telomerase activity

Introduction

Previous studies have demonstrated that DNA bound to telomerase is in an equilibrium between an active state and an alternative state and that the telomerase essential N-terminal (TEN) domain stabilizes short DNA-RNA duplexes in the active state of the enzyme (See Chapter V). We have been conducting ongoing studies to identify the mechanism by which the TEN domain stabilizes the active state. To this end we have designed several single-molecule FRET (smFRET) experiments to identify motions in the enzyme that accompany the transition between the active and the alternative state. These smFRET measurements are designed to measure distance changes across the telomeric DNA substrate, between the DNA substrate and the TEN domain, and between the TEN domain and the remainder of the enzyme. In order to accomplish these experiments, we incorporated recently developed biophysical techniques, including three-color FRET and a novel protein-labeling strategy. Initial three-color smFRET experiments have demonstrated that telomeric DNA undergoes correlated stretching motions during the transition to the alternative state. Initial smFRET experiments involving direct fluorophore labeling of the TEN domain demonstrate that the TEN domain and the DNA move relative to one another

during transitions between the active and alternative states. Future smFRET experiments will focus on improving these technologies, gathering more data, and designing new smFRET experiments to study TEN domain motions.

Three-color FRET measurements probing the nature of the telomerase alternative binding mode

Having identified an alternative DNA binding mode for telomerase-DNA interactions, we were interested in identifying further molecular details of this interaction. In particular, we wanted to identify other conformational rearrangements in the telomerase holoenzyme that may accompany the transition to the alternative binding state. One approach would be to identify other conformational rearrangements that occur on the same time scale as the known time scale of the alternative state transitions. These newly identified rearrangements could also be tested for their response to TEN domain mutations. However, the clearest way to demonstrate that any new conformational rearrangement of the enzyme is responsible for the alternative state would be to conclusively demonstrate that this conformational rearrangement happens at the same time as the transition to the alternative state.

Recent experiments from other groups have demonstrated the capability to perform three-color FRET measurements, simultaneously measuring FRET between two separate dye-pairs[1]. In this type of experiment, one has the capability to test if two FRET changes are coupled. Coupled FRET changes between the two dye-pairs would occur simultaneously and be visible in single-molecule FRET traces and two-

dimensional histograms. In order to perform three-color FRET experiments, Cy3, Cy5, and Cy7 fluorophores are used. Using a technique known as alternating laser excitation (ALEX)[2], 532 nm and 650 nm lasers are rapidly alternated. In this setup, FRET is measured for one frame between Cy3 and Cy5 under 532 nm excitation, and in the next frame between Cy5 and Cy7 under 650 nm excitation.

We created new detection optics in our microscope in order to perform three-color experiments. In these optics, an additional dichroic mirror was used to separate light from Cy5 and Cy7 and 3 separate detection channels were created to measure Cy3, Cy5, and Cy7 intensity simultaneously (Figure 1A&B). In addition, a 4-band filter was used to filter out both 532 nm and 650 nm light from the laser excitation. On the software side, custom Matlab scripts were written to identify molecules in 3 separate channels and monitor their intensities. Matlab scripts were also written to separate the raw intensities back into two traces, one for 532 nm excitation, and the other for 650 nm excitation.

Several corrections are necessary for three-color FRET experiments. There are two particular corrections that are the most important. First, Cy7 traces need to be corrected for the lower quantum yield and detection efficiency of Cy7 fluorophores. This correction can be applied by measuring the change in Cy5 and Cy7 intensities upon Cy7 photobleaching. The ratio of these numbers is known as the gamma factor, and represents the ratio of the two quantum yields[3]. The Cy7 intensity is divided by the gamma factor to generate corrected Cy7 traces. This method was previously used in the original description of this three-color FRET method[1].

The second important correction for Cy7 traces is to correct for bleed-through of Cy5 intensity into the Cy7 channel due to incomplete separation by the dichroic mirrors. Characterization of this bleed-through suggests that it is 15% of the Cy5 intensity. This value was subtracted from Cy7 traces.

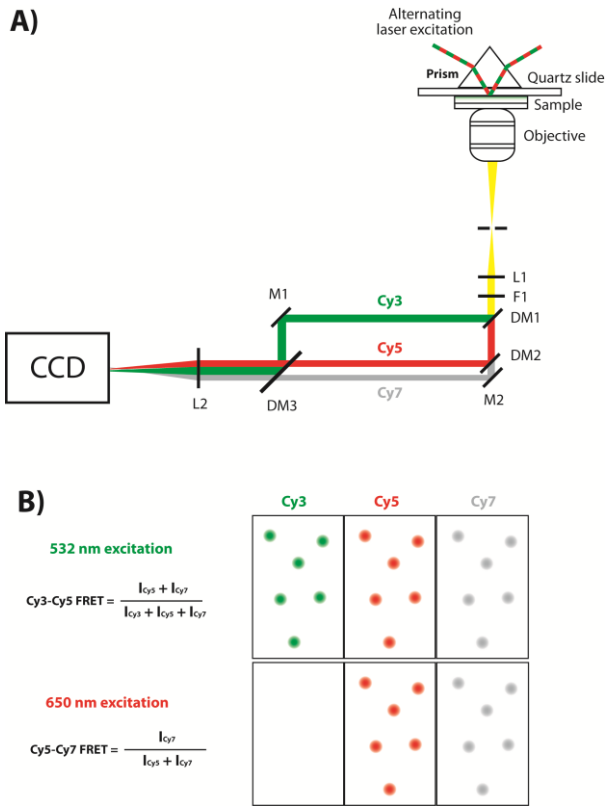


Figure 1: A) Schematic diagram of three-color FRET microscope. ALEX is used to generate alternating pulses of 532 nm and 650 nm laser light. This is used in a prism-type total internal reflection fluorescence (TIRF) microscope. Light from the objective is split with a dichroic mirror (DM1) to separate out the Cy3 light. The remaining light is split with a second dicroic mirror (DM2) to split the Cy5 and Cy7 light. The three separate channels are re-focused with a lens (L2) and imaged using a CCD camera. **B)** The intensities of diffraction-limited spots from the Cy3, Cy5, and Cy7 channels are measured under 532 nm and 650 nm excitation. For a given spot Cy3-Cy5 FRET was calculated as the ratio of Cy5 + Cy7 intensity to total intensity under 532 nm excitation. Cy5-Cy7 FRET was calculated as the ratio of Cy7 intensity to Cy5 + Cy7 intensity under 650 nm excitation.

Three-color FRET experiments measure motions across the DNA primer during telomerase binding

Three-color FRET experiments were designed to measure FRET across the DNA primer during telomerase binding events. In these experiments, a Cy3 dye was placed on a DNA handle immobilized on a slide by a biotin-streptavidin linkage. This handle was annealed to a $(TG)_8$ telomeric DNA primer[4] labeled with Cy5. Cy7 was placed on U36 of TER (Figure 2A). Binding events were measured as the onset of Cy5-Cy7 FRET upon arrival of the Cy7 fluorophore. Three-color FRET traces consist of two traces collected simultaneously: measuring Cy3-Cy5 FRET and Cy5-Cy7 FRET. As the amount of Cy3-Cy5 FRET includes all of the light that passes through Cy5, including light that was relayed to Cy7 by FRET, the Cy3-Cy5 FRET traces were plotted with Cy3 in one trace and the sum of Cy5 and Cy7 in the second trace.

smFRET traces for a representative molecule are shown in Figure 2B. Several interesting observations arise from this trace. First, our Cy3 and Cy5 probes represent labeling sites on either side of the DNA primer, measuring distance changes across the primer during telomerase DNA binding. Before telomerase binding occurs (arrow marks telomerase binding), primers demonstrate a FRET value of ~0.6 FRET. Upon telomerase binding, Cy3-Cy5 FRET values remain essentially unchanged. This demonstrates that telomerase does not significantly alter the length of the DNA upon binding.

Another interesting observation concerns the Cy3-Cy5 FRET measurements across the DNA during transitions between the active and alternative states. Cy5-Cy7

FRET traces demonstrate clear dynamics reproducing the dynamics observed during the transitions between the alternative and active states (Figure 2B, asterisks). Cy3-Cy5 FRET traces show correlated dynamics that occur simultaneously (Figure 2B, asterisks). Interestingly, this demonstrates that as the DNA moves into the alternative state, the DNA enters a further extended conformation. These correlated FRET changes are confirmed by 2D FRET histograms representing all of the traces observed for this experiment (Figure 2C). In these histograms each point represents the frequency at which a single Cy3-Cy5 FRET value occurred simultaneously with a single Cy5-Cy7 FRET value. The 2D histogram indicates a major population centered at 0.62 Cy3-Cy5 FRET and 0.70 Cy5-Cy7 FRET and a minor population with both lower Cy3-Cy5 FRET and Cy5-Cy7 FRET values.

Experiments with L14A and Q168A TEN domain mutants are known to stabilize the alternative state (see Chapter V). In 2D FRET histograms observed with L14A and Q168A mutants, an alternative 2D FRET distribution is clearly apparent (Figure 2C). Consistent with Cy3-Cy5 and Cy5-Cy7 FRET changes being correlated, the alternative state distribution for L14A mutants has both lower Cy3-Cy5 FRET values and lower Cy5-Cy7 FRET values than the active state at 0.50 and 0.44 FRET respectively.

It is important to note that while Cy3-Cy5 FRET changes and Cy5-Cy7 FRET changes are correlated in our experiment, the Cy3-Cy5 FRET changes are less well-resolved than the Cy5-Cy7 FRET changes, suggesting that the distance change between the Cy3 and Cy5 dyes is less pronounced than the distance change between

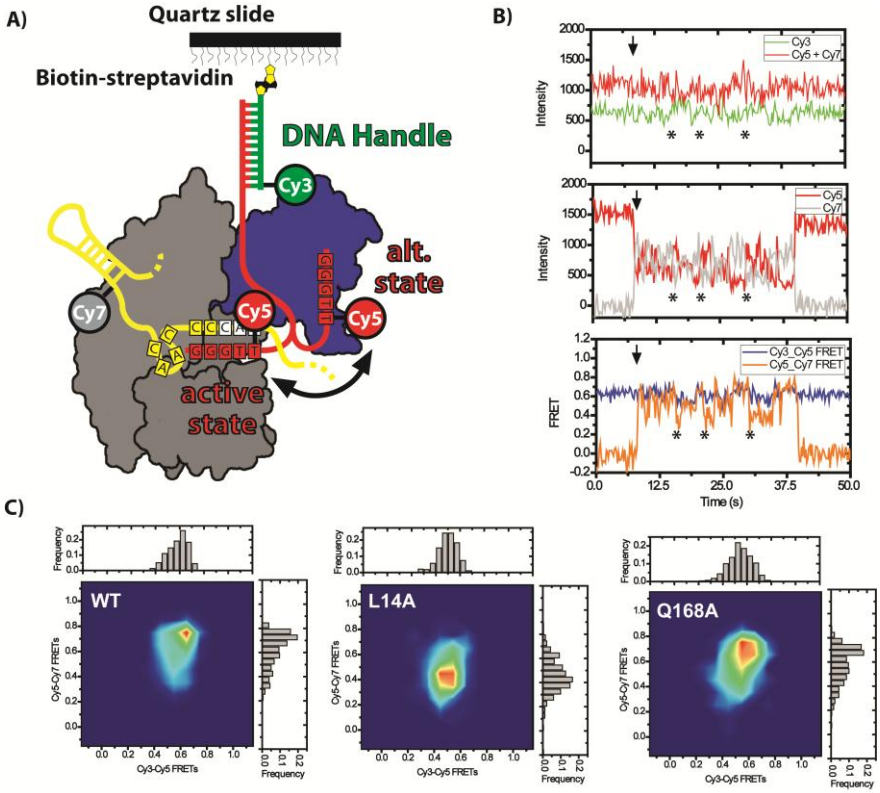


Figure 2: A) Schematic diagram of labeling strategy for three-color FRET experiments. A DNA handle was labeled with Cy3 and immobilized on a quartz microscope slide. This handle was annealed to a telomeric DNA primer corresponding to primer D5 labeled with Cy5. Telomerase was flowed over containing a Cy7 fluorophore label at residue U36 of TER. Changes in FRET between Cy3 and Cy5 and between Cy5 and Cy7 were measured as the DNA transitioned between the active and alternative DNA binding modes. **B)** Representative three-color FRET traces. Cy3 and (Cy5 + Cy7) intensity were measured under 532 nm excitation (top panel). Cy5 and Cy7 intensity were measured under 650 nm excitation (middle panel). Calculated Cy3-Cy5 FRET (blue) and Cy5-Cy7 FRET (orange) were measured over time (bottom panel). The initiation of the binding event is marked with an arrow. Transitions from the active to the alternative state are marked with asterisks. **C)** 2D FRET histograms for WT, L14A, and Q168A enzymes measured with primer D6 in the above three-color FRET experiment. Histograms were measured as the frequency that a given Cy3-Cy5 FRET value (x-axis) was observed at the same time as a given Cy5-Cy7 FRET value (y-axis). 1D histograms of Cy3-Cy5 FRET and Cy5-Cy7 FRET are shown above each axis. 2D FRET histograms suggest correlated changes in FRET as the active state (observed in the WT distribution) has both higher Cy3-Cy5 FRET and Cy5-CY7 FRET values than the alternative state (observed in the L14A distribution). The Q168A distribution shows a mixture of the two states as previously observed (see Chapter V).

the Cy5 and Cy7 dyes. The magnitude of the Cy5-Cy7 FRET change suggests that the Cy5 dye is moving across a relatively large distance in three dimensional space. The fact that that full distance is not reflected in Cy3-Cy5 FRET measurements suggests that the Cy3 fluorophore is positioned in such a way that it is roughly equivalently far from both positions of the Cy5 dye in three dimensional space, though still slightly closer to the Cy5 dye position in the active site.

These correlated FRET changes have implications on the mechanism of the alternative state. If the alternative state simply represented a release of the 3' end of the DNA from the active site, we would expect the DNA to recoil back towards the Cy3 labeling site resulting in higher Cy3-Cy5 FRET values and anti-correlated changes with the Cy5-Cy7 FRET measurements. Instead, we observe correlated changes between Cy3-Cy5 FRET and Cy5-Cy7 FRET. This suggests that the released DNA is specifically stabilized by an alternative binding site within the enzyme and the alternative state of the enzyme represents a discrete alternative binding site for the DNA within the enzyme.

Future three-color FRET experiments

While the three-color experiments have started with promising results, a lot more should be done in order to improve data collection and analysis. Two particular improvements could be made to the system. At the moment, the gamma correction for the differences in Cy5 and Cy7 quantum yields and detection efficiencies is quite large, requiring about a three-fold correction in Cy7 intensity. This is slightly higher

than was previously observed in another system[1]. Initial experiments in our system have suggested that by switching the Cy3 and Cy7 fluorophore positions in our experimental setup we may be able to partially solve this problem as the quantum yield of Cy7 appears slightly higher when the label is placed on the DNA. Experiments with the Cy3 and Cy7 fluorophore positions switched should therefore be tested.

Improvements should also be made to the data analysis of our smFRET traces. At the moment, the custom Matlab script requires several additions. For instance, the ability to use several different ALEX programs for different interval laser pulses should be implemented into the program. Also, the program should be able to detect automatically whether the first frame started on the 532 nm or 650 nm laser pulse. Finally, the program should also incorporate the capability to correct trace intensity based on photobleaching events for each of the three dyes.

There are also some important control experiments that should be performed. In particular, it is important to test for Cy3-Cy7 FRET occurring in our experiments. FRET between these two dye-pairs could complicate our measurements, as it could affect the accuracy of our Cy3-Cy5 FRET measurement under Cy3 excitation. Ideally, the experiment should be designed to maximize the Cy3-Cy7 inter-dye distance so as to minimize Cy3-Cy7 FRET. Since the Cy3-Cy7 dye-pair has a shorter Förster radius than the Cy5-Cy7 dye-pair, and therefore requires shorter distances for FRET, it is likely that our current setup does not have a significant Cy3-Cy7 FRET contribution in our setup. Furthermore, comparing Cy5-Cy7 FRET under 532 nm and

650 nm excitation demonstrates roughly equivalent values, suggesting there is not a significant contribution of Cy3-Cy7 FRET to the Cy7 intensity. Nevertheless, the contribution of Cy3-Cy7 FRET should be fully characterized for these experiments.

Another consideration is the use of DNA handles to tether the telomeric primer to the microscope slides. Initial smFRET experiments demonstrate that the DNA handle hybridized with a telomeric DNA primer can occupy two FRET states, an ~0.6 FRET state and a ~0.3 FRET state (Supplemental Figure 1). However, telomerase binding events were only observed on DNA molecules in the 0.6 FRET state, suggesting that the 0.3 FRET state represents a non-productive conformation of the DNA model. The 0.3 FRET state becomes progressively more occupied the longer the DNA is incubated with telomerase primer (Supplemental Figure 1). Given that the 0.3 FRET state represents a large majority of the DNA molecules after 20 minutes, it appears that FRET experiments utilizing a DNA hnalde should be conducted quickly (within 20 minutes of adding telomerase to the slide) in order to be able to observe as many traces as possible. This effect should be kept in mind for other experiments using telomeric primers immobilized via a DNA handle.

We believe that the three-color FRET system is now an important platform for future experiments utilizing different labeling sites to test different motions within the enzyme. In the future, three-color FRET measurements could be combined with protein-based fluorophore-labeling of TERT in order to monitor conformational rearrangements of the TEN domain and how they may relate to changes between the active and alternative DNA binding modes or how they may relate to RNA

translocation during repeat addition processivity. There are some important design criteria to keep in mind when designing future three-color FRET experiments. The experiment should be designed to maximize the distance between Cy3 and Cy7 while keeping the distance between the Cy3-Cy5 and Cy5-Cy7 FRET pairs near the Förster radii. This requirement definitely limits the dye positions that can be used together and special care should be taken when designing three-color FRET experiments.

Potential studies on telomerase TEN domain motions

Previous work has established that there exist two conformations of DNA bound to the telomerase holoenzyme, an active and an alternative conformation. Furthermore, we have demonstrated that key conserved TEN domain residues regulate the equilibrium between these two states by stabilizing the active conformation. However, several questions remain about the mechanism by which the TEN domain stabilizes the active state of the enzyme. In particular, we are interested in establishing if the TEN domain represents a mobile element of the structure, or if it remains static during telomerase activity. It has long been known that the TEN domain is connected to the rest of TERT by means of a long flexible linker. As a result it has been theorized that the TEN domain is a mobile element in the enzyme[5]. In order to test if TEN domain motions are responsible for transitions between the active and alternative states, we have developed a system for direct labeling of TERT for our single-molecule FRET (smFRET) binding assay.

In this system, our cyanine dyes can be incorporated into the protein using a recently developed labeling strategy[6]. In this technique, a sulfur-reactive dye is first reacted with a coenzyme A (CoA) molecule. Next, a protein known as Sfp phosphopantetheinyl transferase transfers the cyanine dye coupled to the phosphate of the CoA molecule onto a serine in a specific 11-amino-acid sequence incorporated into the protein of interest. This 11-amino-acid tag can be engineered into an unstructured loop region of the protein domain so as to have a minimal effect on the activity of the enzyme. It is important to note that in this strategy, the dye is now on an effectively longer linker, consisting of the 6-carbon linker on the dye, the phosphate bond to the serine, and the 11 amino acid tag. As a result, the dye will now report less accurately on the position of the original labeling site than a conventional FRET probe. Nevertheless, this strategy will likely suffice to report on gross structural changes between different conformations. In this type of experiment, one would be looking for changes in the FRET distribution in the presence or absence of certain mutants and would be less concerned with the absolute position of the dyes involved.

In our initial work, we designed six separate labeling sites within the TEN domain to insert the 11-amino-acid tag, known as a ybbR tag. These were designed into unstructured loops found in the crystal structure of the *Tetrahymena thermophila* TEN domain[7], as well as at the N-terminus of the protein. The ybbR tag was incorporated at residues 1, 26, 48, 82, 102, and 125 of the TEN domain. Telomerase extension assays demonstrated that the ybbR tag did not disrupt of the enzyme for the

majority of labeling sites (including labeling sites, 1, 48, 82, 102, and 125). In fact, only labeling site 26 showed a significant defect in activity upon insertion of the ybbR tag (Figure 3A).

Next, we labeled coenzyme A with a sulfur-reactive Cy5 dye. The Cy5-CoA conjugate was added to telomerase RRL expression reactions containing the constructs for the ybbR-tagged TERT as well as Sfp enzyme and $MgCl_2$. After an hour at room temperature, the complex was immunopurified and tested for activity. RRL reactions were tested by SDS PAGE and fluorescence scanned to confirm that Cy5 was covalently labeled to TERT (FIGURE 3B) and telomerase extension assays confirmed 4 of the 5 remaining labeling sites showed no decrease in activity upon Cy5 labeling (FIGURE 3C).

Immunopurified TERT labeled with Cy5 in the TEN domain was saved for future smFRET experiments. SDS PAGE gels of immunopurified TERT suggests a highly pure product at the approximate molecular weight of TERT of 133 kDa (Figure 3D). Currently, we have conducted two types of measurements in our smFRET assay using dye-labeled protein: FRET measurements between Cy5-labeled protein and Cy3-labeled DNA and FRET measurements between Cy5-labeled protein and Cy3-labeled RNA.

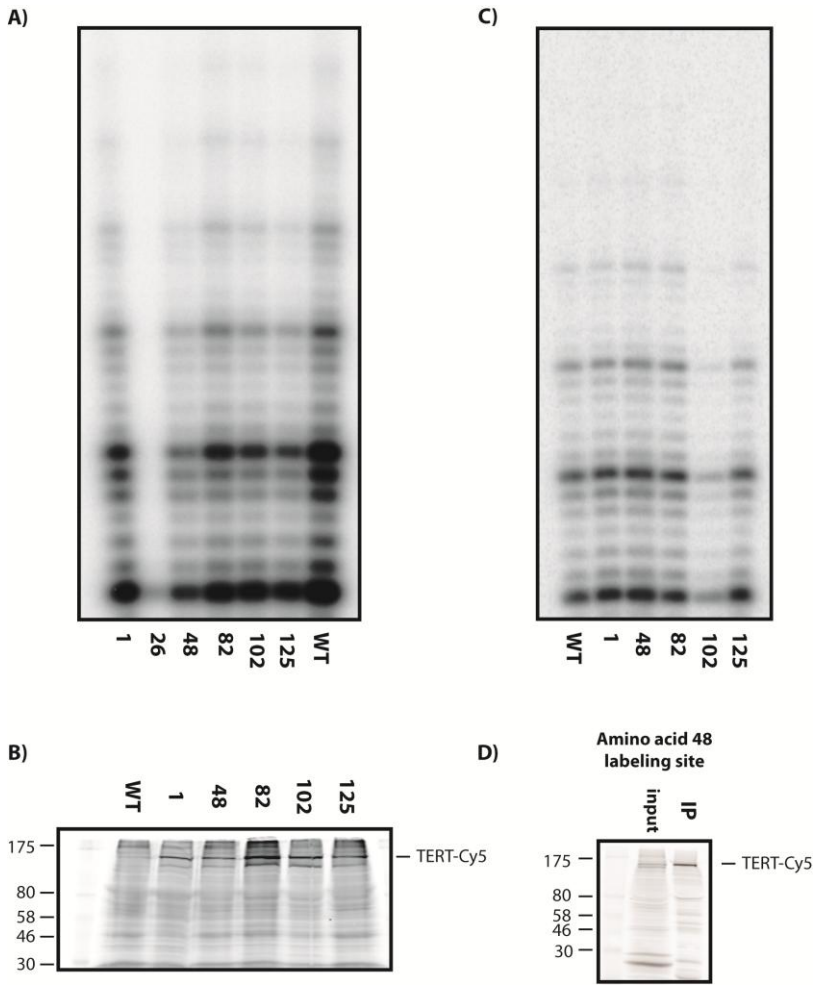


Figure 3: A) Telomerase activity observed in the presence of ybbR tag inserts at residues 1, 26, 48, 82, 102, and 125. Most labeling sites demonstrated high levels of telomerase activity, with the exception of labeling site 26. **B)** A 650 nm laser scan of and SDS-PAGE gel containing rabbit reticulocyte lysate (RRL) with WT (unlabeled) TERT and TERT labeled with Cy5 using a ybbR tag at the indicated TEN domain residue. ybbR-labeled TERT demonstrates a clear band corresponding to the presence of Cy5-labeled TERT. **C)** Telomerase activity assay observed in the presence of Cy5-labeled TERT at the indicated residues using the ybbR tag. Labeling sites 1, 48, 82, and 125 showed robust telomerase extension activity in the presence of both the ybbR tag and the Cy5 fluorophore label. Labeling site 102 demonstrated reduced telomerase activity. **D)** A 650 nm laser scan of an SDS-PAGE gel containing RRL with TERT labeled at labeling site 48 (input) and immunopurified TERT labeled at labeling site 48, demonstrating purified Cy5-TERT.

smFRET measurements between the TEN domain and the DNA primer demonstrate feasibility of protein-labeling studies

For experiments measuring FRET between the DNA primer and the TEN domain, Cy5-labeled TERT was incubated with Cy3-labeled primer which had been immobilized on a quartz microscope slide (Figure 4A). As in previous experiments, the primers used in the experiment were progressively longer on their 3' end to mimic different intermediates in the catalytic cycle. Each primer was given a letter designation consistent with previous nomenclature[4] and a number designation based on the number of potential basepairs between the DNA primer and the template RNA (Figure 4B). For these experiments, primers D5, F7, and B9 were used corresponding to primers that could form 5, 7, and 9 basepairs with telomerase RNA, respectively.

smFRET traces were collected for each of the primer with 3 separate labeling sites on the TEN domain (residues 48, 82, and 125). smFRET traces demonstrated that Cy5-labeled TEN domains were functional in smFRET binding assays (Figure 5A). However, there appear to be fewer traces in our protein-labeled studies than in our RNA-labeled studies. In addition, the average trace duration appears to be shorter, indicating a faster photobleaching rate. These limitations could be overcome in several ways. First, simply collecting more data over longer periods of time should circumvent the problem of adequate data collection Secondly, protein-labeling with Cy3 could be tested, as Cy5 is generally more susceptible to photobleaching in our experience. In any case, for these experiments, it is important to note that special care

should be taken to ensure that results are reproducible as the limited data collection could result in less well-defined FRET distributions.

Our initial smFRET data demonstrated reasonably well-distributed histograms which could be easily fit with either one or two Gaussian distributions. In addition, labeling sites 48, 82, and 125 demonstrated a clear trend across primers as the DNA primer was progressively extended, suggesting that the primer moves towards the TEN domain during enzyme activity, consistent with the prediction of our model (Figures 4C-E). Labeling site 125 appeared to be closer to the primer than sites 48 and 82, demonstrating higher FRET values in our experiments. Interestingly, in the crystal structure of the TEN domain, residue 125 appears to be closer to the critical residues L14, Q168, and F178 than residues 48 and 82[7]. Therefore, the smFRET data is consistent with residues L14, Q168, and F178 making direct contact with the DNA with residues 48 and 82 being farther away from the active site of the enzyme.

smFRET traces for primer D5 with labeling site 48 demonstrated dynamics between a major conformation at ~0.35 FRET and a minor conformation at ~0.55 FRET. The rate constants of these dynamics appear similar to those observed between the active and alternative conformations of the enzyme observed when primer D5 was tested with an RNA labeling site (Figures 5A & B). To confirm that these transitions are indeed related to the transitions between the active and alternative states, we next tested labeling site 48 Cy5-labeled TERT containing an L14A mutation. The distribution for L14A mutant TERT shifted considerably towards the ~0.55 FRET state (Figures 5A & B), suggesting that indeed the 0.35 and 0.55 FRET states

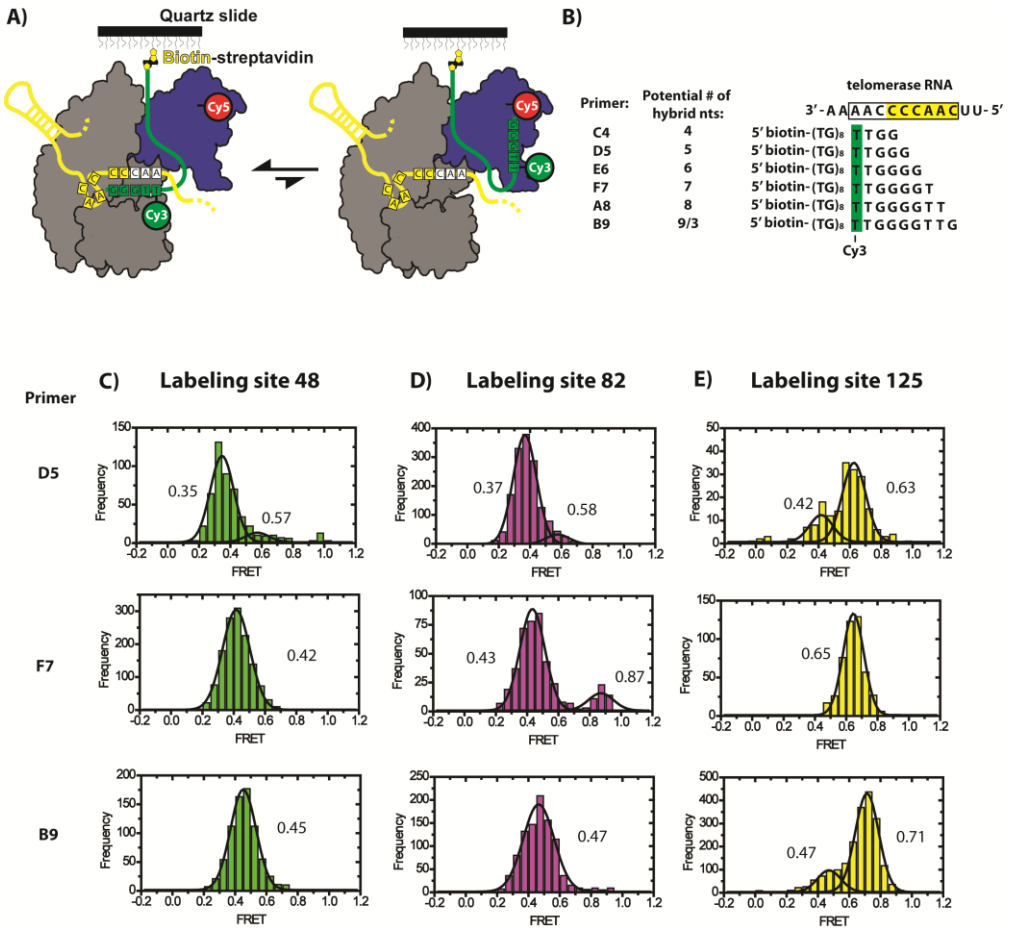


Figure 4: A) Schematic diagram indicating labeling strategy. Cy3 was placed on the DNA primer as in previous experiments and Cy5 was placed on the TEN domain of TERT. FRET was measured between the dyes on immobilized telomerase molecules, measuring FRET between the DNA primer and the TEN domain. **B)** List of DNA primers used in smFRET experiments. Primers consisted of a $(TG)_8$ repeat followed by a telomeric 3' end. Each primer was successively extended on the 3' end by an additional nucleotide. Primers were given a letter designation and a number designation based on the number of potential basepairs they could form with TER. **C)** smFRET histograms measured between the DNA primer and the TEN domain labeled at residue 48. smFRET histogram demonstrate a clear Gaussian distribution with a trend across primers as the primer is successively extended. The distributions were fit to one or two Gaussians with the center of the distribution given. **D)** smFRET histograms of FRET measured between the DNA primer and labeling site 82. These traces also demonstrate a trend across primers. **E)** smFRET distributions observed with FRET measured between the DNA primer and labeling site 125. FRET values in these measurements were significantly higher than observed in the other two TEN domain labeling sites.

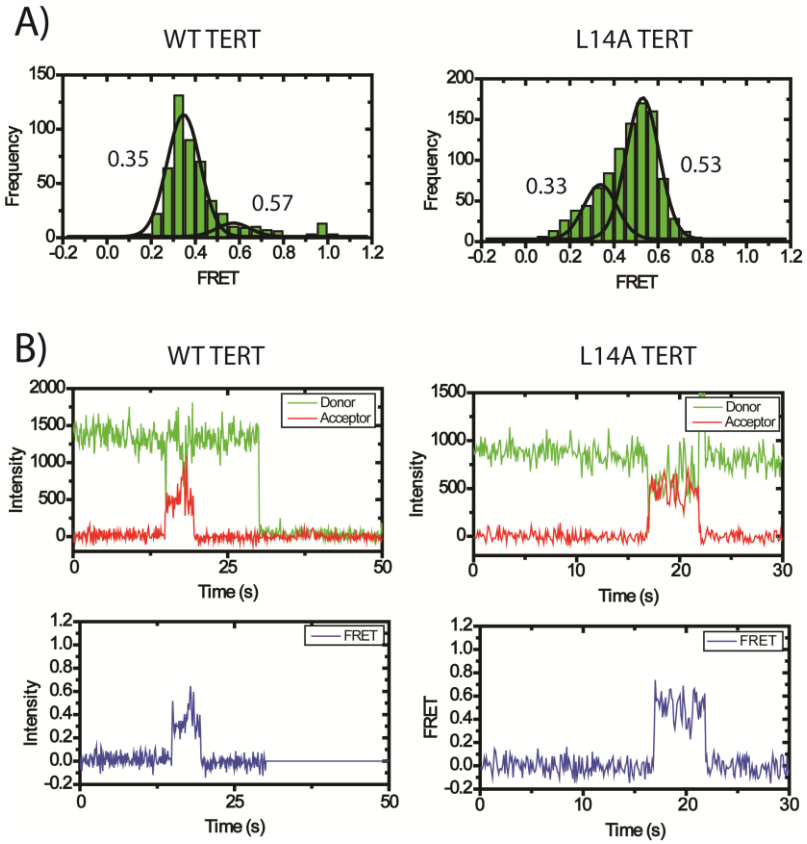


Figure 5: Comparison of WT and L14A TERT in TEN-DNA smFRET assays. **A)** Primer D5 was measured with telomerase containing WT and L14A TERT. In both experiments, FRET was measured between the primer and position 48 in the TEN domain. smFRET histograms suggest that the presence of the L14A mutant significantly shifts the FRET distribution. This suggests the ~0.55 FRET state is the alternative state and the 0.35 FRET state is the active state. **B)** Representative smFRET traces for WT and L14A TERT measurements demonstrating conformational dynamics between the two states.

represent the active and alternative FRET states respectively. This demonstrates that the TEN domain and DNA primer move with respect to one another during the active-alternative state transition. Initially, this would appear to support a static TEN domain during this process, as we know that the DNA moves with respect to the rest of the telomerase complex, however it appears likely that some level of DNA-TEN

domain relative motion would also be present in a model that implicates gross TEN domain motions within the telomerase holoenzyme.

In order to complete studies on TEN domain motions relative to the DNA, several objectives must be met. First, additional data must be collected to confirm that the initial data is clearly reproducible. This is particularly important due to the relavtively short duration of our traces due to Cy5 photobleaching. Secondly, it would be valuable to determine whether switching the dye labels between the protein and the DNA would improve data throughput. Finally, to completely confirm that TEN-DNA relative motions are coupled to the active-alternative state transition, 3-color FRET experiments could be designed, with Cy3 on the TEN domain, Cy5 on the primer, and Cy7 on U36 of the RNA. In this experiment, if our protein-DNA and DNA-RNA relative motions are truly coupled, we should observe that they occur simultaneously in our 3-color FRET setup.

Measuring FRET between the TEN domain and telomerase RNA

In order to distinguish between a moving TEN domain and a static TEN domain, we must measure FRET between the TEN domain and a relatively fixed position within the rest of telomerase. Ideally, this fixed labeling site would be at a distance of 50-60Å from the TEN domain labeling site in order to maximize the sensitivity of our FRET response. RNA represents a natural choice for our fixed labeling site as we know that we can easily label an RNA with a cyanine dye. TER residue U36 represents a logical target for labeling within the telomerase RNA as it is

thought to be a non-mobile element within the enzyme, being fixed by the interaction of the template boundary element region of the RNA with the RNA binding domain of TERT[8,9].

Initial FRET studies on Cy3-labeled U36 tested with Cy5-labeled TEN domain TERT were designed. In these experiments an Alexa488-labeled primer was immobilized on the surface via a biotin-streptavidin interaction (Figure 6A). Telomerase was flowed on to the slide, containing Cy3-labeled TER and Cy5-labeled TERT (Figure 6A). A blue laser was used to identify the location of Alexa488-labeled primers and the illumination was then switched to a green laser to measure Cy3-Cy5 FRET within the enzyme. Single-molecule traces observed by this technique were marked by an initial period of illumination (due to Alexa488 fluorescence) and then were observed as the onset of FRET due to binding of Cy3 and Cy5 containing telomerase. smFRET traces corresponding to telomerase binding events were observed, demonstrating the feasibility of this approach (Figure 6B). As observed before for TEN-Cy5 labeled traces, Cy5 appeared to rapidly photobleach in these experiments.

Using these particular labeling sites, a FRET distribution centered around 0.2 FRET was observed (Figure 6B & C). At this FRET value, our FRET response is unlikely to be sensitive enough to detect subtle distance changes. As a result, we decided to measure FRET between the TEN domain and U63 on telomerase RNA. This labeling site was used in previous experiments to measure FRET changes between RNA and DNA labeling sites. While this labeling site was well-behaved in

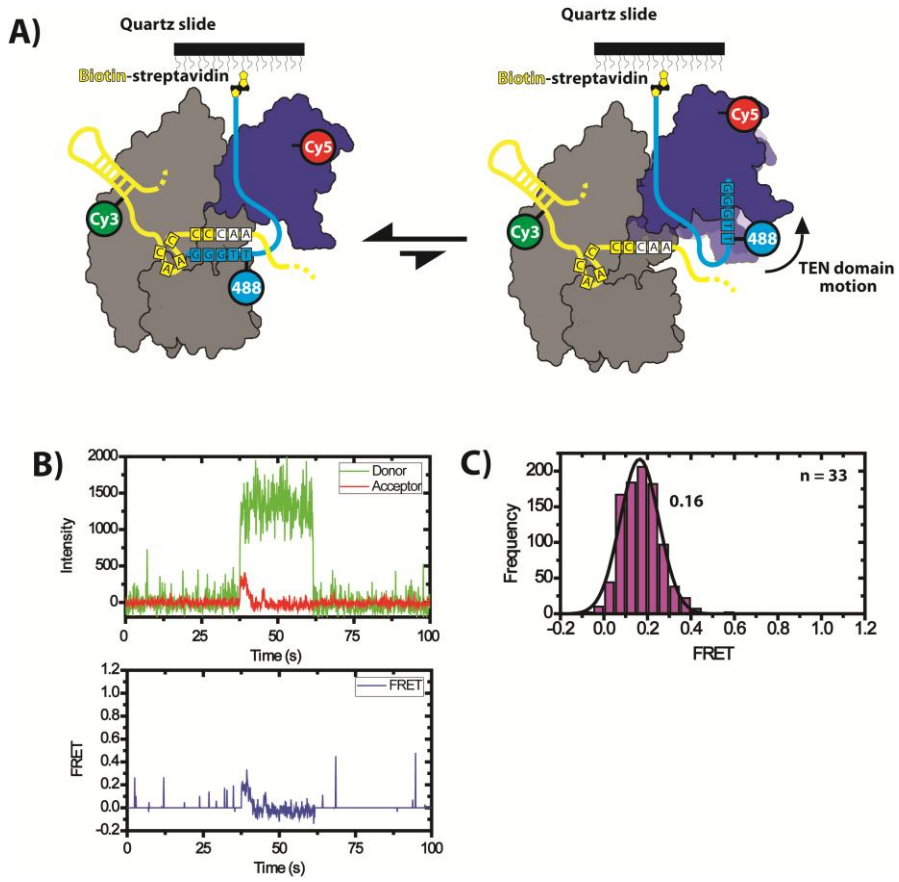


Figure 6: A) Schematic diagram of smFRET TEN domain labeling experiment. Alexa-488 primers were used to identify the location of the primers under blue laser excitation. FRET was measured between Cy3 at TER position U36 and Cy5 on the TEN domain. TEN domain dynamics should manifest themselves as a change in FRET between Cy3 and Cy5. **B)** Representative smFRET trace. Telomerase diffuses onto the primer, indicated by the onset of Cy3 and Cy5 intensity. Cy5 then photobleaches, returning to 0.0 FRET. **C)** smFRET histogram demonstrating the distribution of FRET values observed for 33 separate traces. The distribution was fit to a Gaussian distribution suggesting a 0.16 FRET state.

these studies, it is important to note that there is no biochemical evidence to demonstrate that U63 is a fixed site in the enzyme in the same manner that U36 has been implicated as a fixed site by its known interaction with the RBD[8].

As with U36-labeled RNA tested with TEN-Cy5-labeled TERT, U63-labeled RNA demonstrated smFRET traces (Figures 7A&B). Histograms revealed a broad FRET distribution roughly centered at~0.5 FRET (Figure 7C). It is unclear what phenomenon accounts for the broadness of the distribution. At least part of the broadness appears to be the result of FRET dynamics within smFRET trajectories. However, there are likely additional sources of experimental noise. One source of noise is the use of the Cy3-label on the RNA, causing additional background illumination due to excited dye molecules in solution. An additional source of noise is found in the inability to correct for differences in Cy3 and Cy5 quantum yields by traditional methods. Ordinarily, this correction would be measured upon Cy5 photobleaching, however a large number of traces diffuse off of the surface before Cy5 photobleaching can occur. This limitation could be overcome by only analyzing primers in which Cy5 photobleaching occurs during the observation window.
The transitions between states observed with U63 and TEN domain labeling sites do not appear to be correlated with the rate constants of the transitions between the active and alternative states. It is likely that these transitions measure an alternative source of dynamics, either due to an uncorrelated TEN domain motions, motions within the TEN domain linker to the Cy5 dye, or motions by the U63 labeling site itself.

In order to improve protein-RNA FRET experiments, it is likely that further optimization of labeling sites will be required. Cy5-labeled TEN domain observed with Cy3-labeled DNA demonstrated a relatively well-defined distribution of FRET

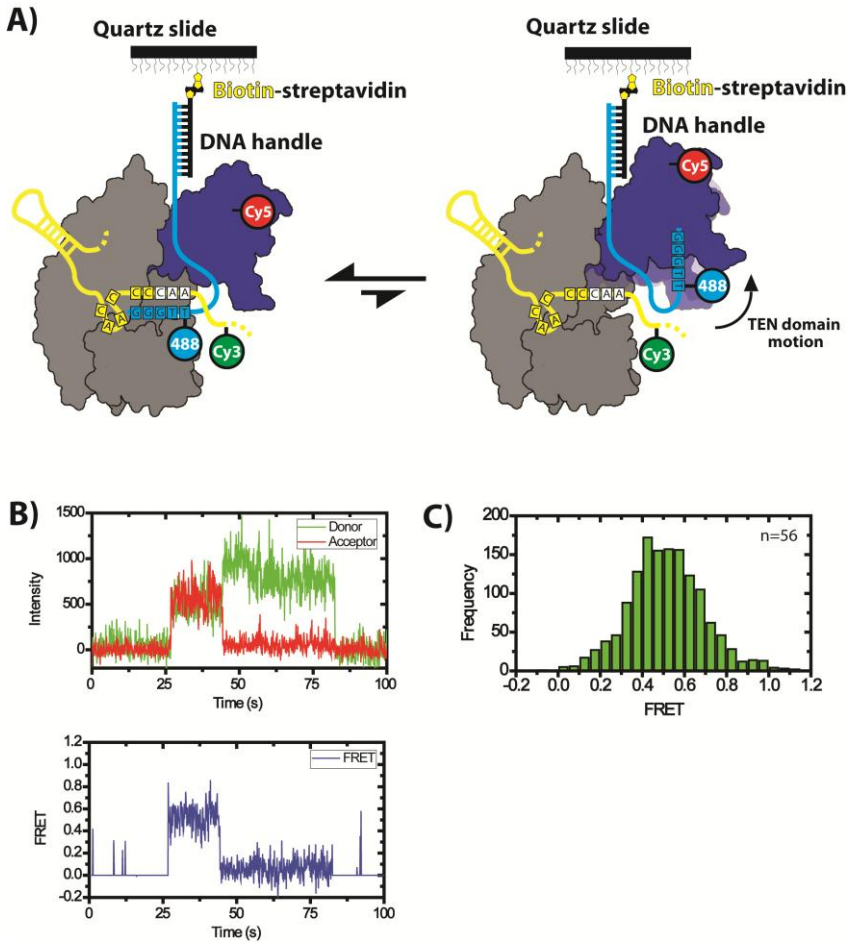


Figure 7: A) Schematic diagram of TEN domain FRET experiments. The TEN domain is labeled with Cy5 using an incorporated ybbR tag. U63 of TER is labeled with Cy3. Potential TEN domain motions would be measured as changes in FRET between the two labeling sites. **B)** Representative smFRET trace for TEN domain FRET experiments. **C)** smFRET histogram of TEN domain FRET measurements using the TEN domain labeling site at residue 48. FRET histograms demonstrate a broad distribution centered at ~0.5 FRET.

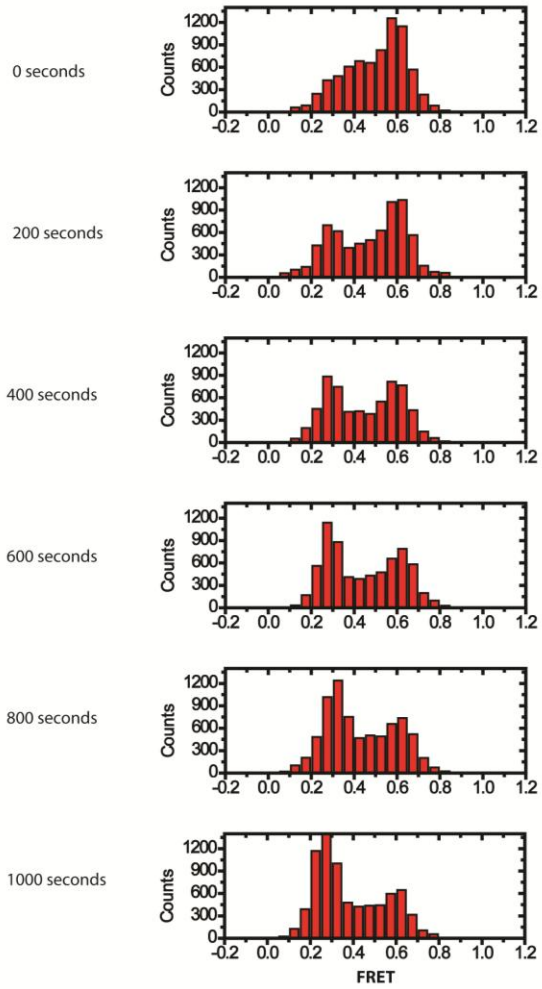
values, suggesting that the abnormally wide distribution found with protein-RNA FRET measurements may in fact be due to conformational rearrangements at the U63 RNA-labeling position. This indicates that the U63 labeling site may not be the ideal candidate for RNA labeling.

It remains to be seen whether we will detect any TEN domain motions coupled with the transition between the active and alternative states. It is important to note as well that in smFRET experiments, failure to detect a change in the FRET signal does not mean a change does not occur. TEN domain motions could occur in such a way as they do not alter the distance between labeling sites in three dimensional space. Furthermore, TEN domain motions may not be coupled to the shift between the active and alternative states, but could be coupled to other transitions in the enzyme. A logical alternative source of TEN domain motions would be translocation of TER during repeat addition processivity, and this possibility should be tested as well.

Human studies on TEN domain mutations

Having established the role of the *Tetrahymena thermophila* TEN domain in stabilizing DNA-RNA duplexes, a logical next step would be to test the effect of TEN domain mutations on the human enzyme. Recent studies suggest that *T. thermophila* TEN domain residues L14 and Q168 have analogs in the human residues L13, L14, and Q169. Furthermore, it has been demonstrated that mutations to these residues exhibit similar defects in telomerase processivity[5,10]. Our lab has recently developed a platform to study the human enzyme by smFRET. This would be an ideal system in which to study the effects of TEN domain mutants on DNA binding dynamics in the human enzyme.

Supplemental Figure 1



Supplemental Figure 1. Time course of Cy3-Cy5 FRET values observed across the DNA primer in three color FRET experiments with primers immobilized via a DNA handle (as in Figure 2A). A bimodal distribution was observed, with primers demonstrating both a ~0.6 FRET distribution and a ~0.3 FRET distribution. In the presence of immunopurified telomerase, this distribution shifted over time. Initially, the 0.6 FRET state is the predominant FRET state, but by 1000 seconds the 0.3 FRET state has become the predominant FRET state. Only the 0.6 FRET state appears to be competent to bind telomerase in our smFRET assay. Therefore it is important that in smFRET assays utilizing a DNA handle measurements are limited to about the first 20 minutes, as it is likely that many fewer binding events will be observed after the first 20 minutes.

References

1. Lee, S., Lee, J. & Hohng, S. Single-molecule three-color FRET with both negligible spectral overlap and long observation time. *PLoS One* **5**, e12270 (2010).

2. Kapanidis, A.N. et al. Fluorescence-aided molecule sorting: analysis of structure and interactions by alternating-laser excitation of single molecules. *Proc Natl Acad Sci U S A* **101**, 8936-41 (2004).

3. Ha, T. et al. Single-molecule fluorescence spectroscopy of enzyme conformational dynamics and cleavage mechanism. *Proc Natl Acad Sci U S A* **96**, 893-8 (1999).

4. 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).

5. Zaug, A.J., Podell, E.R. & Cech, T.R. Mutation in TERT separates processivity from anchor-site function. *Nat Struct Mol Biol* **15**, 870-2 (2008).

6. Yin, J., Lin, A.J., Golan, D.E. & Walsh, C.T. Site-specific protein labeling by Sfp phosphopantetheinyl transferase. *Nat Protoc* **1**, 280-5 (2006).

7. 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**, 218-25 (2006).

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

9. Akiyama, B.M., Gomez, A. & Stone, M.D. A conserved motif in Tetrahymena thermophila telomerase reverse transcriptase is proximal to the RNA template and is essential for boundary definition. *J Biol Chem* **288**, 22141-9 (2013).

10. 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).

CHAPTER VII: Potential future experiments on telomerase structure and dynamics

Introduction

A large number of questions concerning telomerase structure, assembly, and dynamics remain to be addressed. One prescient question regards the nature of the RNA pseudoknot. Pseudoknots are a highly conserved feature of telomerase RNAs, however it is unclear exactly what role they play in telomerase activity. In this chapter I discuss potential experiments to localize the RNA pseudoknot within the telomerase holoenzyme, and determine whether or not a pseudoknot-hairpin equilibirium plays a role in telomerase RNA translocation. Further questions remain regarding other interactions between the protein and the RNA. In this chapter, I also describe future crystallography experiments that could be used to characterize sites of protein-RNA interaction in the *Tetrahymena thermophila* telomerase holoenzyme.

I also describe our efforts to characterize the rate constants of various steps in the telomerase catalytic cycle using smFRET activity assays. Previous experiments have characterized the rate of repeat addition processivity (RAP), which occurs on the timescale of minutes[1]. However, it has been more challenging to measure the rate of nucleotide addition processivity (NAP). Single-molecule experiments offer the possibility to directly observe the rate constants of individual steps in the telomerase catalytic cycle in real time. However, our initial experiments have yielded confusing results. It is likely that further breakthroughs in telomerase expression and

purification will be necessary to complete these studies and I discuss some ideas for improving our smFRET activity assay.

Monitoring telomerase activity in real time using smFRET

smFRET experiments on active telomerase in the presence of nucleotides offer the opportunity to measure rate constants of the various steps of the telomerase catalytic cycle. In particular, the experiments could measure the rate of nucleotide addition processivity. Although the rate of RAP has been previously studied, NAP has been more difficult to characterize. This is because RAP is the rate limiting step of telomerase extension, and so processivity measurements are dominated by this slower step. In principle, it may be possible to measure the rate of NAP using a stop-flow apparatus, however the extremely poor yield of heterologously expressed telomerase and the known heterogeneity in reconstituted telomerase molecules could complicate the interpretation of this data. Single molecule experiments can circumvent these limitations, as single molecule experiments do not require a lot of material and have the ability to separate out separate populations of molecules.

Experiments with telomerase extension intermediates have demonstrated a clear change in FRET as Cy5-labeled telomerase progressively extends a Cy3-labeled DNA primer (Figure 1A&B). This offers a clear prediction on the FRET changes that should occur when a Cy3-labeled DNA primer is allowed to extend to the end of a single telomerase repeat (Figure 1C). In order to test this prediction, primer E6 labeled with Cy5 was incubated with telomerase labeled with Cy5 at position U36 in

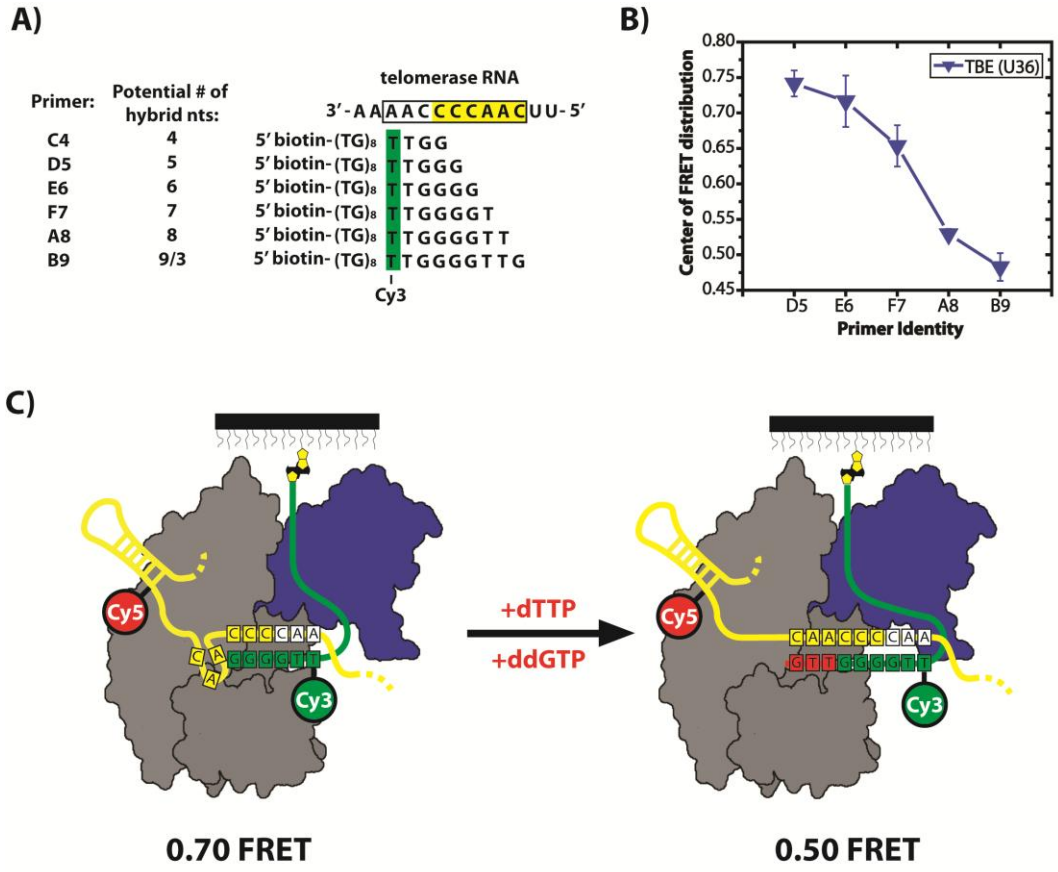


Figure 1 A) List of primers used in previous smFRET telomerase binding studies. Primers were 5' labeled with biotin in order to immobilize on slides for smFRET studies. Primers contained a $(TG)_8$ motif followed by a telomeric DNA sequence which hybridizes with telomerase RNA. Primers were given a letter designation and a number designation based on the maximum number of DNA-RNA basepairs the primer could form. Primers were labeled with Cy3 at the indicated position. **B)** Diagram representing the center of the smFRET distribution obtained with primers D5-B9 in smFRET binding experiments in the absence of nucleotide, when primers from Figure 1A were tested with telomerase labeled in the TER subunit at residue U36. As the primer is progressively extended, the FRET distribution shifts from ~0.70 FRET to ~0.50 FRET, suggesting a FRET signal that should be present in real-time activity assays. **C)** Schematic diagram of real-time activity assays. Cy3-labeled primer E6 is immobilized on a smFRET microscope slide and U36-labeled telomerase is flowed on in the presence of dTTP and ddGTP. In this setup, telomerase is only capable of adding a single repeat, and will not undergo RAP, allowing us to measure the rate of addition of a single repeat.

the telomerase RNA in the presence of dTTP and ddGTP. The use of this chain-terminating dideoxy nucleotide allows us to control the extension of the telomeric DNA primer such that it can extend to the end of a single telomeric repeat but no farther. This simplifies the interpretation of our experiment by limiting the extension to three nucleotides over which we have a clear expectation of the expected FRET change. smFRET histograms demonstrate that the addition of dTTP and ddGTP to our smFRET experiments shifts our FRET distribution from one centered around ~0.70 FRET to one centered around ~0.50 FRET, consistent with our expectations. (compare Figures 2A-C). These effects are nucleotide-dependent (Figure 2A) and require the active site telomerase residues, as a catalytically inactive telomerase fails to alter the FRET distribution even in the presence of nucleotide (Figure 2E).

When individual smFRET traces for Cy3-labled primer E6 in the presence of dTTP and ddGTP were studied there were predominantly two types of traces. One subset of traces appears to have already been extended to the end of the first telomeric repeat by the time of our initial observation (Figure 3A). There are two explanations for this phenomenon. The first is that another telomerase molecule bound and extended this primer before the binding event that we were able to directly observe occurred. The other explanation is that NAP occurs on a time scale faster than our ability to measure and that these traces represent molecules that were instantaneously extended from the point of view of our 0.1s time resolution.

In another subset of molecules, we directly observe transitions from the ~0.70 FRET state to the ~0.50 FRET state. These transitions occur very quickly, near the

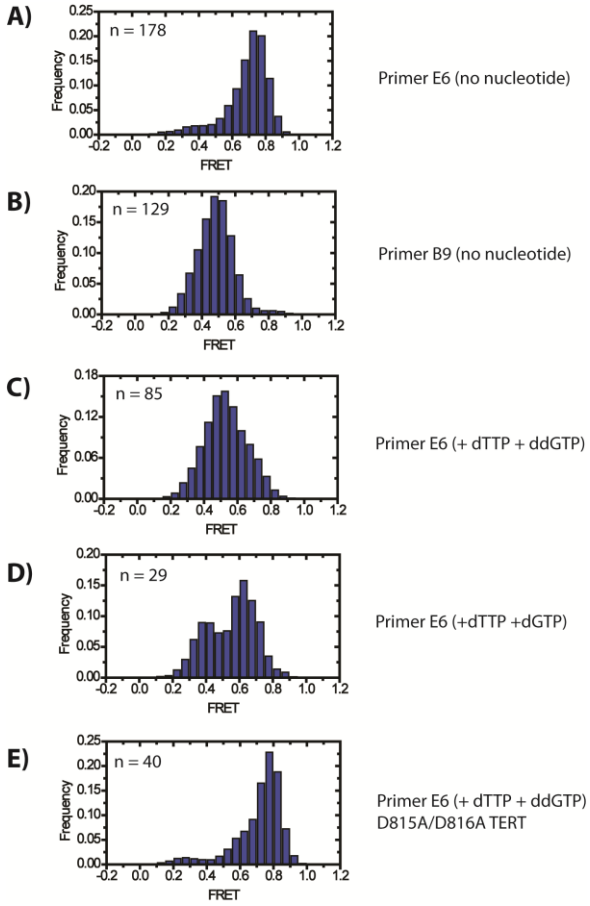


Figure 2. smFRET histograms of primer E6 in the presence of actively extending telomerase enzyme. smFRET histograms of primer E6 **(A)** and primer B9 **(B)** in the absence of nucleotide, demonstrating the expected beginning and end points of telomerase extension. **C)** smFRET histogram of primer E6 in the presence of dTTP and ddGTP. In the presence of these nucleotides, telomerase appears to extend for a single repeat, as the FRET distribution shifts from the normal E6 distribution (Figure 2A) to a distribution more resembling the B9 distribution (Figure 2B). **D)** smFRET histogram of the FRET distribution observed with primer E6 incubated with dTTP and dGTP. This primer is capable of extending for multiple repeats. The distribution suggests two major populations at ~0.4 and ~0.6 FRET. In principle, these could represent the pre-translocated and post-translocated states of the enzyme. **E)** smFRET histogram of primer E6 in the presence of nucleotide tested with telomerase bearing two mutations in the telomerase active site. The smFRET distribution most closely resembles primer E6 in the absence of nucleotide, confirming that the previous shifts in smFRET distributions are due to telomerase activity.

detection limit of our measurements (Figure 3B). These transitions follow a short pause during the ~0.70 FRET state before primer extension is initiated, perhaps related to a conformational rearrangement that must occur upon binding before primers are locked in for activity.

A further subset of the molecules demonstrating transitions appear to transition from the ~0.70 FRET state to the ~0.50 FRET state and back again, alternating back and forth between the two states (Figure 3C). The cause of these particular dynamics is unknown. It is tempting to speculate that these are molecules that are attempting to undergo translocation at the end of a single round of repeat addition processivity. In this case the Cy3 label on the DNA could be moving to a less defined position as the DNA-RNA duplex is denatured. However, we do not have any data to conclusively support this hypothesis.

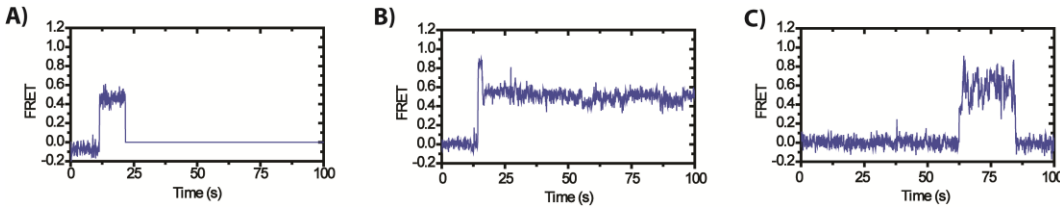


Figure 3. Representative smFRET traces for primer E6 incubated with Cy5-labeled telomerase incubated with dTTP and ddGTP. In these smFRET traces, the enzyme is only capable of extending for a single repeat. Three different types of behaviors were observed. **A)** In the first type of behavior, primers demonstrate a rapid decrease in FRET and remained stable at the ~0.5 FRET state. **B)** In the second behavior, molecules begin at the ~0.5 FRET state and remain stable at this state. **C)** In the third type of behavior, primers rapidly alternate between a ~0.70 FRET state and a ~0.5 FRET state.

In general, the subset of traces demonstrating clear transitions between FRET values represented a low percentage of the total traces as a whole, as low as 10% of the total, with the vast majority of the changes demonstrating a stable ~0.50 FRET state. This complicates our data interpretation because it is difficult to maintain with certainty that we are not simply choosing the traces which conform to our predictions for our analysis. Furthermore, even if we were certain that our observed traces are a valid representation of the kinetic steps of telomerase activity, it would be difficult to collect sufficient traces for statistically significant measurements. Therefore it is likely that further breakthroughs in the expression, reconstitution, and biochemical purification of telomerase will be necessary to perform adequate measurements of telomerase kinetics, even at the single molecule level.

This point becomes further apparent when we consider telomerase incubated with primers in the presence of both dTTP and dGTP. In this case, there is no chain-terminating nucleotide and telomerase molecules are free to extend for several successive repeats. smFRET histograms demonstrate that on average these molecules occupy a lower FRET state, suggesting that further DNA extension occurs, pushing the Cy3-label farther from the telomerase active site and forming two distinct FRET distributions at ~0.4 and ~0.6 FRET (Figure 2D). However, when one analyzes the single-molecule FRET traces (Figure 4) it becomes apparent that primers extruded beyond a single repeat exhibit complex motions that defy a simple assignment as to how far they have been extended. This is consistent with primer DNA occupying a

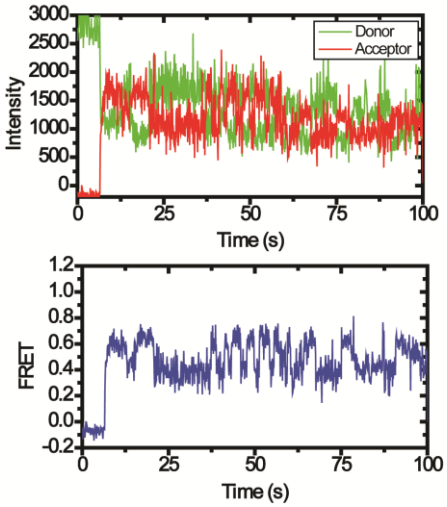


Figure 4. smFRET trace of extending telomerase enzyme. Cy3-labeled primer E6 was incubated with U36-labeled telomerase in the presence of dTTP and dGTP. Under these conditions, telomerase is capable of extending for multiple rounds of telomerase extension. The observed primers demonstrate an equilibrium between a ~0.4 and a ~0.6 FRET state. It is unclear what is the cause of these dynamics.

less well-defined position in successive DNA repeats as the DNA is extruded from the active site and no longer stabilized by telomerase-DNA interactions.

Future smFRET measurements of active telomerase are likely going to involve more highly expressed and better purified enzyme. This could involve the use of an improved heterologous expression technique or the ability to label telomerase purified from endogenous sources with a site-specific fluorophore molecule. Recent experiments from other groups have demonstrated the importance of accessory factors in the telomerase holoenzyme, including processivity factors Teb1, p50, and the p7-1-4 complex[2-4]. Inclusion of these factors in a single complex may be critical for stabilizing an exit channel of telomere DNA from telomerase, generating more interpretable FRET traces. However, thusfar we have not been able to perform an

efficient reconstitution of the full telomerase holoenzyme *in vitro*. One future strategy may be to immobilize one of the critical processivity factors on the surface of the slide, and tether the remainder of the telomerase holoenzyme to the surface via its interaction with that processivity factor. In this way we could be certain that all immobilized molecules contained the processivity factor, eliminating a source of heterogeneity (Figure 5).

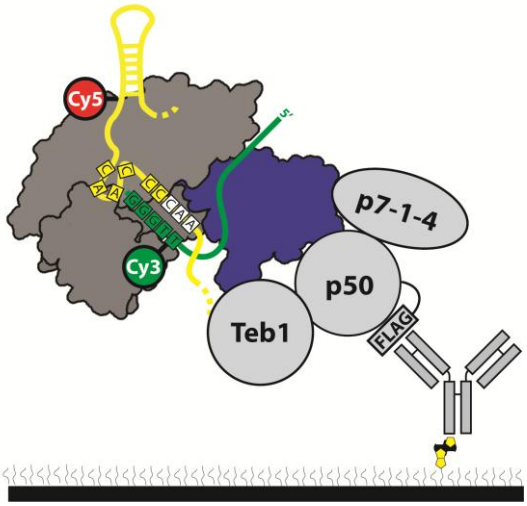


Figure 5. Immobilization strategy for *in vitro* reconstituted telomerase holoenzyme. In order to study *Tetrahymena* telomerase in the context of its natural processivity factors, telomerase could be reconstituted *in vitro* with its known accessory factors (Teb1, p50, and the p7-1-4 complex). Previous attempts to purify the complex by means of a tag on the TERT subunit have resulted in sub-stoichiometric quantities of the accessory factors. In order to circumvent this problem, telomerase could be directly purified on smFRET microscope slides by its interaction with telomerase accessory factors. In this experiment, anti-FLAG IgG would be immobilized on a quartz microscope slide via a biotin-streptavidin linkage. RRL reactions containing reconstituted telomerase, Teb1, p7-1-4, and a FLAG-tagged p50 subunit would be flowed onto the slide and subjected to several washing steps. Dye-labeled DNA primers could then be introduced, and FRET could be measured between a label on telomerase RNA and a label on the DNA.

At the moment, it is difficult to conclude anything definitive from our initial smFRET measurements on telomerase activity. From the smFRET histograms (Figure 2A-E) it is apparent that we can reconstitute telomerase activity on the surface of a smFRET slide. However, the resultant smFRET traces have unexpected complexities which makes them difficult to interpret. Our initial observations suggest that the rate of NAP is likely quite fast compared to RAP, likely occurring faster than the seconds time scale. However, we do not have data of sufficient quality to accurately quantify this rate constant. Attempts to characterize the rate constants associated with RAP are further hampered by our inability to determine how far a primer has extended by its association with a single FRET state. It appears likely that further developments in the expression, reconstitution, and purification of telomerase may be required to perform smFRET telomerase activity measurements.

Potential studies on the telomerase RNA pseudoknot

One of the most fascinating questions regarding the mechanism of telomerase activity concerns the role of the RNA pseudoknot (PK). This three-stranded RNA structure is highly conserved in telomerase RNAs[5]. Mutations to the pseudoknot can severely affect telomerase *in vitro* and has been linked to heritable disease *in vivo*[6]. However, it is not currently understood why telomerase RNAs require a pseudoknot for activity. A recent structure of the human telomerase RNA PK has been solved, demonstrating the role of conserved base triples in maintaining this stable RNA tertiary structure[6]. The realization that the pseudoknot can alternate between a hairpin

and a pseudoknot state suggested that this equilibrium could be functionally coupled to RNA motions within the enzyme[7]. However, more recent work suggests that in the presence of physiological magnesium ions, the PK occupies a more stable fold than previously thought, suggesting that it may not be a mobile element within the holoenzyme[8]. In order to fully understand how the RNA PK contributes to telomerase activity two questions must be answered. First, we must have a better understanding of the location of the telomerase RNA PK within the context of the assembled telomerase holoenzyme. Secondly, we must understand if the PK/hairpin equilibrium really does drive productive conformational changes within the enzyme.

Addressing the location of the PK in the enzyme is a difficult question. Interactions between the PK and any conserved protein motifs have not been discovered. In theory a site-specific biochemical probing, such as the site-directed hydroxyl radical probing described in Chapter IV could be used. Alternatively, a cross-linking experiment could be designed by incorporating chemically reactive groups into the RNA PK itself. There is some evidence in yeast that the PK can cross-link to the DNA primer, suggesting a close physical interaction, and phylogenetic analysis from multiple species suggests that the PK remains adjacent to the template RNA in a conserved "core" domain[9]. smFRET could also be used to provide distance constraints, although initial experiments suggest that the PK remains close to many other RNA elements, consistent with the telomerase RNA "core" model[9,10].

In order to address whether or not the PK participates in a PK-hairpin equilibrium during telomerase processivity, it may be possible to design site specific

RNA cross-links in order to covalently close the RNA PK and restrict the ability of the PK to rearrange. This experiment would be best suited to the human system where the existing RNA PK structure could be used as a guide. Amine-modified RNA bases could be cross-linked together using a bifunctional linker, and appropriately cross-linked RNAs could be purified by denaturing PAGE. The RNA PK could be splint-ligated into the full-length telomerase RNA and tested activity. The activity could be compared to WT activity as well as the activity of the RNA bearing the amine modification in the absence of cross-linking reagent. The presence of telomerase activity in the presence of a covalent cross-link preventing denaturation of the RNA PK would definitively demonstrate that no PK-hairpin rearrangement occurs during telomerase activity. Conversely, if the modification prevents telomerase activity but the non-cross-linked amine modifications do not, this would be a strong argument in favor of a functional PK-hairpin equilibrium.

Further characterization of protein-RNA interactions in telomerase

Biochemical studies on telomerase indicate that multiple sites of protein-RNA interaction occur within the telomerase holoenzyme. In *Tetrahymena* telomerase, the template boundary element (TBE), stem I, and stem-loop IV have all been implicated in protein-RNA interactions[11-13]. Our recent site-directed hydroxyl radical probing experiments have identified a site of direct interaction between the TBE and the RBD of TERT[14]. Using these studies as a guide, we have designed a new RBD construct that contains all of the essential conserved RNA-binding elements within the RBD.

This construct has been amenable to crystallography studies. Efforts to co-crystallize this RBD construct with a minimal RNA construct containing only the TBE residues in stem II have yielded initial hits in crystallization screens. Following optimization of these crystallography conditions, we are confident that we will be able to see the molecular details of TBE-RBD interactions.

However, there remain other interactions between telomerase RNA and the RBD that have yet to be fully characterized by biochemical or structural methods. Initial studies in our lab have demonstrated that the addition of stem I and stem IV RNA to minimal TBE constructs significantly increase the affinity of RBD for telomerase RNA constructs. (Figure 6A-C). When combined with biochemical evidence demonstrating that these regions are important in protein-RNA interactions[11], these suggest a direct interaction between regions of stem I and stem IV and the TERT RBD. Therefore, following the conclusion of our TBE-RBD crystallography studies, a logical next step would be to design minimal stem I and minimal stem-loop IV RNA constructs and test these constructs for RBD interactions. After identifying new minimal RNA constructs that form complexes with RBD *in vitro*, we could use these new RNA constructs in future crystallography studies to build a more complete model of the telomerase holoenzyme.

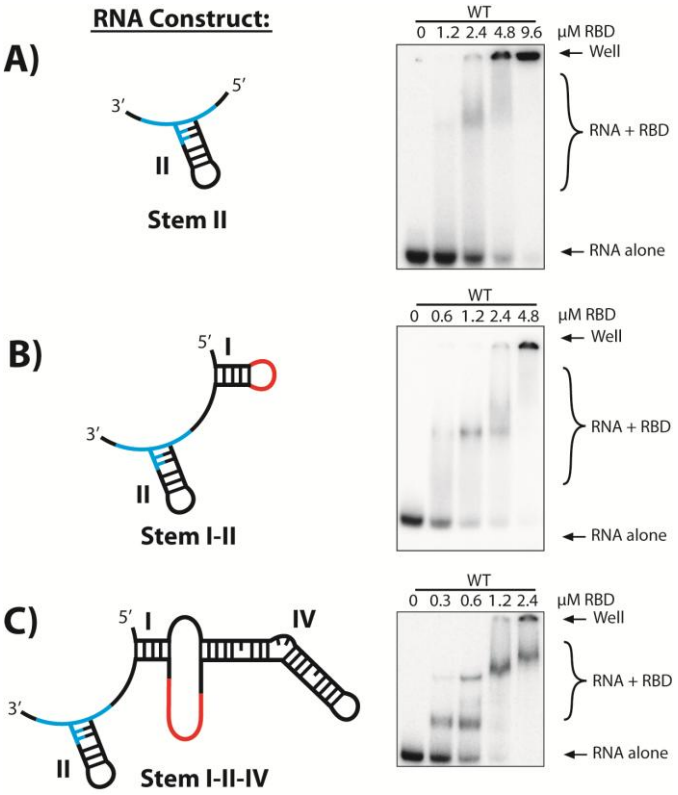


Figure 6. *Tetrahymena* telomerase RNA constructs demonstrate multiple interactions between the TERT RBD subunit and telomerase RNA. **A)** An RNA construct containing only stem II of TER demonstrates relative weak affinity for telomerase RBD. Electophoretic mobility shift assays (EMSAs) were performed using radioalabeled RNA constructs and purified RBD protein. The stem II construct demonstrates a K_d in the 5 µM range. **B)** Addition of stem I to the stem II RNA improves the K_d to around 1 µM. **C)** Addition of stem IV to the Stem I-II construct further improves the affinity to a K_d of about ~0.5 µM.

References

1. Greider, C.W. Telomerase is processive. *Mol Cell Biol* **11**, 4572-80 (1991).

2. Min, B. & Collins, K. An RPA-related sequence-specific DNA-binding subunit of telomerase holoenzyme is required for elongation processivity and telomere maintenance. *Mol Cell* **36**, 609-19 (2009).

3. Min, B. & Collins, K. Multiple mechanisms for elongation processivity within the reconstituted tetrahymena telomerase holoenzyme. *J Biol Chem* **285**, 16434-43 (2010).

4. Jiang, J. et al. The architecture of Tetrahymena telomerase holoenzyme. *Nature* **496**, 187-92 (2013).

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

6. Theimer, C.A., Blois, C.A. & Feigon, J. Structure of the human telomerase RNA pseudoknot reveals conserved tertiary interactions essential for function. *Mol Cell* **17**, 671-82 (2005).

7. Theimer, C.A., Finger, L.D., Trantirek, L. & Feigon, J. Mutations linked to dyskeratosis congenita cause changes in the structural equilibrium in telomerase RNA. *Proc Natl Acad Sci U S A* **100**, 449-54 (2003).

8. Hengesbach, M., Akiyama, B.M. & Stone, M.D. Single-molecule analysis of telomerase structure and function. *Curr Opin Chem Biol* **15**, 845-52 (2011).

9. Lin, J. et al. A universal telomerase RNA core structure includes structured motifs required for binding the telomerase reverse transcriptase protein. *Proc Natl Acad Sci U S A* **101**, 14713-8 (2004).

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

11. Lai, C.K., Mitchell, J.R. & Collins, K. RNA binding domain of telomerase reverse transcriptase. *Mol Cell Biol* **21**, 990-1000 (2001).

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

13. O'Connor, C.M., Lai, C.K. & Collins, K. Two purified domains of telomerase reverse transcriptase reconstitute sequence-specific interactions with RNA. *J Biol Chem* **280**, 17533-9 (2005).

14. Akiyama, B.M., Gomez, A. & Stone, M.D. A conserved motif in Tetrahymena thermophila telomerase reverse transcriptase is proximal to the RNA template and is essential for boundary definition. *J Biol Chem* **288**, 22141-9 (2013).