UC Santa Cruz

UC Santa Cruz Electronic Theses and Dissertations

Title

Structure and Dynamics of Telomerase

Permalink

https://escholarship.org/uc/item/6sk3j31x

Author

Jansson-Fritzberg, Linnea

Publication Date

2018

Peer reviewed|Thesis/dissertation

eScholarship.org Powered by the California Digital Library
University of California

UNIVERSITY OF CALIFORNIA

SANTA CRUZ

STRUCTURE AND DYNAMICS OF TELOMERASE

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

DOCTOR OF PHILOSOPHY

in

MOLECULAR, CELL and DEVELOPMENTAL BIOLOGY

by

Linnea I. Jansson-Fritzberg

September 2018

The Dissertation of Linnea I. Jansson-Fritzberg is approved:

_____

Professor Michael D. Stone, Chair

_____

Professor Harry F. Noller

_____

Professor Alan M. Zahler

_____

Lori Kletzer
Vice Provost and Dean of Graduate Studies

Copyright by

Linnea Jansson-Fritzberg

2018

TABLE OF CONTENTS

TABLE OF CONTENTS Page

List of Figures and tables……. v

Abstract…… viii

Acknowledgments…… x

CHAPTER 1: Introduction

General introduction…… 1

References…… 9

CHAPTER 2: Structural basis for template boundary definition in Tetrahymena telomerase

Abstract…… 14

Introduction……. 15

Results…….. 19

Discussion…… 30

Materials and Methods…… 34

References…… 37

Supplementary information……. 41

CHAPTER 3: DNA structure dependent modulation of telomerase activity

Abstract…… 46

Introduction…….. 47

Results…….. 52

Discussion…… 67

Materials and Methods…… 72

References…… 80

Supplementary information…… 87

CHAPTER 4: Ongoing single molecule FRET studies probing dynamics of the telomerase enzyme during catalysis

Introduction……. 98

Human telomerase reconstitution and purification……. 99

Site specific labeling of hTERT…….. 100

Activity-dependent purification of human telomerase…… 108

Effects of $MgCl_2$ on telomerase activity……. 111

Effects of POT1/TPP1 on telomerase activity and dynamics……. 113

Future optimization of human telomerase smFRET experiments…….. 118

Ongoing single molecule studies of telomerase dynamics and kinetics in real time…….. 119

Future ZMW experiments…….. 128

References…….. 130

List of figures and tables

CHAPTER 1

Figure 1: Telomeres shorten with each cell division…….. 1

Figure 2: The telomerase catalytic cycle…… 3

Figure 3: Telomerase components……. 4

CHAPTER 2

Figure 1: *Tetrahymena thermophila* telomerase composition and catalytic cycle…… 17

Figure 2: Structure of the *T. thermophila* TRBD (tTRBD)-TBE complex…… 22

Figure 3: Protein-protein interactions between conserved tTERT RBD motifs…… 24

Figure 4: Protein-RNA interactions within the tTRBD-TBE complex……. 27

Figure 5: Conservation and function of observed interactions in the tTRBD-TBE complex……... 30

Figure 6: Structural model of *T. thermophila* telomerase RNA template connectivity to the TER template boundary element…… 32

Table 1: Data collection and refinement statistics (molecular replacement)… 34

Supplemental Figure 1: Cross species TERT-RBD sequence alignments…… 42

Supplemental Figure 2: Ciliate TERT-RBD sequence alignments……... 43

Supplemental Figure 3: RNA binding activity of new tTRBD protein construct…… 44

Supplemental Figure 4: Crystal packing arrangement of the tTRBD-TBE RNA complex……... 45

Supplemental Figure 5: Comparison of TER secondary structure…… 46

CHAPTER 3

Figure 1: Human telomerase function…… 50

Figure 2: Product distribution profile is dependent on amount of consecutive TTAGGG repeats…….. 56

Figure 3: Product distribution profile is dependent on monovalent salt identity…… 58

Figure 4: smFRET studies of telomerase in the presence of K^+ and Li^+…….. 60

Figure 5: Human telomerase kinetics…… 65

Figure 6: Model of GQ folding within actively extending telomerase complex……... 69

Supplemental Figure 1: Using radiolabeled dGTP or radiolabeled primer result in similar product accumulation profiles……. 90

Supplemental Figure 2: Telomerase is processive in the presence of 50nM primer…….. 91

Supplemental Figure 3: Telomerase activity is slower in Li^+…… 92

Supplemental Figure 4: Real-time smFRET traces……... 93

Supplemental Figure 5: Example Monte-Carlo analysis to illustrate confidence of fitting and covariance of parameters for the scheme shown in Figure 6b…… 94

Supplemental Figure 6: Reproducibility of kinetic analyses…… 95

Supplemental Figure 7: Model of a GQ on the telomerase enzyme……. 96

Table 1: Oligos used for this study…… 97

CHAPTER 4

Figure 1: Sfp labeling…… 101

Figure 2: Current ybbR label sites on TERT…… 102

Table 1: Sites used for labeling tag insertion within TERT…… 103

Figure 3: smFRET experimental setup……. 105

Figure 4: FRET distribution for ybbR-TEN enzymes…….. 107

Figure 5: Activity-dependent purification of telomerase……. 109

Figure 6: Excess $MgCl_2$ disrupts telomerase activity…….. 112

Figure 7: Reducing $MgCl_2$ in the RRL improves telomerase activity……. 114

Figure 8: POT1/TPP1 induces a shift in TEN-ybbR mutant

FRET populations…… 115

Figure 9: POT1/TPP1 effects on TEN-ybbR mutants……. 117

Figure 10: Zero Mode Waveguide technology…… 120

Figure 11: Telomerase is active in presence of PacBio nucleotide analogs… 122

Figure 12: Telomerase nucleotide titration…….. 123

Figure 13: Telomerase activity time course……..... 124

Figure 14: Telomerase ZMW traces…… 126

Figure 15: Real-time detection of telomerase activity…… 128

Abstract

Linnea Jansson-Fritzberg

Structure and Dynamics of Telomerase

Telomeres are repetitive, G-rich DNA sequences, along with DNA-associated proteins, that cap the ends of linear chromosomes. Telomeres serve a two-fold function; they protect the ends of chromosomes from being recognized as sites of DNA damage and they prevent fraying of chromosomes into coding regions during successive rounds of DNA replication. In proliferative cells where unregulated chromosome end shortening would be a critical problem, telomere lengths are maintained by the specialized reverse transcriptase enzyme, telomerase. Because of its requirement in proliferative cells, telomerase is also upregulated in about 90% of cancer, making it an attractive target for cancer therapeutics.

Telomerase is a ribonucleoprotein, composed of a protein component, TERT, and an RNA component, TR. Telomerase reverse transcribes telomere repeats onto the 3' end of chromosomes through its integral template within TR. Telomerase is unique among reverse transcriptases in that it is able to add multiple telomere repeats during a single binding event. The precise conformational rearrangements required for this processive telomerase action are not well understood.

In this thesis, I focus on the interaction between telomerase and the nascent telomere and how reverse transcription of multiple repeats affects the actively extending telomerase complex. First, I focus on how template boundary definition ensures the faithful synthesis of the required hexameric telomere sequence and

demonstrate that critical regions within both TERT and TR are responsible for establishing strict template boundary definition. Second, I focus on the interplay between DNA structure formation within the nascent telomere and the actively extending telomerase complex and show that formation of G-quadruplexes within the telomere sequence affects telomerase activity. Lastly, I discuss ongoing work using single molecule techniques to investigate dynamics of human telomerase during the telomerase catalytic cycle and the problems that must be overcome to complete this work.

Acknowledgements

First, I would like to thank Dr. Michael Stone for being such an amazing mentor to me during my graduate work. He has shown me what a great mentor looks like and I hope to one day use what I have learned from him to become a great mentor myself. He is not only a great scientist but also a really good person and I am really glad that I have gotten the chance to get to know him. I would additionally like to thank Dr. Harry Noller and Dr. Alan Zahler who have provided excellent advice as part of my thesis committee. I would also like to thank the Stone lab, it has been such a fun environment to conduct science in. I would like to thank my family, especially my parents Janet and Christer Jansson, who have been so supportive of my goal to become an independent scientist. They are both great role models who I constantly look up to. Finally, I would like to thank my husband David, who has been my personal champion and whose support of me pursuing my dream career has never wavered.

CHAPTER 1: Introduction

The ends of linear chromosomes are capped by repetitive G-rich stretches of DNA that are the foundation for telomere structure [1,2]. Due to incomplete DNA replication of chromosome termini caused by the end-replication problem, telomere DNA gradually shortens with each cell division (**Fig. 1**) [3-5].

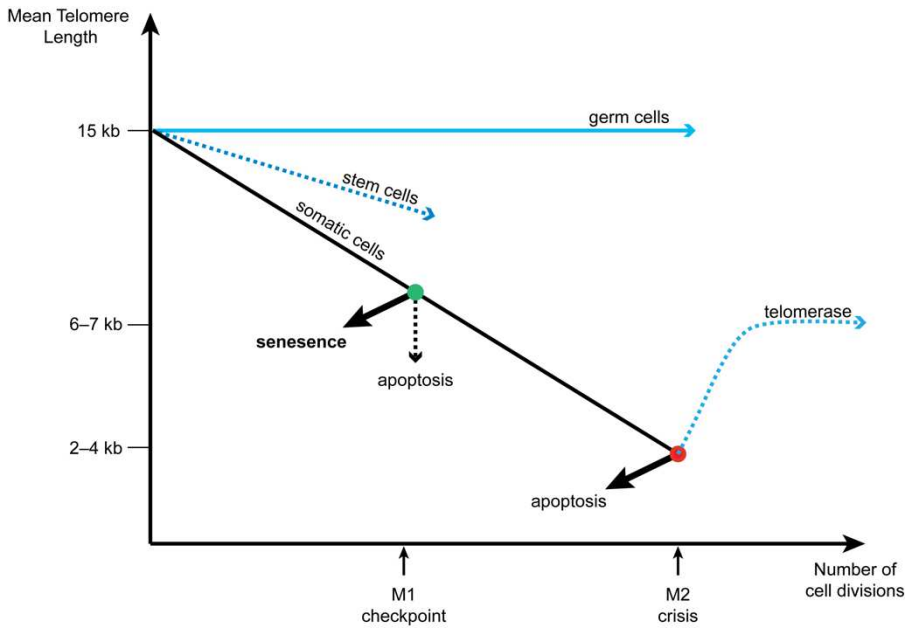


Figure 1. Telomeres shorten with each cell division. Telomere length in somatic cells shorten with each round of DNA replication due to the end replication problem. At a critical telomere length, cells either senesce or undergo apoptosis (green dot). If cells keep dividing they will eventually reach the M2 crisis point (red dot) which triggers apoptosis. If telomerase is upregulated at any point before this crisis point however, telomere length is maintained and cell division is allowed to continue. This aberrant upregulation of telomerase is a common hallmark of cancer cells. Telomerase expression is also upregulated in healthy germ cells, and to a lesser degree, stem cells, in order to provide these cells with unlimited or enhanced dividing potential, respectively.

Without a mechanism to counteract this shortening, telomeres become critically short and serve as a signal for the cell to undergo senescence or programmed cell death.

This senescent state and eventual cell death is a necessary mechanism to prevent unrestricted cell division in somatic cells [6]. In contrast, cells that need to maintain higher levels of proliferative potential, such as stem cells and germ cells, express the specialized enzyme, telomerase, to maintain telomere length [7]. The same principle holds true for most human cancer cell types, making telomerase an important biomedical drug target [8].

Telomerase is a ribonucleoprotein complex composed of a protein subunit, TERT, and an integral long noncoding RNA component, TR [9,10]. TERT and TR constitute the telomerase catalytic core, whereas a number of additional species-specific proteins can also associate with the telomerase complex [2]. TERTs are generally well conserved and share significant sequence homology with viral RT enzymes [11,12]. Telomerase is a canonical RT in the sense that it catalyzes a reverse transcription reaction to synthesize telomere DNA from its RNA template. However, telomerase also has additional specialized features that support its ability to synthesize multiple telomere DNA repeats during a single binding event [13,14]. During catalysis, telomerase docks on to the 3' single stranded end of the telomere DNA tail, and reverse transcribes the integral template region within TR onto the end of the chromosome (**Fig. 2**). Following the addition of a complete telomeric DNA repeat, telomerase must realign the nascent DNA with downstream template residues in order to prime a subsequent round of telomere repeat synthesis. The unique property of telomerase to recycle its internal RNA template and synthesize multiple telomere

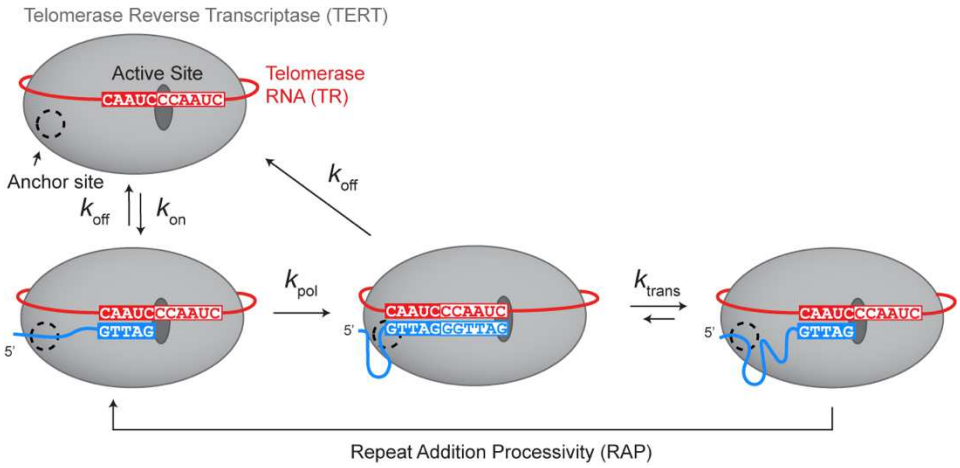


Figure 2. The telomerase catalytic cycle. Telomerase anneals to the 3' end of the telomere (blue) and reverse transcribes a telomere repeat based off of its intrinsic RNA template within TR (red). After the addition of a full telomere repeat, telomerase translocation occurs in order to allow the new telomere end to reanneal to the 3' end of the template to prime for another round of repeat addition. This process is termed Repeat Addition Processivity (RAP).

DNA repeats without dissociating from the telomere is termed Repeat Addition Processivity (RAP).

The TERT protein is composed of four conserved domains; The Telomerase Essential N-terminal (TEN) domain, the RNA Binding Domain (RBD), the Reverse Transcriptase (RT) domain and the C-terminal Extension (CTE) (**Fig. 3a**). The TEN domain is critical for maintaining RAP but is not required during the addition of a single telomere repeat [15,16]. In addition, the TEN domain has been shown to harbor an anchor site for telomere sequence upstream of the active site [15-17]. However, the exact nature of this anchor site and which residues are important for interacting with the telomere remain mostly unknown. The RBD is responsible for establishing RNA-protein interactions between TERT and TR and is absolutely essential for telomerase

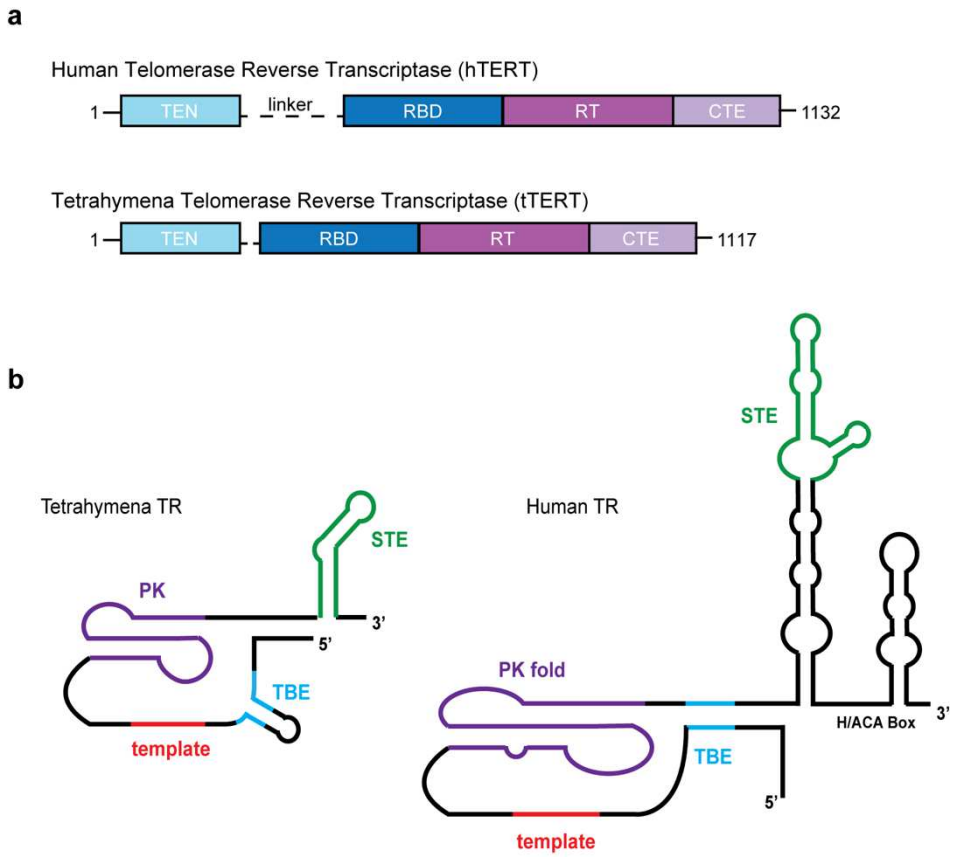


Figure 3. Telomerase components. (**a**) *Tetrahymena* and human TERT. The four conserved protein domains, TEN, RBD, RT and CTE are indicated in light blue, dark blue, purple and violet, respectively. Human TERT additionally has an extended linker region (space between TEN and RBD domains in hTERT). (**b**) *Tetrahymena* and human TR share common structural motifs, these include the template (red), Pseudoknot (PK, purple), Stem terminal element (STE, green) and Template Boundary Element (TBE, blue) although the exact nature of the human TBE remains unclear. The vertebrate specific H/ACA box is indicated for human TR.

activity [18-20]. The RT and CTE domains are homologous to the palm and thumb domain of canonical polymerases, respectively [21,22]. The RNA component of telomerase is also composed of several conserved structural domains, although the sequence and length of TR varies widely between species [23] (**Fig. 3b**). Two regions of

TR that appear to be generally conserved are the pseudoknot (PK)/template domain and the stem terminal element (STE) [24], both of which are required to support wild-type levels of telomerase function in vitro and in vivo [20,24-27]. While the functional role of the RNA template is self-evident, the contributions of the TR PK and STE remain the subject of continued investigation.

Another area that remains poorly understood is the interplay between telomerase and the nascent telomere during active telomere synthesis. How telomerase navigates along the newly synthesized telomere during active elongation remains largely uncharacterized. A large portion of this thesis focuses on addressing the interactions between TERT, TR and the nascent telomere.

In chapter 2 of this thesis I focus on how template boundary definition is established during multiple rounds of repeat addition. Telomeres are reverse transcribed using a hexameric sequence in the TR template, but how this sequence gets faithfully utilized repeatedly has remained an area of great interest. This interest is understandable considering that if any non-template sequence of TR gets reverse transcribed, the newly transcribed telomere will no longer properly anneal to the RNA template, resulting in dissociation and loss of telomerase activity. To study this problem, we chose to use the *Tetrahymena Thermophila* telomerase system due to the availability of years of biochemical data as well as the ability to in vitro express and purify certain individual TERT domains. Telomerase is well conserved between *Tetrahymena* and human (**Fig. 3**), making *Tetrahymena* an ideal model organism for telomerase studies. A decade of biochemical studies on *Tetrahymena* telomerase has

demonstrated that a region within TR called the Template Boundary Element (TBE), is essential for maintaining proper template boundary definition [18,28-31]. In addition, the TBE was shown to interact with the RNA Binding Domain (RBD) of TERT and was additionally shown to be in close proximity to the CP2 motif, one of the three conserved RNA binding motifs within RBD [32]. To complement the biochemical studies of this interaction, we set out to structurally characterize how the TBE and RBD physically interact. We succeeded in solving a crystal structure of the RBD/TBE complex, which highlighted the importance of previously determined critical RNA and protein residues for establishing both RNA/protein interactions as well as for enforcing template boundary definition. Our structure led to a model of successive scrunching and stretching of the RNA template during each repeat addition and suggested that the force of the stretch induced at the end of each repeat synthesis is too much to overcome without destabilizing critical protein/RNA interaction and is therefore energetically unfavorable, resulting in a pause at the end of the RNA template during each repeat addition. This model was further supported by single molecule analyses that demonstrated compaction and extension of the RNA template region during catalysis [33]. Whether this method of template boundary definition is conserved in human telomerase remains to be determined.

In chapter 3, I focus on how DNA structure formation within the nascent telomere affects the actively extending human telomerase complex. Telomeres are naturally G-rich, allowing for the formation of G-quadruplexes (GQs). G-rich sequences are common across the human genome and specifically occur at sites of

complex gene regulation [34]. How the potential to form GQs during active telomere elongation affects human telomerase has not been well characterized. Studies on ciliate telomerase have suggested that there may be GQ-induced telomerase dissociation during telomere elongation [35,36]. Additionally, studies on human telomerase have shown that telomerase has difficulty loading onto primers that terminate in a GQ [37], but how DNA structure formation in the nascent telomere affects an actively extending telomerase enzyme remains unknown.

Traditional telomerase assays are not optimized to investigate the effects of DNA structure on telomerase activity. These assays have long been performed with low micromolar amounts of radiolabeled dGTP together with much higher concentrations of dATP and dTTP. These unnatural nucleotide concentrations and stoichiometries introduce factors that affect the rate constants governing telomerase processivity and dissociation. The primary factor affecting RAP in these assays is the limiting dGTP concentration, which likely induces stalling or dissociation when the enzyme is waiting to incorporate this nucleotide. In contrast, our recent studies have been performed using equimolar and physiologically relevant nucleotide concentrations. This is possible by using a radiolabeled primer instead of radiolabeled dGTP, which otherwise inherently limits how much of this nucleotide can be used. When performing telomerase activity assays using equimolar dNTP concentrations, we observed product accumulation profiles that cannot be fit using an exponential function. While previous literature has estimated that the rate constants governing processivity and dissociation stay constant during each repeat addition resulting in a

strictly decreasing product accumulation decay profile, our results indicate that when a product length capable of forming a GQ is synthesized, both the dissociation rate and the forward rate increase. These results led us to believe that the formation of a GQ within the actively extending telomerase complex, can be both disruptive as well as favorable for telomerase activity. We propose a model where depending on where the terminal repeat is annealed when a GQ is formed and whether anchor site interactions are able to stay intact, either telomerase dissociation occurs or the enzyme is primed for another round of repeat addition.

In chapter 4, I discuss the ongoing work to study the dynamics of human telomerase during catalysis. First, I discuss both the progress and pitfalls for investigating TERT dynamics through site-specific labeling of various TERT domains and using smFRET to investigate potential conformational rearrangements during telomerase activity. I then discuss preliminary data from a collaboration with PacBio and the Puglisi lab at Stanford, where we are attempting to investigate the real-time kinetics of telomerase nucleotide incorporation. This work has provided interesting, although hard to interpret, preliminary results, and with time, will no doubt provide insight regarding the intricacies of the telomerase catalytic cycle.

References

1. de Lange, T. How shelterin solves the telomere end-protection problem. *Cold Spring Harb Symp Quant Biol* **75**, 167-77 (2010).
2. de Lange, T. Shelterin: the protein complex that shapes and safeguards human telomeres. *Genes Dev* **19**, 2100-10 (2005).
3. Olovnikov, A.M. [Principle of marginotomy in template synthesis of polynucleotides]. *Dokl Akad Nauk SSSR* **201**, 1496-9 (1971).
4. Watson, J.D. Origin of concatemeric T7 DNA. *Nat New Biol* **239**, 197-201 (1972).
5. Allsopp, R.C. et al. Telomere length predicts replicative capacity of human fibroblasts. *Proc Natl Acad Sci U S A* **89**, 10114-8 (1992).
6. Hayflick, L. The cell biology of aging. *Clin Geriatr Med* **1**, 15-27 (1985).
7. Blackburn, E.H. et al. Recognition and elongation of telomeres by telomerase. *Genome* **31**, 553-60 (1989).
8. Kim, N.W. et al. Specific association of human telomerase activity with immortal cells and cancer. *Science* **266**, 2011-5 (1994).
9. Feng, J. et al. The RNA component of human telomerase. *Science* **269**, 1236-41 (1995).
10. Nakamura, T.M. et al. Telomerase catalytic subunit homologs from fission yeast and human. *Science* **277**, 955-9 (1997).

11. Belfort, M., Curcio, M.J. & Lue, N.F. Telomerase and retrotransposons: reverse transcriptases that shaped genomes. *Proc Natl Acad Sci U S A* **108**, 20304-10 (2011).

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

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

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

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. Robart, A.R. & Collins, K. Human telomerase domain interactions capture DNA for TEN domain-dependent processive elongation. *Mol Cell* **42**, 308-18 (2011).

17. 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).

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

19. Jansson, L.I. et al. Structural basis of template-boundary definition in Tetrahymena telomerase. *Nat Struct Mol Biol* **22**, 883-8 (2015).

20. Huang, J. et al. Structural basis for protein-RNA recognition in telomerase. *Nat Struct Mol Biol* **21**, 507-12 (2014).

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

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

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

24. 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).

25. Cash, D.D. et al. Pyrimidine motif triple helix in the Kluyveromyces lactis telomerase RNA pseudoknot is essential for function in vivo. *Proc Natl Acad Sci U S A* **110**, 10970-5 (2013).

26. Kim, N.K. et al. Solution structure and dynamics of the wild-type pseudoknot of human telomerase RNA. *J Mol Biol* **384**, 1249-61 (2008).

27. 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).

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

29. 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).

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

31. 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).

32. 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).

33. 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).

34. Maizels, N. & Gray, L.T. The G4 genome. *PLoS Genet* **9**, e1003468 (2013).

35. Zahler, A.M., Williamson, J.R., Cech, T.R. & Prescott, D.M. Inhibition of telomerase by G-quartet DNA structures. *Nature* **350**, 718-20 (1991).

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

37. Hwang, H. et al. Telomeric overhang length determines structural dynamics and accessibility to telomerase and ALT-associated proteins. *Structure* **22**, 842-53 (2014).

CHAPTER 2: Structural basis for template boundary definition in Tetrahymena telomerase

(originally published in the journal Nature Structure and Molecular Biology)

Abstract

Telomerase is required to maintain repetitive G-rich telomeric DNA sequences at chromosome ends. To do so, the telomerase reverse transcriptase (TERT) subunit reiteratively uses a small region of the integral telomerase RNA (TER) as a template. An essential feature of telomerase catalysis is the strict definition of the template boundary to determine the precise TER nucleotides to be reverse transcribed by TERT. We report the 3 Å crystal structure of the *Tetrahymena* TERT RNA binding domain (tTRBD) bound to the template boundary element (TBE) of TER. tTRBD is wedged into the base of the TBE RNA stem-loop and each of the flanking RNA strands wraps around opposite sides of the protein domain. The structure illustrates how the tTRBD establishes the template boundary by positioning the TBE the correct distance from the TERT active site to prohibit copying of non-template nucleotides.

Introduction

Telomeres are essential nucleoprotein structures that help maintain genome integrity by differentiating natural chromosome ends from sites of DNA damage [1]. In most cells, a progressive shortening of telomere length with each round of cell division provides a molecular signal for cell aging and regulates entry into permanent cell growth arrest or apoptosis [2]. In contrast, cells possessing a high level of proliferative capacity (i.e. stem cells, inflammatory cells, and other self-renewing tissues) maintain telomere length through the enzymatic action of the specialized reverse transcriptase, telomerase[3]. The discovery of loss-of-function telomerase mutations in patients with diseases of the hematopoietic system (ie. Dyskeratosis Congenita and Aplastic Anemia) illustrates the requirement for telomerase in highly proliferative tissues [4]. On the other hand, aberrant activation of telomerase is deleterious, providing a mechanism for ~ 90% of human cancers to bypass the tumor suppressing activity of telomere shortening [5].

The minimal requirements for *in vitro* reconstitution of the active telomerase ribonucleoprotein (RNP) include a telomerase RNA (TER) and the telomerase reverse transcriptase (TERT) protein (**Fig. 1a,b**) [6]. The telomerase RNP directs *de novo* DNA synthesis at chromosome 3' ends utilizing a unique reverse transcription mechanism wherein an internal region of the telomerase RNA serves as the template (**Fig. 1c**) [7]. During telomere repeat synthesis, telomerase binds a single-stranded DNA substrate via Watson-Crick base-pairing with the RNA template and through additional TERT-DNA interactions termed 'anchor sites' [8-11]. Next, TERT catalyzes the synthesis of a

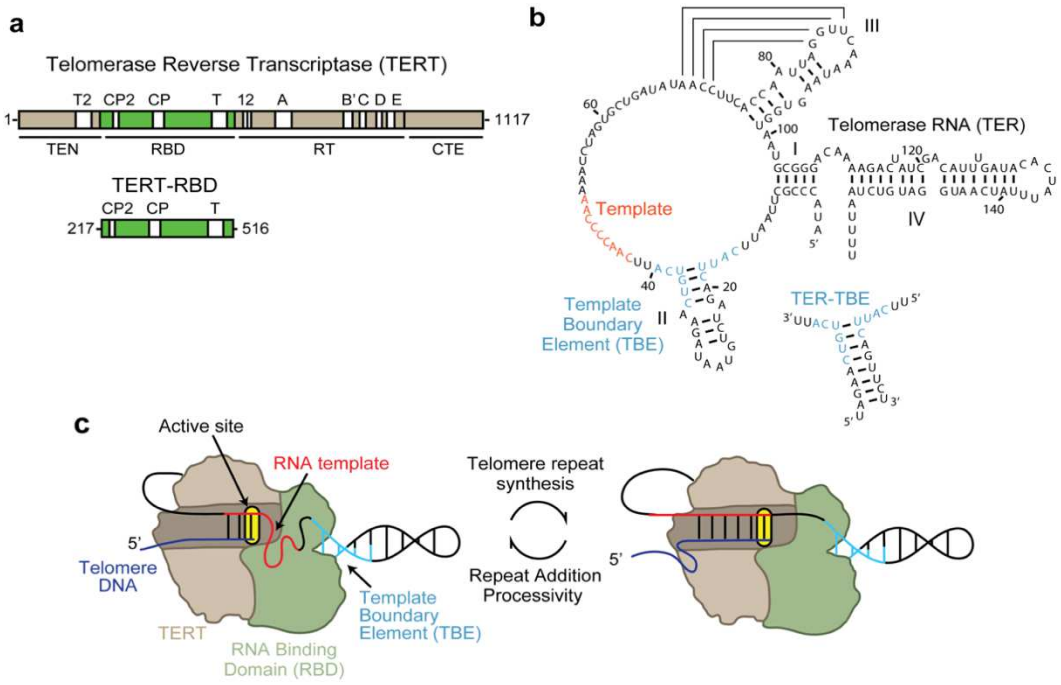


Figure 1. *Tetrahymena thermophila* telomerase composition and catalytic cycle. (**a**) Primary structure of the *T. thermophila* telomerase reverse transcriptase (TERT). The TEN, RT and CTE domains are shown in brown and the RNA binding domain (RBD) is shown in green. Conserved sequence motifs are shown as white boxes. The RBD construct used in this study is shown in green under the full-length TERT. (**b**) Secondary structure model of the *T. thermophila* telomerase RNA (TER). The RNA template is shown in red and the template boundary element (TBE) is shown in cyan (inset). The modified TBE RNA construct used in the present study. (**c**) Model for *T. thermophila* telomerase catalysis with the TERT RT domain in brown, RBD domain in green, the template RNA in red and the telomere DNA in blue.

telomere DNA repeat according to the sequence specified by the RNA template. Completion of a telomere repeat initiates telomerase RNP translocation to reposition the TERT active site, telomerase RNA, and the DNA substrate to the original DNA primer alignment configuration to allow for further rounds of repeat addition [12,13].The precise definition of a region of TER that may access the TERT active site and

template the synthesis of a telomere DNA repeat is a hallmark of telomerase function; however, the precise structural details of how this template boundary is defined have not been characterized.

TERTs from diverse organisms are highly conserved and share a common domain organization including: the essential N-terminal domain (TEN), RNA binding domain (RBD), reverse transcriptase domain (RT), and C-terminal extension (CTE) (**Fig. 1a**). The contribution of individual TERT domains to telomerase assembly and catalytic activity has been studied using *in vitro* deletion and mutagenesis studies, as well as single molecule FRET. The TEN domain makes contacts with both DNA and RNA and is essential for telomerase processivity [10,14-16]. RBD binds TER using several conserved regions known as the T-motif, CP-motif, and ciliate-specific CP2-motif (**Supplementary Fig.1** and **Supplementary Fig.2**) [17,18]. TERs are far more divergent with respect to both size and sequence, but maintain a conserved organization of multiple RNA elements including: a template boundary element (TBE), the template region, an RNA pseudoknot domain, and a distal RNA stem-loop (**Fig. 1B**) [19]. In telomerase from the model organism *Tetrahymena thermophila*, the base of stem-loop II and its flanking single-stranded RNA comprise the TBE of TER, and is proposed to interact with the T-, CP-, and CP2-motifs within the TERT-RBD [20,21]. Consistent with this notion, sequence mutations within the TBE or each of the conserved TERT-RBD motifs result in a loss of protein-RNA binding affinity, as well as template read-through defects when assayed in direct primer extension assays *in vitro* [17,18,20-22].

The high-resolution structure of the *Tetrahymena* TERT-RBD (tTRBD) domain revealed the organization of the T- and CP-motifs, which together form a putative RNA binding pocket [23]. However, the tTRBD structure did not include RNA and lacked the CP2-motif, which is essential for high affinity RNA interactions [21]. The more recently reported structures of TRBDs from diverse organisms indicate that this domain shares a common overall folding topology [24-26]. Moreover, the structure of the *Tribolium castaneum* (tc) TERT bound to a model RNA-DNA hybrid in the active site revealed the orientation of the RNA template with respect to the tcTRBD and supports the model that a high affinity RNA interaction in this region of TERT may establish the template boundary [11]. The first vertebrate TRBD structure was reported from the fugu fish *Takifugu rubripes* (tr), revealing the position of a conserved TFLY motif near the putative RNA binding region of trTRBD [25]. In this study, the TFLY motif was proposed to serve a similar role to the ciliate specific CP2-motif in promoting a high-affinity RNA interaction; however, the lack of RNA in the structure precluded a detailed description of the TRBD-RNA interaction network.

We set out to structurally determine the basis of template boundary definition in *Tetrahymena* telomerase. Here, we report the x-ray crystal structure of tTRBD bound to its cognate TBE RNA fragment. The structure highlights how the T-, CP-, and CP2-motifs within tTRBD cooperate to mediate interactions with the base of stem-loop II in *Tetrahymena* TER. The structure of tTRBD in complex with the TBE RNA explains previously reported biochemical mutagenesis experiments [17,18,20-22],

provides new insights into the mechanism of telomerase catalysis, and places novel constraints on the organization of the rest of the telomerase RNA subunit within the telomerase ribonucleoprotein (RNP) complex.

Results

Structure of TERT-RBD bound to the TBE RNA

To determine the molecular mechanism of template boundary definition in *Tetrahymena* telomerase we sought to develop protein and RNA constructs amenable to high-resolution structure determination. The previously reported structure of the tTRBD does not include the cognate telomerase RNA or the ciliate specific CP2 peptide [23]. Prior mutagenesis experiments established the importance of the CP2-motif for protein-RNA interactions that direct template boundary definition [20-22], and subsequent site-directed hydroxyl radical probing showed the CP2 peptide is in close spatial proximity to the TER TBE located at the base of stem-loop II [21]. Beginning with a tTRBD fragment (aa 195-516) that harbors multiple sites of RNA interaction including the CP2-motif [14,21], we made a set of truncation mutants to identify a highly soluble tTRBD domain (aa 217-516) that retains sequences required for specific high-affinity interaction with TER (**Fig. 1a**).

We optimized the TER stem-loop II construct by replacing the native A22-A34 bulge with a canonical U22-A34 base pair (**Fig 1b**). This base substitution is observed naturally in several closely related ciliate TERs [27], and does not impact *Tetrahymena* telomerase function when assayed *in vitro* and *in vivo* [20,28]. Indeed, full-

length TER containing the A22U substitution supported wild type levels of binding to tTRBD (**Supplementary Fig. 3a**), as well as catalytic activity and template boundary definition (**Supplementary Fig. 3b**). To improve the folding efficiency of the isolated RNA stem construct, we separated stem-loop II into two short oligonucleotides (**Fig 1b, inset**). The tTRBD protein and the optimized TER-TBE RNA were mixed together at a 1:1.5 stoichiometry to yield a mono-dispersed protein-RNA complex as judged by size exclusion chromatography (**Supplementary Fig. 3c).** We crystallized the tTRBD-TBE complex and solved the structure to 3 Å resolution by molecular replacement using the previously reported tTRBD structure as a search model (**Table 1**) [23]. The calculated electron density map permitted unambiguous modeling of most of the protein-RNA complex, with the exception of a small region of the tTRBD protein and several RNA nucleotides that are likely disordered in the crystal. The tTRBD-TBE complex crystallized in a 1:1 stoichiometry with two protein-RNA complexes (A and B) per asymmetric unit (**Supplementary Fig. 4a**). Complexes A and B demonstrated a high degree of overall similarity with an RMSD of 0.98 $Å^2$; however, regions of the protein-RNA binding interface in complex B were observed to be close to crystal packing contacts (**Supplementary Fig. 4b and 4c**.) and were substantially more disordered than in complex A, in which this same interface was far from any lattice contacts. We therefore focused our structural analyses on the more ordered complex A.

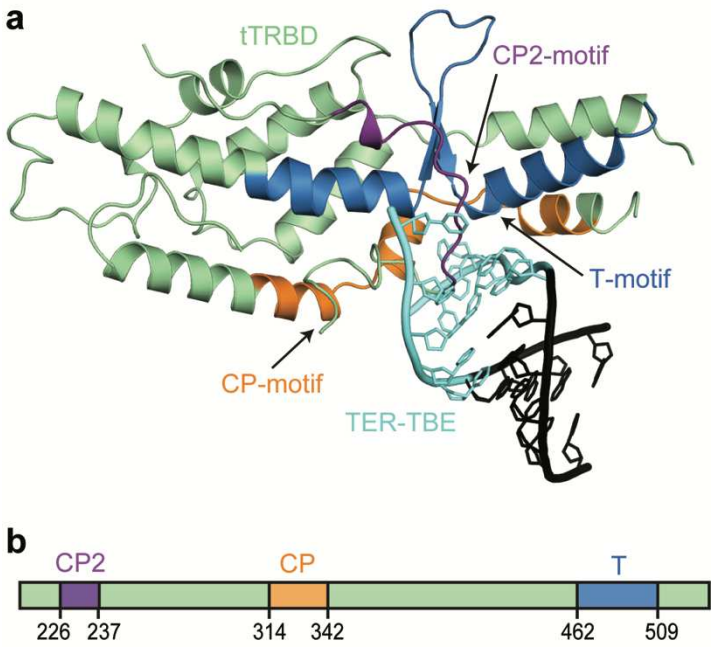


Figure 2. Structure of the *T. thermophila* TRBD (tTRBD)-TBE complex. (**a**) Cartoon representation of the tTRBD-TBE complex. tTRBD is shown in light green with the CP2-, CP-, and T-motifs shown in purple, orange and blue, respectively. The TBE is shown in cyan and the remaining stem II RNA in black. (**b**) Primary structure organization of tTRBD. Coloring scheme for each of the conserved sequence motifs is as described in **a**.

In the structure, the tTRBD is wedged into the base of the TER-TBE RNA stem II, diverting the paths of the 5'- and 3'-flanking strands in opposite directions (**Fig. 2a**). All protein-RNA contacts occur at the base of stem II, with the distal region of stem II (corresponding to the location of the native stem-loop II) extruding away from the protein interface. This arrangement is consistent with TER mutagenesis experiments and site-directed hydroxyl radical probing, which demonstrated that the base of stem II interacts with the tTRBD [21], as well as with the

observation that the distal region of stem-loop II is permissive to both sequence and structure modifications [20,28]. The three conserved sequence motifs within tTRBD (CP2, CP, and T) comprise the core RNA binding platform (**Fig. 2a,b**). Surprisingly, the majority of the protein-RNA contacts are not sequence specific with respect to TER, with a few notable exceptions (see below). Rather, the base of the TER-TBE is bound through an extensive network of polar contacts between the three conserved sequence motifs within tTRBD and the sugar-phosphate backbone of the RNA.

The CP- and T-motifs position the CP2-motif for RNA binding

The previously reported structure of the tTRBD (aa 254-519) revealed the organization of the conserved CP- and T-motifs [23]. Together, these regions of tTRBD form an electropositive groove that was hypothesized to represent a site of RNA interaction. However, our structure indicates that a major function of the T-CP pocket is to bind and orient the CP2 peptide (**Fig. 3a**), which is essential for RNA binding and was lacking in the previously crystallized construct [23]. The position of the CP2 peptide is fixed by a network of polar interactions between conserved amino acid side chains, as well as a number of backbone contacts. For example, the side chain of Arg226 in CP2 makes several hydrogen bonding interactions with the backbone of Phe408, Met411, Lys412 and Gln415, which together form a pocket that anchors the N-terminal region of the CP2 peptide within tTRBD (**Fig. 3b**). The

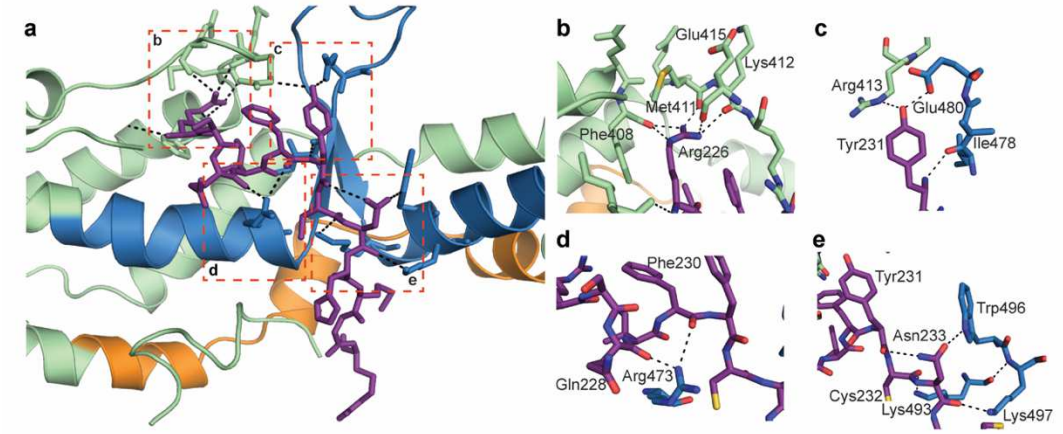


Figure 3. Protein-protein interactions between conserved tTERT RBD motifs. (**a**) Cartoon representation highlighting CP2 peptide (purple) interactions. Coloring scheme is same as in **Fig. 2a**. Interactions outlined by red dashed boxes are highlighted in panels **b-e**. (**b**) Arg226 makes polar contacts with the core of RBD. (**c**) Tyr231 interacts with the T-motif residue Glu480. (**d**) Gln228 and Phe230 make polar contacts with the T-motif residue Arg473. (**e**) Asn233 interacts with T-motif residues Trp496 and Lys497, as well as with CP2 peptide residue Tyr321.

hydroxyl group of Tyr231 in CP2 is hydrogen bonded to the guanidinium and carboxylate groups of Arg413 and Glu480, respectively (**Fig. 3c**). These polar contacts between Tyr231 and other regions of tTRBD induce a nearly 90^{o} kink in the protein backbone, redirecting the path of the CP2 peptide toward the base of the TBE RNA. The side chain of Arg473 in the T-motif makes hydrogen bonds with the backbone of CP2 residues Phe230 and Gln228 (**Fig. 3d**). In addition, the N1 position of the T-motif residue Trp496 indole group is hydrogen bonded with the carboxamide group of Asn233 in CP2. Finally, the terminal amino groups of two conserved lysines within the T-motif (Lys493 and Lys497) make polar backbone contacts with residues

Cys232 and Asn233 within the CP2 peptide (**Fig. 3e**). Thus, our structure of the tTRBD bound to the TBE RNA stem demonstrates that many of the conserved residues in the T-motif are indirectly required for RNA binding, and function to position the CP2 peptide in a geometry that is favorable for making RNA contacts.

tTRBD-TBE interactions define the template boundary

The TBE RNA stem forms an A-form helix (**Fig. 4a**), consistent with an earlier NMR structure of a model *Tetrahymena* stem-loop II sequence lacking the entire TBE region [29]. The crystal structure presented here reveals two critical RNA base pairs located at the interface between the tTRBD and the TBE RNA: a canonical Watson-Crick base pair between C19 and G37 and a previously unidentified wobble pair between U18 and U38 (**Fig. 4b** and **Supplementary Fig. 5**). The observation of these base pairing interactions and their central role in mediating RNA binding explains the high degree of sequence conservation observed for these nucleotide positions among ciliate telomerase RNAs [27]. Conserved residues from the CP2-, CP-, and T- motifs all contribute to the tTRBD surface that binds the TBE RNA (**Fig. 4c**). Within the CP2 peptide, the guanidinium group of Arg237 extends into the major groove of the TBE RNA stem, making hydrogen bonding contacts with the Hoogsteen face of G37 and with the carbonyl group of U38, while the imidazole group of His234 makes a stabilizing polar contact to the carbonyl of U38 (**Fig. 4c,d**).

In addition to interactions with the helical portion of the TBE RNA, tTRBD makes a large number of polar contacts between the 5'-flanking single-stranded RNA

and conserved T- and CP-motif side chains. Specifically, the Watson-Crick face of RNA nucleotide C15 makes

hydrogen bonds with the side chains of Arg492 (T-motif) and Asn324 (CP-motif), and the 2' hydroxyl group of the C15 ribose moiety contacts the terminal amino group of Lys332 (CP-motif) (**Fig. 4c,e**). The side chains of several additional CP-motif residues (Lys328, Tyr337, Gln338, and Lys341) make polar contacts with the RNA phosphate backbone (**Fig. 4c,e**). With respect to the 3' single stranded RNA flanking the base of the TBE RNA stem, only a single nucleotide (C39) could be unambiguously fit into the electron density. Although nucleotides A40-U42 and U13 are distant from any lattice contacts, they are likely disordered in the crystal and therefore could not be fit into the electron density map. Several tTRBD residues make polar contacts with the phosphate backbone in the 3' single stranded region, including the guanidinium group of Arg473 (T-motif). Thus, our structure reveals the sequence and structure specific interactions between the CP2-,CP-, and T-motif residues that mediate the protein-RNA interactions governing template boundary definition.

Conservation and function of tTRBD-TBE interactions

Many of the residues that participate in orienting the CP2 peptide, or that interact directly at the tTRBD-TBE binding interface (**Fig. 5a,b**), are highly conserved among ciliate TRBDs (**Supplementary Fig. 2**). The positions of these amino acids in our structure are in accord with this sequence conservation, as well as

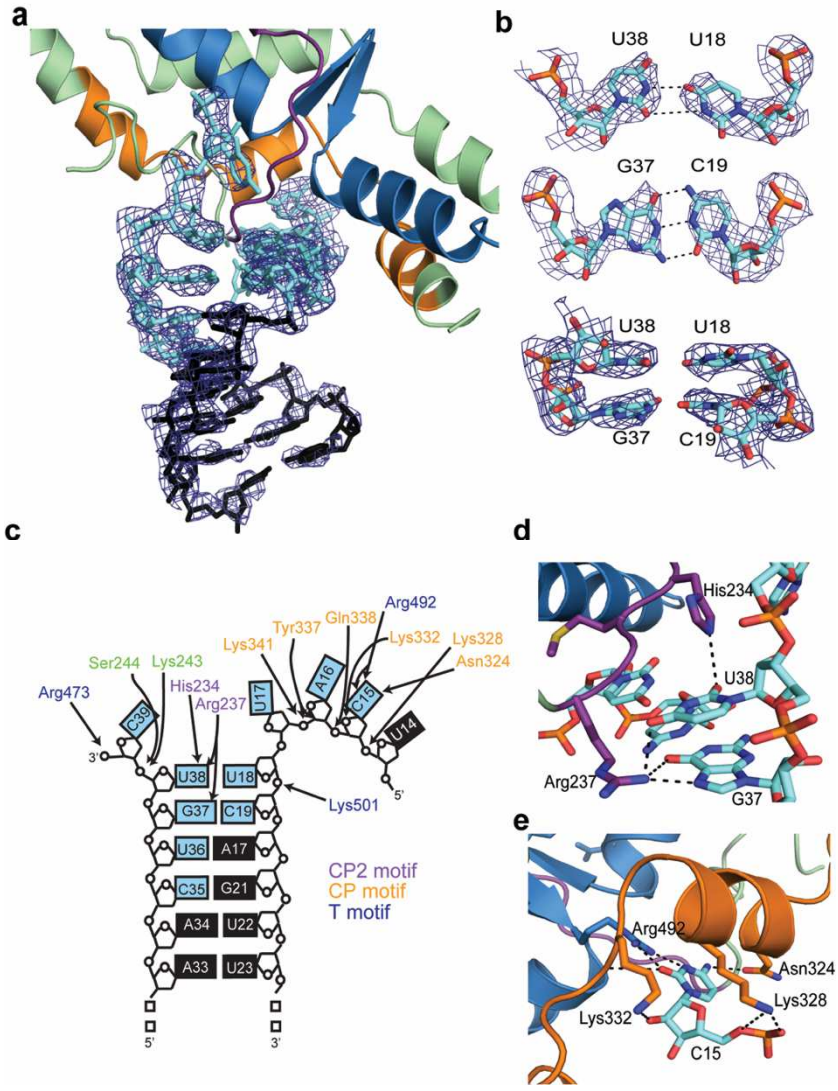


Figure 4. Protein-RNA interactions within the tTRBD-TBE complex. (**a**) Simulated annealing omit map of the TER-TBE RNA at 2.0σ contour level superimposed on cartoon model of tTRBD-TBE structure. (**b**) Simulated annealing omit map of the U38-U18 and G37-C19 base pairs at contoured to 1.5σ. The bottom panel shows a side view of the stacked U23-U18 and G37-C19 base pairs. (**c**) Schematic diagram of tTRBD-TER TBE interactions. The TBE RNA nucleotides are colored in cyan and the rest of the RNA in black. tTRBD residues from CP2-, CP-, and T-motif residues are shown in purple, orange and blue, respectively. Arrows indicate sites of polar interactions. (**d**) Arg237 interacts with nucleotides U38 and G37 at the base of the TBE RNA stem. His234 also interacts with U38. (**e**) The TBE RNA nucleotide C15 fits into a polar pocket between the CP- and T-motifs. C15 makes polar contacts with T-motif residue Arg492 and CP-motif residue Asn324, Lys332 and Lys328.

with the observation that alanine substitution at each of these sites can induce defects in RNA binding and/or telomerase activity [17,20-22]. Notably, the two T-motif residues Glu480 and Arg492 are invariant among TRBDs from ciliates, yeasts, and vertebrates (**Supplementary Fig. 1**). The side chain of Arg492 in the T-motif makes base specific contacts with C15 in the TBE RNA (**Fig. 5b**), raising the possibility that this residue may be directly involved in RNA binding interactions in TRBDs from other organisms. A second invariant T-motif residue, Glu480, is hydrogen bonded to the side chain of Tyr231 in CP2 (**Fig. 5a**). Interestingly, the recently reported structure of the *Takifugu rubripes* TRBD identified the TFLY-motif, which is conserved among vertebrate TRBDs [25]. Comparison of the sequence of the CP2- and TFLY-motifs suggests that certain residues appear to be conserved (**Fig. 5c**). In order to ascertain whether the ciliate specific CP2-motif is structurally homologous to the vertebrate TFLY-motif, we aligned our structure with the two available vertebrate TRBD structures from *Takifugu rubripes* (tr) and the closely related Medaka fish *Oryzias latipes* (ol) (**Fig. 5c**). Indeed, in all three TRBD structures, the invariant T-motif glutamate side chain is well aligned, while Tyr321 from the olTRBD TFLY-motif is overlapping with Tyr231 from the tTRBD CP2-motif (**Fig. 5d**). In contrast, the corresponding tyrosine side chain from trTRBD (Tyr305) is significantly displaced. This alternative positioning of the TFLY-motif in the trTRBD structure is likely due to the fact that this region is directly involved in making crystal-packing contacts [25]. Further comparison of the tTRBD CP2-motif and the olTRBD TFLY-motif

structures suggests additional regions of homology. For example, the sidechains of Phe230 in tTRBD and Phe319 in olTRBD are in close proximity in the structural alignment. Lastly, two conserved basic tTRBD CP2 residues (His234 and Arg237) that comprise a sequence specific RNA binding patch are also present within olTRBD (Arg324 and His327), albeit with an inverted sequence order (**Fig. 5e**). Taken together, the results of our structural analysis suggest that the ciliate specific CP2-motif is structurally homologous to the vertebrate TFLY-motif, which may serve a similar function to mediate sequence specific protein-RNA interactions.

A molecular model of template boundary definition

In order to further understand the mechanism of template boundary definition in telomerase, we aligned the tTRBD-TBE structure with the complete TERT structure obtained from the flour beetle *Tribolium castaneum* bound to a model RNA-DNA hairpin in the active site [11]. The resulting model demonstrates the relative positions of the TBE and the template RNA within the telomerase complex (**Fig. 6a**). In this model, residue C39 is positioned ~17 Å from the nearest template RNA residue, roughly corresponding to the contour length of the three RNA nucleotides not present in the structural model. Previous work has demonstrated that the TBE is a site of protein-RNA interaction that is critical for preventing entry of non-template residues into the TERT active site [20,21]. The structural model we propose now directly demonstrates the mechanism by which this is accomplished.

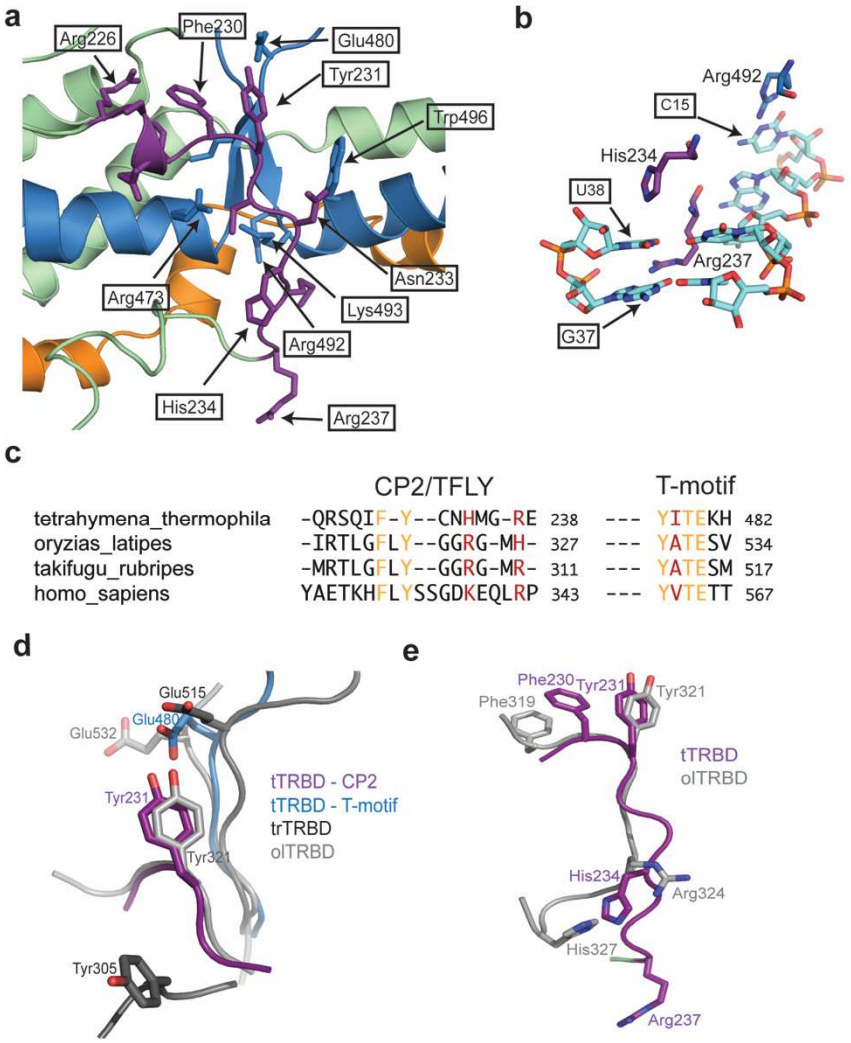


Figure 5. Conservation and function of observed interactions in the tTRBD-TBE complex. **(a)** Conserved residues in the CP2- and T-motifs that are important for RNA binding and/or telomerase function. Boxed residues represent CP2 residues that are conserved across ciliates and T-motif residues that are conserved across species. **(b)** Conserved RNA residues that are critical for RBD binding. Boxed nucleotides represent nucleotides that are conserved across ciliates and that have previously been demonstrated to be important for RBD binding. **(c)** CP2/TFLY alignment. The CP2 motif of *Tetrahymena* was aligned manually to the TFLY motif of vertebrates. Invariant residues are shown in orange and residues with conserved chemical properties are shown in red. A segment of the conserved T-motif alignment is also shown. **(d)** Structural alignment of tTRBD, trTRBD and olTRBD. tTRBD CP2-motif and T-motif are shown in purple and blue, respectively. trTRBD and olTRBD are dark grey and light grey, respectively (PDB 4LMO[25]) (PDB 4O26[26]). **(e)** Structural alignment of the tTRBD CP2 motif and the olTRBD TFLY motif. Color scheme is same as in **d.**

The location of the tTRBD-TBE complex is precisely positioned to prevent entry of non-template residue U42 into the active site of the enzyme, thereby establishing the template boundary.

Discussion

We report the structure of the tTRBD in complex with the TBE of TER. The position of stem II with respect to TERT observed in our high-resolution structure is consistent with the proposed location of this same RNA element in a structural model derived from a recent 25 Å cryo-EM *Tetrahymena* telomerase holoenzyme structure [30]. The most salient features of the tTRBD-TBE structure are the interactions between three conserved sequence motifs (CP2, CP, and T) and the manner in which these motifs cooperate to position the TBE RNA with respect to tTRBD. Previously reported biochemical experiments support a model in which a high-affinity interaction between tTRBD and the TBE RNA establishes template boundary definition by limiting the amount of RNA that may access the active site of the tTERT reverse transcriptase domain [17,18,20,21]. More recent bulk biochemical and single-molecule experiments demonstrated that the RNA template region of TER exists in either a compressed or extended conformation, depending on the stage of the telomere repeat synthesis reaction [13]. This finding supports the notion that the tTERT reverse transcriptase domain exerts a stretching force on the RNA template during telomere DNA synthesis, and that a point of protein-RNA contact must resist this

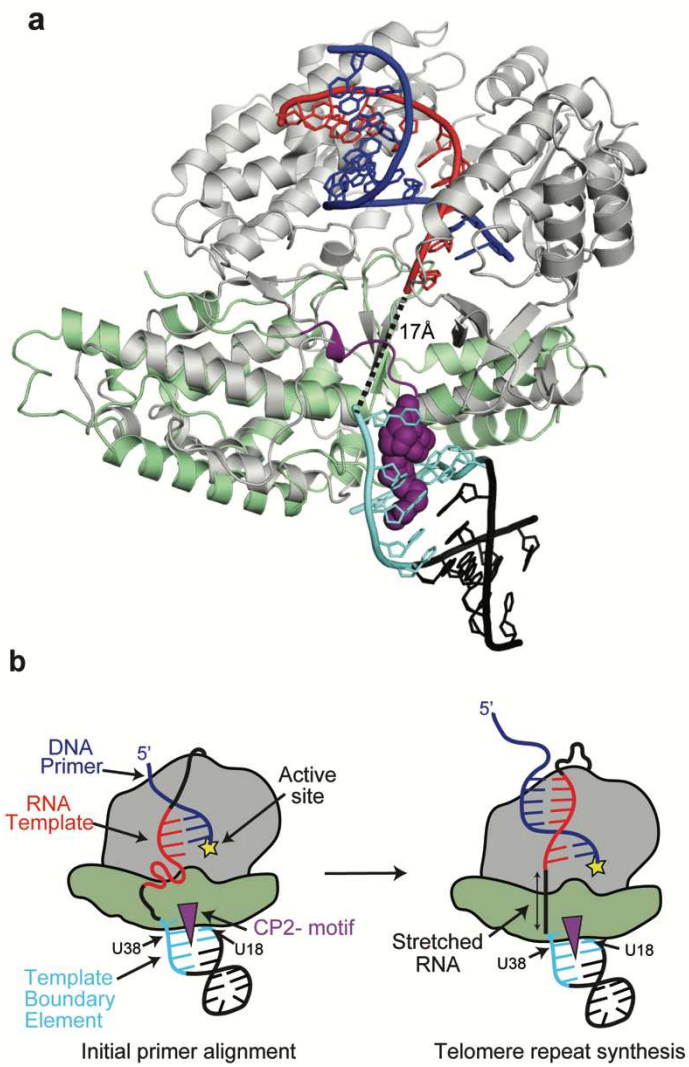


Figure 6. Structural model of *T. thermophila* telomerase RNA template connectivity to the TER template boundary element. (**a**) Structure of the tTRBD (light green) bound to the TER-TBE RNA (cyan and black) aligned to the *T. castaneum* TERT structure (grey) bound to a model RNA template (red) and DNA primer (blue) (PDB 3KYL[11]). Dashed line indicates distance (~17 Å) between the 3' terminus of the TER-TBE and the 5' terminus of the RNA template in model. (**b**) Cartoon model for template boundary definition in *Tetrahymena* telomerase. The tTRBD is shown in green, the TERT active site as a yellow star, the DNA primer in blue, the RNA template in red, the TBE RNA in cyan, and the CP2 peptide as a purple triangle.

pulling force to establish the template boundary. In our tTRBD structure bound to the TBE RNA, the protein domain is wedged against the base of the TBE RNA stem, making several sequence-specific contacts in the RNA major groove (**Fig. 6a**). This arrangement is well suited to provide mechanical stability against the reverse transcriptase pulling force (**Fig. 6b**), since both protein-RNA and RNA-RNA contacts would need to be disrupted to permit more RNA to enter the reverse transcriptase active site.

Further structural analysis is required to determine whether the mechanism of template boundary definition revealed by the tTRBD-TBE structure is a general feature found in telomerases from other organisms. The P1 stem found in vertebrate telomerase RNAs has been proposed to serve a functionally analogous role to *Tetrahymena* TER stem-loop II [31]. However, more recent experiments demonstrate that the sequence of the vertebrate telomerase RNA template can itself govern the location of the template boundary [32]. A recent study of the *Takifugu rubripes* TRBD structure identified a conserved protein motif (the TFLY-motif) present in vertebrate TERTs [25]. The TFLY-motif is positioned in the *Takifugu* TRBD structure nearby the position of the CP2-motif in the *Tetrahymena* TRBD-TBE structure, indicating that the TFLY-motif may be the vertebrate homolog of the CP2-motif in ciliates. An alignment between TRBD structures from *Takifugu rubripes*, the closely related *Oryzias latipes* [26], and our *Tetrahymena* TRBD structure demonstrates a significant overlap in the position between the CP2-motif and the vertebrate TFLY-motifs (**Fig. 5d,e**). Indeed, Tyr231in the tTRBD structure is nearly identical in position to Tyr321

in the olTRBD, suggesting that the universally conserved tyrosine in the vertebrate TFLY-motif is the functional homolog of the tyrosine in the CP2-motif. This discovery should motivate future structural studies of vertebrate TRBDs bound to their cognate template proximal RNA fragments in order to determine the molecular basis of template boundary definition in vertebrates.

Table 1. Data collection and refinement statistics (molecular replacement)

	RBD-TBE
Data collection	
Space group	P212121
Cell dimensions	
a, *b*, *c* (Å)	58.53, 117.77, 131.18
α, β, γ (°)	90.00, 90.00, 90.00
Resolution (Å)	41.51(3.00) *
R_{merge}	0.231(0.969)
$I / \sigma I$	7.8(2.1)
Completeness (%)	99.5(99.3)
CC(1/2)	0.990(0.720)
Redundancy	6.8(6.8)
Refinement	
Resolution (Å)	41.51-3.00
No. reflections	18649
R_{work} / R_{free}	0.239/0.290
No. atoms	4793
Protein	4767
Ligand/ion	2
Water	24
B-factors	
Protein	62.90
Ion	65.50
Water	43.80
R.m.s. deviations	
Bond lengths (Å)	0.003
Bond angles (°)	0.82

One crystal was used. * Values in parenthesis are for highest-resolution shell.

Materials and methods

RBD – TBE complex formation.

His-tagged RBD (217-516) was expressed in *E.coli* BL21-CodonPlus (DE3)-RP competent cells (Agilent) and purified by nickel exchange chromatography. RBD was further purified by size exclusion chromatography (Superdex 200 column). The TBE RNA construct was prepared by annealing two individual strands, each containing one side of Stem II (**Fig. 1b**). (Strand A: 5' UUCAUUCAGUUCU 3' Strand B: 5' UAGAACUGUCAUU 3'). To anneal the construct, the two strands were heated to 95^{o}C for 3 minutes in annealing buffer (500 mM NaCl, 20mM Tris-HCl pH 7.8) and then allowed to slowly cool to room temperature. After cooling to room temperature $MgCl_2$ and DTT were each added to a final concentration of 1mM. The tTRBD protein was mixed with the annealed TER-TBE RNA at a stoichiometry of 1:1.5 in high salt buffer (500mM NaCl, 20mM Tris-HCl pH 7.8, 1mM $MgCl_2$ and 1mM DTT). This mixture was then dialyzed overnight (~16 hours) into low salt buffer (100mM NaCl, 20mM Tris-HCl pH 7.8, 1mM MgCl2 and 1mM DTT). Formation of protein-RNA complex was analyzed using diagnostic sizing chromatography on a Superdex 200 column (**Supplemental Fig. 3c**) prior to setting crystallization trays.

Protein crystallization and data collection.

tTRBD-TBE RNA crystals that diffracted to 3Å appeared overnight and grew to final size of ~10-20 microns after 4-5 days. Crystals were grown by the sitting-drop vapor-diffusion method using the ARI GRYPHON in Hampton Research Natrix HT

condition A5 (200mM KCl, 10mM MgCl2, 50mM MES monohydrate pH 5.6, 4% PEG 8000) at a 3:1 ratio of sample to screen. Crystals were transferred into cryoprotectant solution consisting of 20% ethylene glycol, 160mM KCl, 8mM MgCl2, 40mM MES monohydrate pH 5.6, 3.2% PEG 8000 at 4^{o}C. Crystals were harvested by flash freezing in liquid nitrogen. X-ray data was collected at the ALS beamline 501 at a wavelength of 1Å and a temperature of 100K. The data was processed using MOSFILM. Crystals are in the P212121 space group and there are two tTRBD-TBE complexes in the asymmetric unit.

Structure determination and refinement.

Phases were calculated in CCP4 by molecular replacement using the previously solved Tetrahymena RBD structure (PDB 2R4G [23]) as the initial search model.. Model building was performed in Coot and refined using PHENIX. The structure was refined to good stereochemistry with 93% of residues in the most favored region, 6% in the allowed region and 1% outliers of the Ramachandran plot.

Supplementary Methods

Electrophoretic Mobility Shift Assay (EMSA)

WT TER and TER A22U were transcribed in vitro using T7 RNA polymerase and PCR templates. The RNA was gel purified, CIP and DNase treated and ethanol precipitated. The RNA was then end-labeled with radiolabeled ATP. RNA was heated to 60°C and let cool to room temperature on bench. RDB was diluted into binding

buffer containing 10% glycerol, 20mM Tris pH 8.0, 100mM NaCl, 1mM $MgCl_2$ and 1mM DTT. Each binding reaction contained the indicated amount of tTRBD plus 0.4 nM end-labeled RNA in binding buffer containing 10% glycerol, 20mM Tris pH 8.0, 100mM NaCl, 1mM $MgCl_2$, 1mM DTT, 0.1mg/ml BSA, 0.1mg/ml yeast tRNA and 0.1ul RNasin. Reactions were incubated for 20 min at 30°C. Samples were run on a 5% native acrylamide gel (37.5:1 acrylamide: bis acrylamide, 4% glycerol, 0.5X TBE) and run for 3h at 4°C at 200V. The gel was dried and imaged overnight followed by PhosphorImager analysis.

Primer extension assay.

Telomerase for *in vitro* primer extension assays was prepared in RRL (Promega) as described previously [13]. 2 μL of the RRL telomerase reconstitution reaction was added to 2 μM DNA primer $(GGGGTT)_3$, 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. To test for template boundary defects, reactions were performed in the presence or absence 100 μM dATP. Reactions were incubated for 1 hour at 30^{o}C and then phenol:chloroform extacted and ethanol precipitated. Extension products were resolved on a 12% PAGE DNA sequencing gel and imaged using a Typhoon scanner with a phosphor imaging screen.

Accession codes.

Coordinates and structure factors have been deposited in the Protein Data Bank under accession code 5C9H.

References

1. Palm, W. & de Lange, T. How shelterin protects mammalian telomeres. *Annu Rev Genet* **42**, 301-34 (2008).
2. Allsopp, R.C. et al. Telomere length predicts replicative capacity of human fibroblasts. *Proc Natl Acad Sci U S A* **89**, 10114-8 (1992).
3. 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).
4. Vulliamy, T.J. & Dokal, I. Dyskeratosis congenita: the diverse clinical presentation of mutations in the telomerase complex. *Biochimie* **90**, 122-30 (2008).
5. Kim, N.W. et al. Specific association of human telomerase activity with immortal cells and cancer. *Science* **266**, 2011-5 (1994).
6. Licht, J.D. & Collins, K. Telomerase RNA function in recombinant Tetrahymena telomerase. *Genes Dev* **13**, 1116-25 (1999).
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. Lue, N.F. A physical and functional constituent of telomerase anchor site. *J Biol Chem* **280**, 26586-91 (2005).

9. 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).

10. 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).

11. 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).

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

13. 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).

14. 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).

15. 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).

16. Akiyama, B.M., Parks, J.W. & Stone, M.D. The telomerase essential N-terminal domain promotes DNA synthesis by stabilizing short RNA-DNA hybrids. *Nucleic Acids Res* (2015).

17. 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).

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

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

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

21. 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).

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

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

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

25. Harkisheimer, M., Mason, M., Shuvaeva, E. & Skordalakes, E. A motif in the vertebrate telomerase N-terminal linker of TERT contributes to RNA binding and telomerase activity and processivity. *Structure* **21**, 1870-8 (2013).

26. Huang, J. et al. Structural basis for protein-RNA recognition in telomerase. *Nat Struct Mol Biol* **21**, 507-12 (2014).

27. McCormick-Graham, M. & Romero, D.P. Ciliate telomerase RNA structural features. *Nucleic Acids Res* **23**, 1091-7 (1995).

28. Cunningham, D.D. & Collins, K. Biological and biochemical functions of RNA in the tetrahymena telomerase holoenzyme. *Mol Cell Biol* **25**, 4442-54 (2005).

29. 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).

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

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

32. Brown, A.F. et al. A self-regulating template in human telomerase. *Proc Natl Acad Sci U S A* **111**, 11311-6 (2014).

Supplementary information

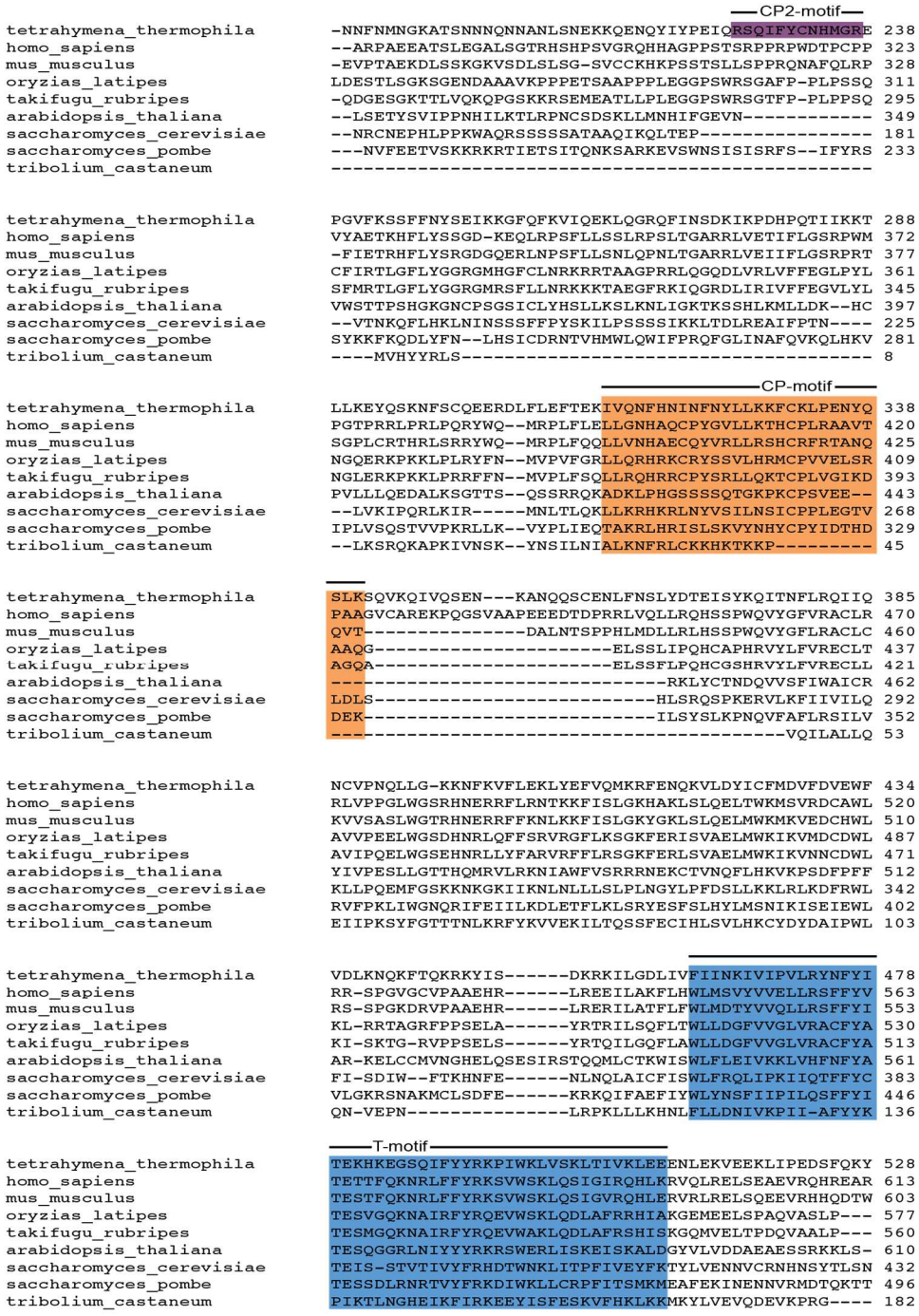


Figure S1. Cross species TERT-RBD sequence alignments. The ciliate-specific CP2 motif is shown for tetrahymena only (purple). The conserved CP- and T-motifs are shown in orange and blue, respectively.

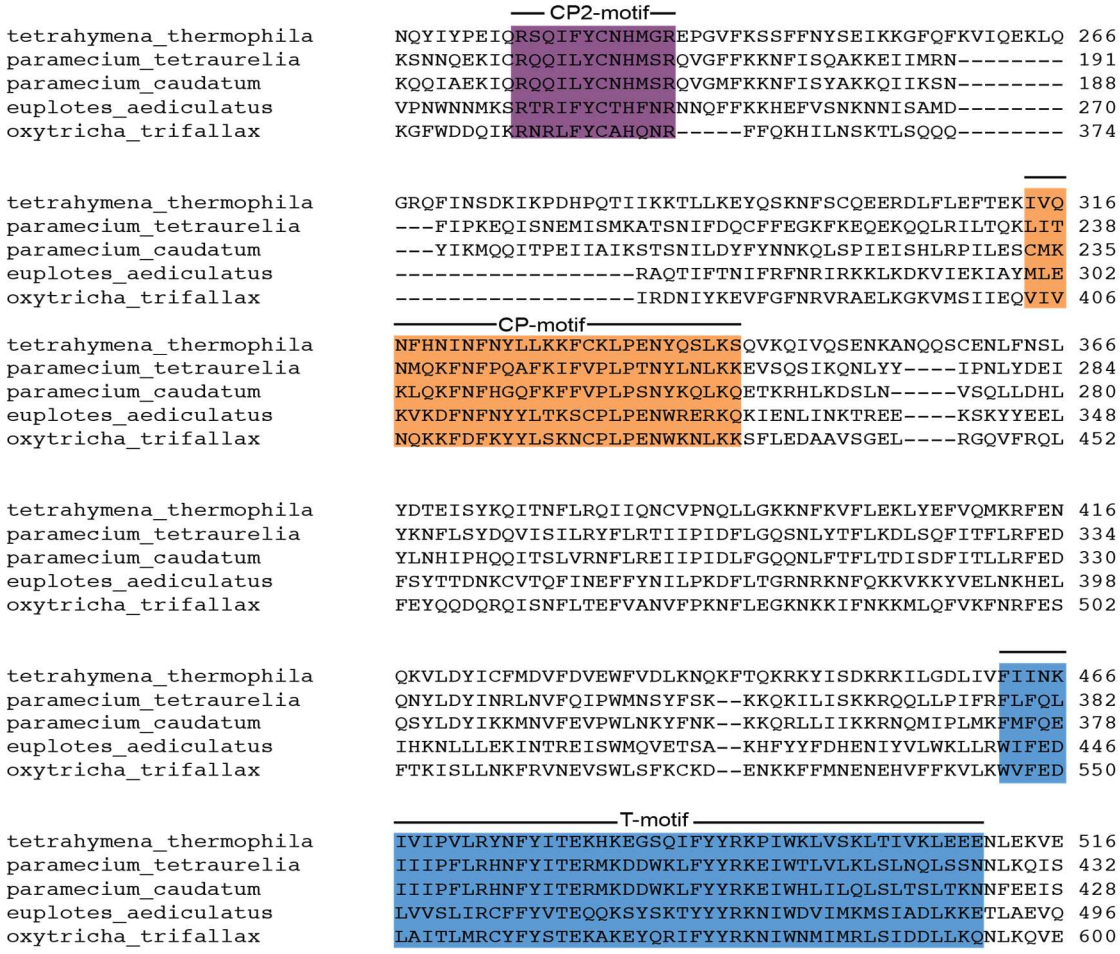


Figure S2. Ciliate TERT-RBD sequence alignments. The ciliate specific CP2 motif is shown in purple. The conserved CP- and T-motifs are shown in orange and blue, respectively.

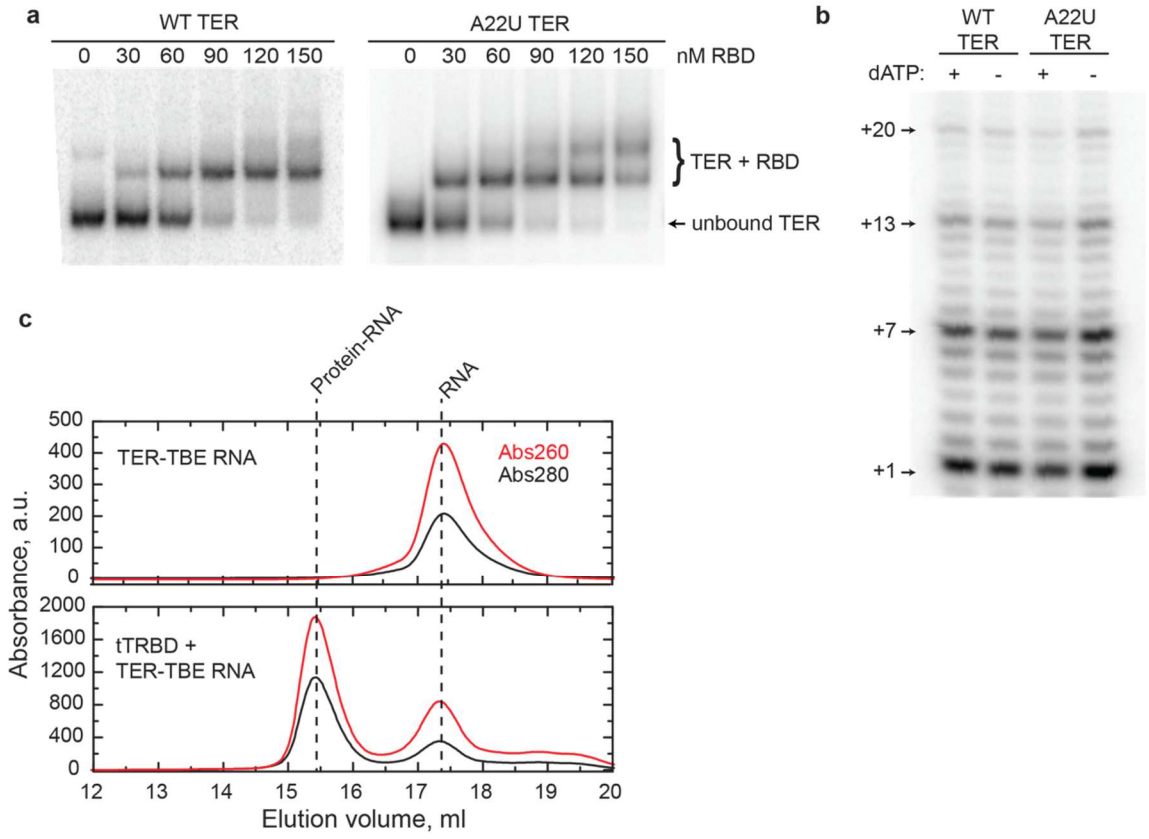


Figure S3. RNA binding activity of new tTRBD protein construct. **(a)** Electrophoretic Mobility Shift Assay (EMSA) of wild type (WT) TER (left panel) and A22U TER (right panel) bound to RBD. **(b)** Primer extension assay of WT and A22U TER in the presence or absence of dATP. Numbers on the left indicate number of nucleotides extended of the primer. A defect in template boundary definition would be expected to arise in the presence of ATP since TER nucleotide U42 would enter the TERT active site as has been shown previously[1,2]. **(c)** FPLC elution profiles of the TBE RNA construct alone (top panel) and of the tTRBD-TBE complex (bottom panel). Red and black lines indicate absorbance measured at 260nm and 280nm, respectively. The tTRBD-TBE complex elutes as a single peak around 15.5 ml and free TBE RNA elutes around 17.4 ml. The samples were run through a superdex 200 column.

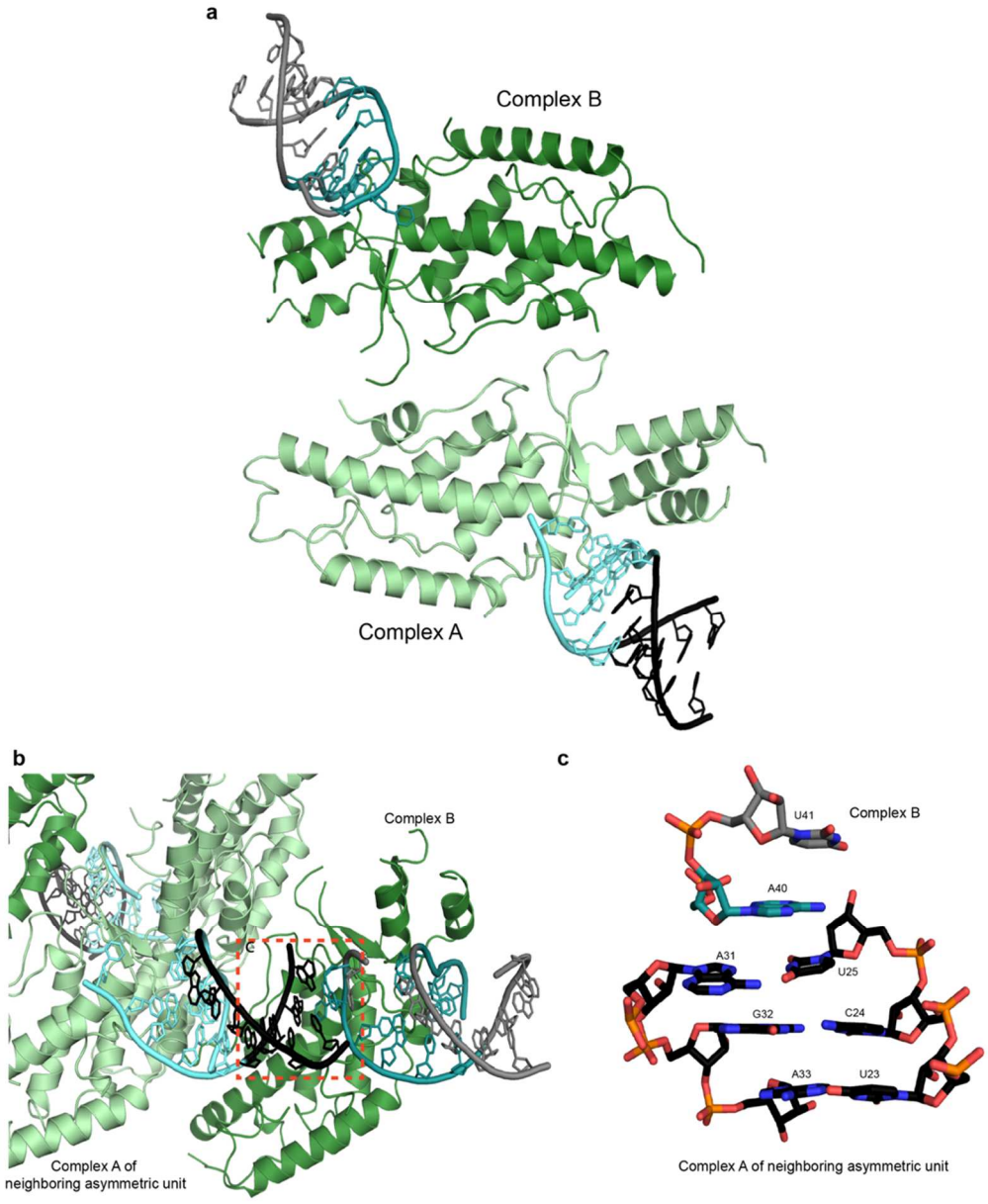


Figure S4. Crystal packing arrangement of the tTRBD-TBE RNA complex. **(a)** There are two protein-RNA complexes in the asymmetric unit. Complex A (light green, cyan and black) is better fit to the electron density than complex B (dark green, teal and grey) and is therefore the molecule chosen for structure determination. **(b)** Crystal contacts are mediated by RNA-RNA base stacking between complex B and complex A of the neighboring asymmetric unit. This interaction is further highlighted in **c** (dashed red box). **(c)** Base stacking interactions between complex B and complex A of the neighboring asymmetric unit. Residues A40 and U41 of complex B are twisted outward compared to the same residues in complex A to mediate crystal contacts with the distal region of stem II of complex A.

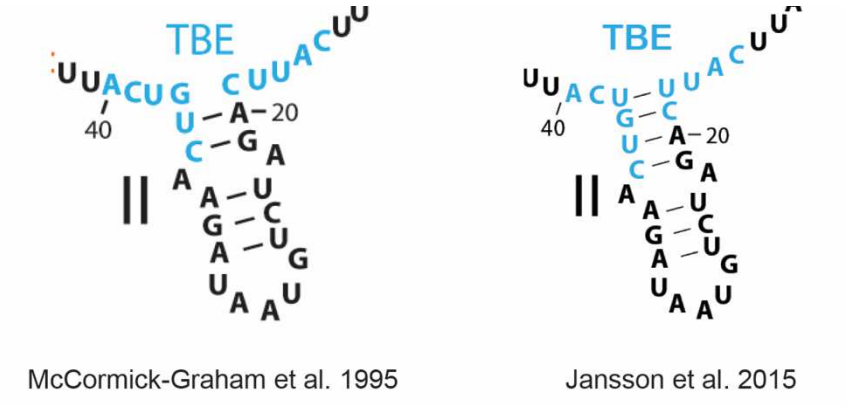


Figure S5. Comparison of TER secondary structure. The left panel shows the secondary structure of TBE-Stem II as determined by McCormick-Graham et al[3]. The right panel shows the updated TER secondary structure based on the observations of the RNA density in our structure.

Supplementary References

1. Lai, C.K., Miller, M.C. & Collins, K. Template boundary definition in Tetrahymena telomerase. *Genes Dev* **16**, 415-20 (2002).
2. 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).
3. McCormick-Graham, M. & Romero, D.P. Ciliate telomerase RNA structural features. *Nucleic Acids Res* **23**, 1091-7 (1995).

CHAPTER 3: Telomere DNA G-quadruplex folding within actively extending human telomerase

(submitted to PNAS)

ABSTRACT

Telomerase maintains telomere length by reverse transcribing short G-rich DNA repeat sequences from its internal RNA template. G-rich telomere DNA repeats readily fold into G-quadruplex (GQ) structures in vitro, and the presence of GQ-prone sequences throughout the genome introduces challenges to replication in vivo. Using a combination of ensemble and single-molecule telomerase assays we discovered that GQ folding of the nascent DNA product during processive addition of multiple telomere repeats modulates the kinetics of telomerase catalysis and dissociation. Telomerase reactions performed with telomere DNA primers of varying sequence or using K^+ versus Li^+ salts yield changes in DNA product profiles consistent with formation of GQ structure within the telomerase-DNA complex. Single-molecule FRET experiments reveal complex DNA structural dynamics during real-time catalysis, supporting the notion of nascent product folding within the active telomerase complex. To explain the observed distributions of telomere products, we fit telomerase time series data to a global kinetic model that converges to a unique set of rate constants describing each successive telomere repeat addition cycle. Our results highlight the potential influence of the intrinsic folding properties of telomere DNA during telomerase catalysis and provide a detailed characterization of GQ modulation of polymerase function.

SIGNIFICANCE

Telomeres protect the ends of linear chromosomes from illicit DNA processing events that can threaten genome stability. Telomere structure is built upon repetitive G-rich DNA repeat sequences that have the ability to fold into stable secondary structures called G-quadruplexes (GQs). In rapidly dividing cells, including the majority of human cancers, telomeres are maintained by the specialized telomerase enzyme. Thus, telomerase and its telomere DNA substrate represent important targets for developing novel cancer drugs. In this work, we provide evidence for GQ folding within the newly synthesized DNA product of an actively extending telomerase enzyme. Our results highlight the delicate interplay between the structural properties of telomere DNA and telomerase function.

INTRODUCTION

Telomeres safeguard the ends of chromosomes from illicit DNA processing events that would otherwise threaten genome stability [1,2]. The foundation of telomere structure consists of short G-rich DNA sequence repeats. The majority of telomere DNA is double-stranded and can be up to several kilobases in length, while the ends are processed to terminate with a short 3' single-stranded G-rich tail (~50-150 nucleotides in vertebrates) [3,4]. Repetitive G-rich DNA sequences are not unique to telomeres and are found throughout the human genome [5]. These G-rich repeats have the capacity to fold into G-quadruplexes (GQs), structures composed of multiple Hoogsteen bonded G-quartet motifs that stack together to yield stable DNA folds [6,7].

GQ folding has been implicated in a variety of biological processes. For example, replication of GQ-prone sequences is problematic and requires contributions from specific DNA helicase enzymes to avoid replication-coupled DNA damage [8-10]. Sequences with GQ-folding potential are enriched within promoter sequences of oncogenes where they are thought to regulate gene expression [11]. Finally, recent evidence suggests GQ folds can form in vivo in a spatially and temporally regulated manner [12-14]. Thus, small molecules that bind and stabilize GQ-folds hold promise as novel cancer drugs; a fact that motivates better understanding of how GQ structure can modulate enzyme function.

Telomerase is an RNA-dependent DNA polymerase that is uniquely adapted to synthesizing G-rich repetitive DNA sequences [15,16]. Telomerase activity combats gradual telomere shortening that occurs with each round of cellular division [17]. Left unchecked, telomere shortening induces senescence or cell death in somatic tissues. In contrast, proliferative cells such as stem cells rely upon telomerase activity to maintain telomeres in order to support continued rounds of cell division [15]. Genetically inherited loss of function mutations in telomerase subunits cause human disorders characterized by deterioration of proliferative tissue types [18-21]. On the other hand, telomerase overexpression contributes to the immortal phenotype of ~90% of human cancers, and is therefore an important target for development of novel cancer therapies [22].

Telomerase is a ribonucleoprotein (RNP) complex that includes the long non-coding telomerase RNA (TR) and the catalytic telomerase reverse transcriptase

(TERT) protein subunit [23,24]. In vertebrates, the 3' end of TR includes canonical H/ACA box RNA motifs and associated proteins that promote TR stability and efficient telomerase RNP biogenesis [25]. To initiate telomerase catalysis, the 3' ssDNA telomeric tail base pairs with the TR template, forming a short RNA-DNA hybrid that is extended in the TERT active site (Fig. 1A). TERT utilizes a limited region of TR to direct synthesis of a defined GGTTAG hexameric telomere DNA repeat sequence (k_{pol}) (Fig 1A). A unique property of telomerase is the ability to translocate on the DNA product (k_{trans}) in order to recycle the integral TR template during processive addition of multiple telomere repeats prior to dissociation from the DNA product (k_{off}) (Fig 1A)[26]. This repeat addition processivity (RAP) implicitly requires multiple points of contact between telomerase and its DNA substrate, a notion that is consistent with data from a variety of telomerase systems identifying 'anchor site' DNA interactions outside the enzyme active site [27-30].

Model telomere DNA substrates harboring integer multiples of four consecutive telomere repeats are inefficient binding substrates for telomerase in vitro while DNA primers with five, six, or seven consecutive repeats are efficiently bound and extended [31]. Thus, while GQ structures can inhibit telomerase association, the presence of a small single-stranded DNA overhang in the substrate is sufficient to recover telomerase loading and function. These previous studies illuminated DNA sequence determinants that mediate the initial binding of telomerase to its substrate; however there remained an untested possibility that GQ structure may regulate the

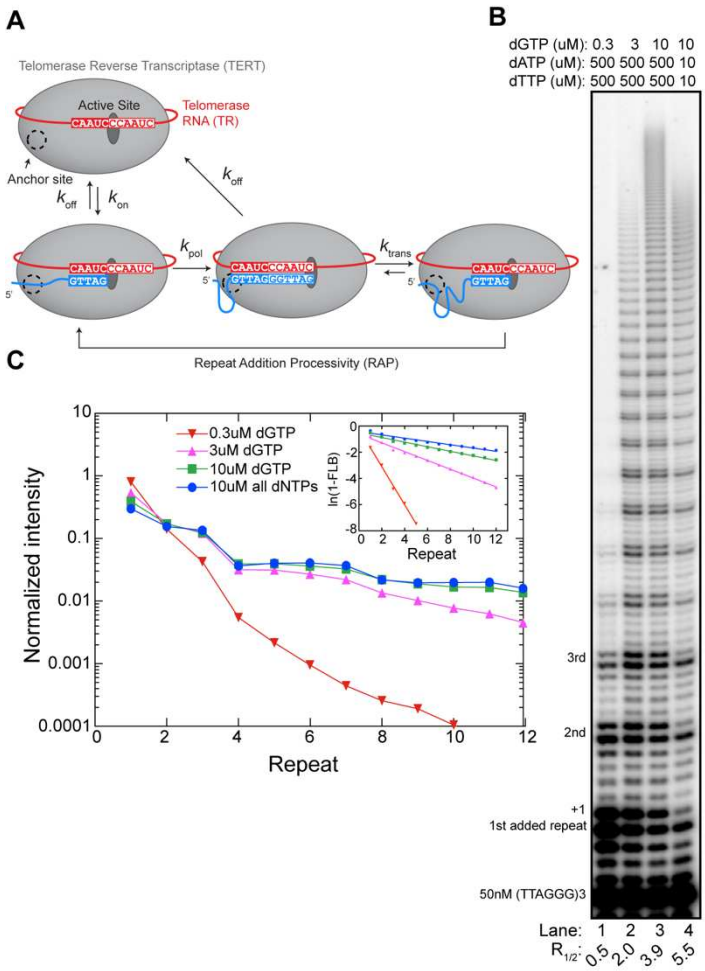


Figure 1. Human telomerase function. (A) Telomerase catalytic cycle. k_{off} and k_{on} represent the rates of dissociation and annealing to the telomere respectively. The rate of nucleotide addition during repeat synthesis is represented by k_{pol} and the translocation rate after the completion of each repeat is represented by k_{trans}. The rate constants governing nucleotide addition and translocation together make up the Repeat Addition Processivity (RAP). TERT is represented by a grey oval and TR is shown simplified in red. The telomerase anchor site is represented by a dashed circle. Telomere DNA is depicted in blue. (B) Telomerase primer extension assay with 50mM ^{32}P-labeled $(TTAGGG)_3$ primer. Nucleotide concentration is indicated above the gel and repeats added to the $(TTAGGG)_3$ primer are indicated on the left. The $R_{1/2}$ value is shown at the bottom of the gel. (C) Normalized intensity plotted vs repeat number. Raw intensity for each band was converted to a fraction by dividing the band intensity by the total counts for the lane and then plotted against repeat number. The plot of ln(1-FLB) vs repeat number used to calculate processivity is inset in the top right corner.

behavior of an actively extending telomerase-DNA complex, as suggested by studies on ciliate telomerase [32,33]. Furthermore, we reasoned that the specialized telomerase system provides a powerful opportunity to investigate the influence of GQ-forming sequences on nucleic acid polymerase function.

To study the relationship between DNA structure and human telomerase catalysis we performed direct primer extension assays using dNTP concentrations similar to those found in the cellular environment [34]. Our experiments reveal a complex pattern of telomerase DNA product accumulation that indicates the efficiency of template recycling is dependent upon the number of synthesized repeats. Experiments using telomere DNA primers of varying sequence and varying salt conditions support the notion that a GQ can form within the telomerase-DNA complex. Single-molecule FRET experiments provide further support for DNA structural dynamics within actively extending telomerase enzymes. To estimate individual rate constants for successive repeat addition cycles we performed global kinetic modeling of telomerase time-series data. Interestingly, our model converges to a unique solution of rate constants that provides a direct measure of processivities for each cycle of telomere repeat addition. These results are consistent telomere DNA GQ folding serving to promote template recycling, as well as to accelerate product dissociation. We present a mechanistic model that provides a framework to understand the delicate interplay of telomere DNA structure and function during telomerase catalysis.

RESULTS

Telomerase product distribution is sensitive to dNTP concentrations and stoichiometry

When measuring telomerase activity in vitro, it is common to employ direct primer extension assays in the presence of ^{32}P-dGTP. This approach permits reactions to be performed with a large excess of unlabeled DNA substrate, benefits from very high sensitivity of product detection, and circumvents PCR-induced artifacts inherent to the telomere repeat amplification protocol (TRAP) assay. However, the use of ^{32}P-dGTP incorporation to detect product accumulation limits the amount of total dGTP that can be used in the assay, leading to the widely reported practice of using non-physiological dNTP stoichiometry that has the potential to significantly alter the telomerase product distribution [35,36]. To circumvent this problem, we used 5' radiolabeled DNA primers and cold dNTPs to monitor telomerase activity (Fig. 1B). We compared this approach to standard assays performed with ^{32}P-dGTP using identical dNTP and primer concentrations (Fig. S1). Although the reaction profiles are qualitatively distinct, we observed quantitatively similar product distributions for the two approaches when normalized for the amount of ^{32}P-dGTP incorporation. Importantly, the majority of the input DNA primers are not extended in our experiments, demonstrating that our reaction conditions are sufficient to limit distributive telomerase activity (i.e. an individual primer being extended by multiple telomerase enzymes). Processive telomerase activity is also evident when analyzing

pulse-chase experiments in which longer DNA products continue to accumulate after addition of a 400-fold excess of cold DNA primer (Fig. S2).

Having established that our end-labeled DNA primer assay is capable of accurately monitoring processive telomerase action, we next sought to analyze the influence of varying dNTP concentrations on the telomerase product distribution (Fig. 1B). Previously published studies defined telomerase processivity as the number of repeats corresponding to the point where the dissociated DNA represents 50% of the total population ($R_{1/2}$) [37,38]. This approach is analogous to half-time ($t_{1/2}$) analysis and can be performed by fitting a linear regression to a plot of ln(1-FLB) versus repeat number, where FLB is the fraction left behind (Fig. 1C inset, see Methods for details). Titrating increasing amounts of dGTP in the presence of a large excess of dATP and dTTP yields a significant boost in RAP as has been reported previously for both human and *Tetrahymena* telomerase (Fig. 1B and 1C) [35,39-43]. However, the use of a large excess of dATP and dTTP is not a good approximation for the physiological dNTP pool which is generally closer to the ~10 uM range [34]. Note that dCTP is not required for the human telomere sequence and its absence does not alter telomerase function (data not shown). When assayed in the presence of equimolar dGTP, dATP, and dTTP, we observe the highest RAP activity of all conditions tested (Fig. 1B, lane 4). Moreover, using this approach, we noticed that the previously observed doublet band at the +1 position after each complete telomere repeat is effectively suppressed (Fig. 1B, compare lanes 1-3 with lane 4). We interpret this +1 band to reflect promiscuous use of the non-template RNA base at hTR position U45, which may be

enhanced by the large excess of dATP. Based upon these results, we elected to perform all subsequent telomerase assays in our study under conditions of equimolar dNTPs that better reflect cellular dNTP concentrations.

G-quadruplex folding alters the pattern of telomerase product accumulation

Established methods for approximating RAP using the $R_{1/2}$ value described above assume an exponential decay of accumulated telomerase products with each telomere repeat added [37,38,44]. However, we noted the appearance of plateaus in the product distribution when using equimolar concentrations of dNTPs (Fig. 1C and 2A). For example, when using a standard telomere DNA primer composed of the sequence $(TTAGGG)_3$, we observe a sudden drop in product accumulation between the bands corresponding to the third and fourth telomere repeats added to the primer (Fig. 2A, red asterisk). Surprisingly, the intensities of the subsequent four added repeats are approximately equal, until a second decrease in accumulation occurs between added repeats seven and eight. This pattern of four equally populated product lengths, followed by a decrease in accumulation, continues throughout the detectable range of telomere DNA products.

Telomere DNA primers with at least four contiguous G-rich repeats can fold into a GQ in vitro [45,46], suggesting the RAP-associated 'pattern of four' we observe in our experiments may be due to GQ folding of the DNA product within an actively extending telomerase complex. To test this hypothesis, we altered the 5' end of the telomere DNA sequence so that it no longer harbored the requisite run of guanines

needed to participate in GQ folding (Fig. 2A, lanes 2 and 3). Altering the primer in this way should change the product length where the 'pattern of four' appears if the newly synthesized DNA folds into a GQ. Indeed, a modified DNA primer with a 5' $(TG)_3$ substitution supports telomerase RAP, but the plateaus in the product profile are delayed by one additional repeat, corresponding to the sequence needed to promote GQ formation in the product DNA (Fig. 2A, compare lanes 1 and 2). Similarly, when the first two repeats in the telomere DNA primer were substituted with the TG dinucleotide sequence ($(TG)_6$), the plateaus are delayed by two additional repeats (Fig. 2A, compare lanes 1 and 3). These results were highly reproducible across three independent experimental trials (Fig. 2B) and agree with the hypothesis that GQ folding within the nascent telomere DNA causes the telomerase product profile to deviate from a strictly decreasing decay.

GQ stabilization alters telomerase kinetics and DNA product structural dynamics

The H-bonding configuration of the G-quartet motifs within a GQ fold can be differentially stabilized by coordination of specific monovalent cations, with a rank order of $K^+ > Na^+ > Li^+$ in terms of degree of stabilization [47]. Therefore, we next set out to further dissect the contribution of GQ folding to telomerase activity by altering the salt conditions of our telomerase assays. We observed robust telomerase activity in all salt conditions tested; however, there was a clear reduction in total product accumulation in Li^+ when compared to Na^+ and K^+ (Fig. 3A). The lower total

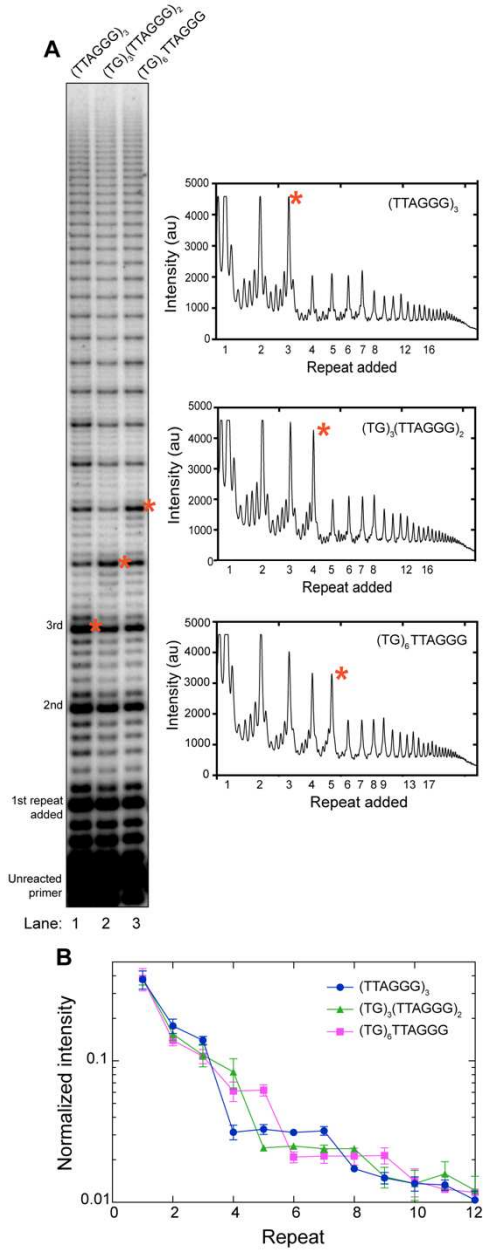


Figure 2. Product distribution profile is dependent on amount of consecutive TTAGGG repeats. (A) Telomerase primer extension assay with primers of varying TTAGGG composition. Primer variants are indicated to the right of the gel. Repeats added to the primer are indicated to the left. Lane profiles with raw intensity vs repeat band for each primer variant are shown on the right. Corresponding bands between the gel and lane profiles are indicated by a red asterisk. (B) Normalized intensity plotted vs repeat number. Fractional intensity of each band compared to total lane counts is plotted against repeat number. Data is averaged from three independent experiments.

product accumulation in the presence of Li^+ is a consequence of slower DNA synthesis kinetics (Fig. S3) and is not completely unexpected based on reported effects of Li^+ on DNA polymerases [48]. Interestingly, we do not observe the robust "pattern of four" RAP product distribution in the presence of Li^+ (Fig. 3A and 3B), the salt condition expected to least stabilize GQ folding during telomerase catalysis. This result lends additional support to the hypothesis that GQ formation in the presence of K^+ or Na^+ impacts the product distribution of an actively extending telomerase complex.

The results of our ensemble telomerase assays suggest that folding of the nascent DNA product can influence telomerase catalysis. However, this approach does not directly interrogate DNA conformation within an active RNP complex. Therefore, we turned to a single molecule Förster Resonance Energy Transfer (smFRET) based approach that directly monitors DNA structure and dynamics within individual telomerase enzymes [49,50]. In order to ensure our smFRET assay supports telomerase activity in both K^+ and Li^+ buffers we employed a recently reported method for in situ detection of extended DNA products at the single-molecule level [51]. Telomerase RNP complexes harboring a Cy3 dye incorporated into the telomerase RNA subunit were bound to a biotinylated DNA primer and then surface-immobilized onto a streptavidin coated glass slide (Fig. 4A). The telomerase-DNA complexes were incubated in either K^+ or Li^+ activity buffer as well as with a Cy5-labeled detection oligonucleotide with a sequence that is complementary to the

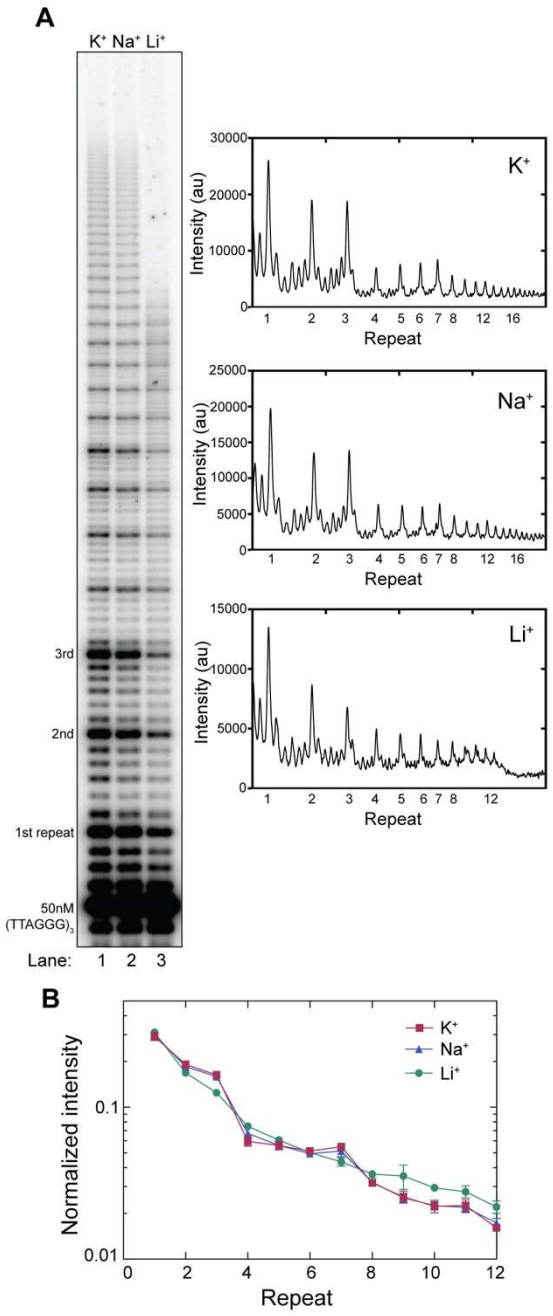


Figure 3. Product distribution profile is dependent on monovalent salt identity. (A) Telomerase primer extension assay in different salt conditions. Repeats added to the primer are indicated to the left. Lane profiles with raw intensity vs repeat band for each lane are shown on the right. (B) Normalized intensity plotted vs repeat number. Fractional intensity of each band compared to total lane counts is plotted against repeat number. Data is averaged from three independent experiments.

telomere product. In this way, telomere primers that are being actively extended by telomerase (i.e. enzyme is still bound and some of the DNA product becomes accessible) are detected as a FRET signal between the Cy3-labeled enzyme and the Cy5-labeled DNA probe. The appearance of the FRET signal was strictly dependent upon addition of activity buffer containing dNTPs and was time dependent (Fig. 4B). After 20 minutes of incubation, comparable levels of telomerase activity, measured as the total number of telomerase-DNA complexes producing a positive FRET signal, were detected in both K^+ and Li^+ activity buffers (Fig. 4B). Taken together, these results demonstrate that telomerase is catalytically active in both K^+ and Li^+ buffers in the single-molecule assay.

Having detected robust telomerase activity in both K^+ and Li^+ buffers, we next performed smFRET assays with Cy3-labeled telomerase and Cy5-labeled telomere primer in order to directly monitor DNA dynamics during active telomere elongation. (Fig. 4C). Stalled telomerase-DNA complexes harboring these site-specific dye modifications were surface-immobilized as described above. Data collected on stalled complexes prepared in this manner yielded a unimodal FRET distribution centered at ~0.75 (Fig. 4D, top left panel). Next, the telomerase complexes were activated for DNA synthesis by introducing dNTPs in telomerase activity buffer (containing 50mM KCl), resulting in a significant shift towards lower FRET values and ultimately yielding a single population centered at near-zero FRET (Fig. 4D, left panels). This FRET change upon activation of DNA synthesis is consistent with the Cy5-label site

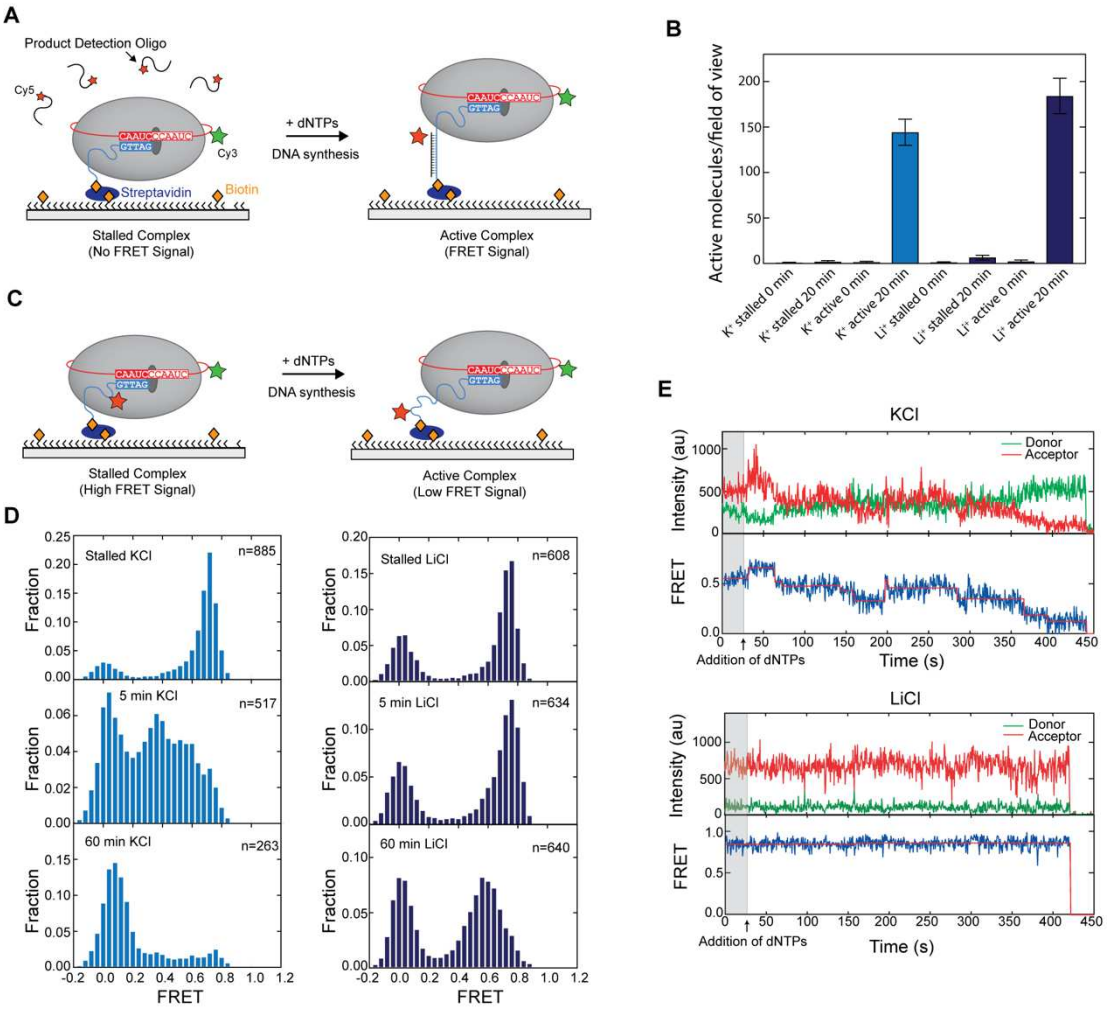

A
Product Detection Oligo
Cy5
Cy3
Streptavidin
Biotin
+ dNTPs
DNA synthesis
Stalled Complex
(No FRET Signal)
Active Complex
(FRET Signal)
B
Active molecules/field of view
K+ stalled 0 min
K+ stalled 20 min
K+ active 0 min
K+ active 20 min
Li+ stalled 0 min
Li+ stalled 20 min
Li+ active 0 min
Li+ active 20 min
C
+ dNTPs
DNA synthesis
Stalled Complex
(High FRET Signal)
Active Complex
(Low FRET Signal)
D
Stalled KCl
n=885
5 min KCl
n=517
60 min KCl
n=263
Stalled LiCl
n=608
5 min LiCl
n=634
60 min LiCl
n=640
Fraction
FRET
E
KCl
Donor
Acceptor
Intensity (au)
FRET
Addition of dNTPs
Time (s)
LiCl
Donor
Acceptor
Intensity (au)
FRET
Addition of dNTPs
Time (s)

Figure 4. smFRET studies of telomerase in the presence of K^+ and Li^+. (A) Schematic of human telomerase smFRET activity assay. Purified telomerase is immobilized to a pegylated and biotinylated quartz slide through binding to a biotinylated telomere primer (blue). TERT is depicted as a grey oval and hTR is shown as a red circle. A Cy3 dye (green star) is conjugated to hTR. A Cy5-labeled detection oligo complementary to two and a half telomere repeats is added to stalled enzyme (left). Upon addition of dNTPs, telomere synthesis begins, and sufficient lengths of telomere DNA for the Cy5-detection oligo to anneal to are extruded from the telomerase active site (right). (B) Telomerase is active in both K^+ and Li^+. smFRET experiments demonstrate a significant increase in active molecules per field of view in both KCl and LiCl buffers only in the presence of dNTPs. Experiments performed in KCl are shown in light blue and experiments performed in LiCl are shown in dark blue. The number of active molecules per field of view was calculated from the number of FRET-positive spots, as an indicator for the presence of Cy5-probe that annealed to actively synthesized telomere repeats. (C) Schematic of human telomerase smFRET experimental design to probe DNA conformational change. The telomere primer (blue) is conjugated to a Cy5 dye (red star). Upon addition of dNTPs, telomere repeat synthesis will move the Cy5 dye on the telomere primer further from the active site, resulting in a lower FRET value. (D) smFRET histograms in K^+ and Li^+. Histograms depict the FRET distribution for stalled telomerase (top panels) as well as after 5 and 60 minutes after addition of dNTPs (middle and bottom panels respectively). Experiments performed in KCl are shown in the left panels and experiments performed in LiCl are shown on the right. (E) Real time smFRET traces. Representative FRET traces are shown for experiments performed in KCl (top panel) and LiCl (bottom panel). Donor intensity is shown in red and acceptor intensity in green. The corresponding FRET value is depicted in blue.

on the telomere DNA moving further away from the telomerase active site, and perhaps folding into a stable structure, during multiple rounds of DNA repeat synthesis. Analysis of real-time single-molecule FRET trajectories collected on actively extending telomerase complexes in this same condition reveals complex DNA conformational dynamics characterized by a general transition to discrete lower

FRET states combined with the occurrence of transient increases in FRET (Fig. 4E and Fig. S4).

In order to dissect the contribution of DNA structure to the observed FRET dynamics we performed the same experiment using LiCl instead of KCl in the telomerase activity buffer. Once again, the initial FRET population of stalled telomerase-DNA complexes was centered at ~0.75 in the presence of Li^+ (Fig. 4D, top right panel). However, in contrast to the K^+ activity buffer, we did not observe a substantial drop in FRET upon addition of dNTPs (Fig. 4C right panels). Instead we observed a modest shift in the major FRET population to ~0.6 after 60 minutes incubation time. This result is consistent with individual smFRET trajectories of telomerase-DNA complexes incubated in Li^+ activity buffer, where the FRET level appears to remain constant over the course of several minutes (Fig. 4D). Together, these results argue that the formation of a GQ in the nascent telomere product contributes to efficient extrusion of the telomere DNA away from the telomerase active site during the course of activity.

Global kinetic modeling provides a measure of telomerase microscopic processivity

Telomerase processivity can be modeled as a series of consecutive reactions in which nucleotide addition is in competition with DNA dissociation at each step of the reaction. To simplify our telomerase kinetics analysis, we focus on the intense repeat addition bands, assuming the intervening nucleotide addition steps are rapid

and accompanied by little DNA dissociation. The macroscopic processivity of telomerase can be conveniently described by the median product length ($R_{1/2}$= ln2/decay constant) (see Fig. 1C) [37]. However, treating the data in this manner has several limitations. First, in a typical primer extension experiment one cannot distinguish between dissociated product and DNA that remained associated with enzyme at the point when the reaction was arrested. In addition, this analysis approach has the effect of suppressing possible deviations in microscopic processivity, which reflects the individual probability of adding another DNA repeat at a specific step of telomerase catalysis. The experiments described in the present study provide clear evidence that telomerase products do not accumulate uniformly and display patterns dependent upon assay conditions, DNA sequence, and/or product length. We therefore developed a kinetic model that can be utilized to globally fit telomerase time-series data in order to extract microscopic processivity values that underlie the observed distributions of telomerase products (Fig. 5A). Using this scheme, we treat this multistep process as a first-order reaction with an effective forward rate constant (k_f) for the transition between each repeat and a dissociation rate constant (k_d) (Fig. 5A) (note that k_f in this model reflects a combination of k_{pol} and k_{trans} described in Fig.1a). Using the kinetic scheme depicted in Fig. 6a, we can then define the microscopic processivity (p) at each step as $p = k_f / (k_f + k_d)$.

DNA dissociation is an effectively irreversible process when a large excess of unreacted primer remains, which outcompetes the re-binding of any product DNA. To ensure our experiments complied with this assumption, we analyzed telomerase

kinetics following a chase with 400-fold excess of unlabeled primer DNA, which serves to block re-association of the labeled DNA primer following telomerase dissociation (Fig. 5B). Telomerase time course assays were performed in the presence of either K^+ or Li^+ activity buffer conditions (Fig. 5B). Activity was initiated at time zero in the presence of end-labeled telomere DNA primer and dNTPs, followed by addition of excess chase primer at 20 minutes. The presence of 400-fold excess cold primer prior to enzyme addition was sufficient to eliminate any observable extension of the 50 nM end-labeled DNA primer used in our assays (Fig. 5B, lanes 1 and 9). Time points were collected at regular intervals out to 90 minutes and the concentration of each repeat species ($B + B^{\#}$, $C + C^{\#}$, etc.) was determined at each time point from the band intensity, knowing that the intensity of the initial primer was 50 nM. Individual rate constants were estimated by fitting the concentration time courses globally, using DynaFit (Fig. 5C, data shown for bands E-J). Global fitting converged to a unique set of rate constants that could be extracted with reasonable precision (see Supplementary Methods for details of kinetic modeling) (Fig. 5D). Comparison of the data obtained in the presence of K^+ and Li^+ revealed that the rate constants, k_f and k_d, decrease with increasing repeat number and result in lower microscopic processivity values for short DNA products as has been noted previously [37]. However, in the presence of K^+, the rate constants and microscopic processivity values show a saw-tooth modulation that gives rise to the 'pattern of four' clustering of products noted earlier (Fig. 5E, closed symbols). Although the effect is relatively

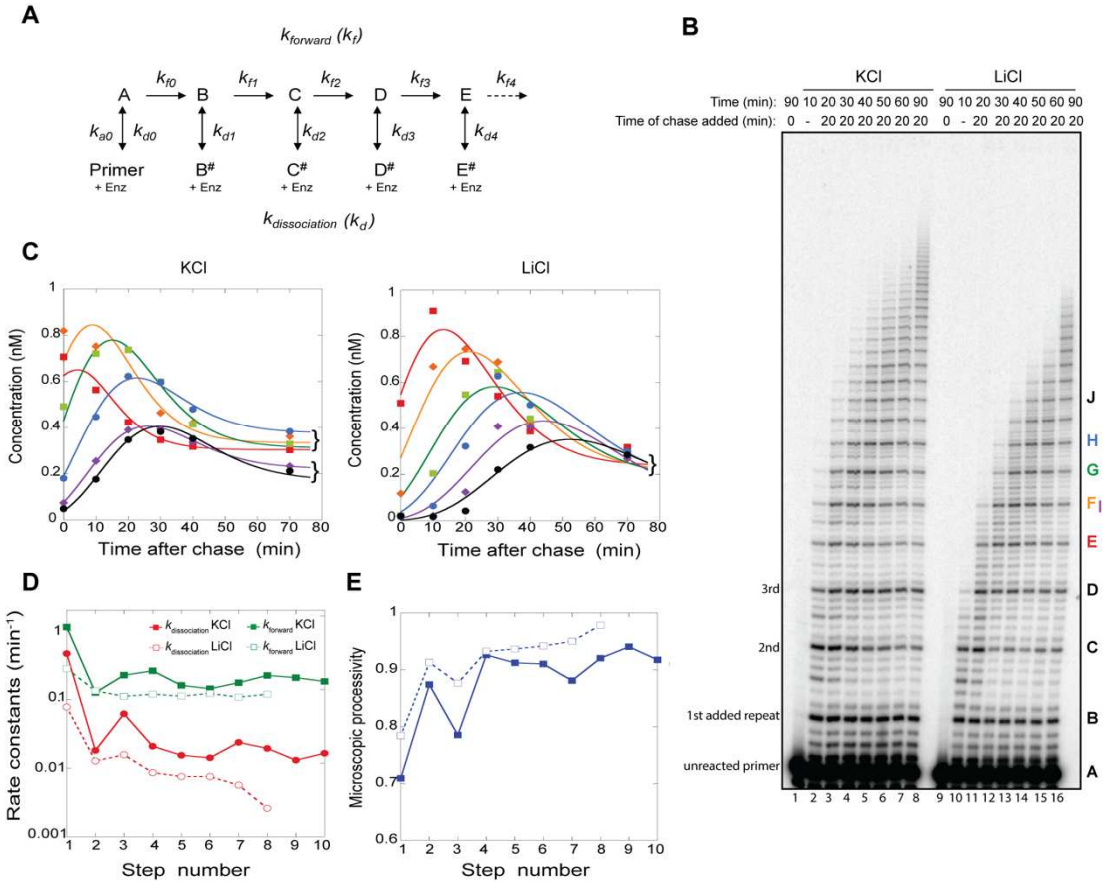

A
$k_{forward}$ (k_f)
A
B
C
D
E
k_{f0}
k_{f1}
k_{f2}
k_{f3}
k_{f4}
k_{a0}
k_{d0}
k_{d1}
k_{d2}
k_{d3}
k_{d4}
Primer + Enz
B# + Enz
C# + Enz
D# + Enz
E# + Enz
$k_{dissociation}$ (k_d)
B
KCl
LiCl
Time (min): 90 10 20 30 40 50 60 90 90 10 20 30 40 50 60 90
Time of chase added (min): 0 - 20 20 20 20 20 20 0 - 20 20 20 20 20 20
3rd
2nd
1st added repeat
unreacted primer
J
H
G
F
E
D
C
B
A
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16
C
KCl
LiCl
Concentration (nM)
Time after chase (min)
D
Rate constants (min⁻¹)
$k_{dissociation}$ KCl
$k_{forward}$ KCl
$k_{dissociation}$ LiCl
$k_{forward}$ LiCl
Step number
E
Microscopic processivity
Step number

Figure 5. Human telomerase kinetics. (A) Kinetic mechanism for processive telomerase activity used to globally fit the primer extension assay shown in (B). The letters refer to the repeat band number (B = 1st added repeat, C = 2nd repeat etc.) and dissociated products are identified with the # symbol. Band intensities are proportional to the sum of products (e.g. B + $B^{\#}$). (B) Extending primer dissociation rate assay in the presence of KCl and LiCl. Primer extension assays were performed with 50 nM ^{32}P-labeled $(TTAGGG)_3$ primer. 20 uM cold $(TTAGGG)_3$ primer was added to the reaction after 20 minutes of activity. A control reaction with 20uM cold primer added at the beginning of telomerase activity was included for both buffer conditions (lanes 1 and 9). Repeat number added to the primer is indicated on the left of the gel. Letters indicating band identity for kinetics modeling is indicated on the right of the gel. Bands E-J are colored according to plot shown in (c). (C) Representative global fits to bands E-J in KCl (left panel) and LiCl (right panel). The concentration of the products (see color code in (a)), based on band intensity relative to the initial 50 nM primer, was plotted against the time after the cold chase. Note the clustering of bands E-H and I-J 70 minutes post chase in the presence of KCl, which corresponds to the four repeats of the first plateau and the first two repeats of the second plateau (cf. Fig. 2). This partitioning is not present in the presence of LiCl. (D) Consecutive rate constant values for forward repeat addition (green squares, k_f) and dissociation (red circles, k_d) returned by DynaFit for data in the presence of KCl (solid symbols) and LiCl (open symbols). The step number refers to the rate constant subscripts shown in (a). Note the overall reaction is slower in the presence of LiCl and beyond band I (8th step) the fitted rate constant values had a large error because the decay phase had barely started by 70 min. Therefore, these values were omitted. (E) Microscopic processivity ($k_f/(k_f + k_d)$) at each step of the reaction calculated from the rate constants shown in (d). Note the saw-tooth structure in the presence of KCl (solid line) compared to the relative lack of structure in LiCl beyond the second step (dashed line).

small it is robust as determined by Monte-Carlo analysis and is reproducible between experiments and telomerase preparations (Fig. S5 and S6). Interestingly, the rate constants for repeat addition (k_f) and dissociation (k_d) were greater in the presence of K^+ than in Li^+, but the resultant microscopic processivity was lower (Fig. 5D and 5E). Taken together, these results suggest that in the presence of K^+, the ability of the DNA product to form a GQ fold (which first arises at step 3, see Discussion for

details) appears to promote translocation (increased k_f) but at the increased risk of DNA product dissociation (increased k_d).

DISCUSSION

The foundation of telomere structure consists of short G-rich repeat sequences, GGTTAG in humans, that have the propensity to fold into G-quadruplex (GQ) structures in vitro and in vivo [6,7]. Short ssDNA oligonucleotides that form GQs are poor telomerase substrates in vitro, due to the occlusion of an accessible single-stranded 3' end [31,46]. Here, we present evidence that GQ formation can occur within an actively extending telomerase complex in vitro and that formation of such GQs affect the kinetic properties of telomerase. Telomerase activity assays using modified telomere DNA primer sequence and varying salt conditions together support the notion that GQ folding occurs during telomere DNA synthesis. Consistent with this concept, single-molecule FRET experiments directly demonstrate complex folding dynamics of the nascent DNA product in the presence of K^+ but not Li^+. We describe a detailed kinetic framework for telomerase catalysis and use this model to globally fit telomerase time-series data in order to extract microscopic processivity values for each cycle of telomere DNA repeat synthesis (see below for details). Our kinetic modeling reveals small but significant GQ-dependent changes in the rate constants describing the DNA repeat synthesis reaction (k_f) and product dissociation (k_d).

Telomerase repeat addition processivity (RAP) can be quantitatively described as the number of repeats corresponding to the point where the dissociated DNA

represents 50% of the total population ($R_{1/2}$) [37]. The median number of repeats is a measure of the macroscopic processivity of telomerase that is implicitly assumed to be independent of product length. This method of analysis has been useful in previous studies that sought to characterize telomerase RAP under varying experimental conditions or with mutant telomerase enzymes [37,52]. A caveat of this approach is that typical telomerase primer extension assays cannot distinguish between products that have already dissociated and DNA that remained bound to enzyme when the reaction was stopped. This caveat can be addressed by the kinetic modeling approach described in the present study, or by introducing additional experimental steps to separate dissociated and bound products prior to gel analysis [43]. However, one additional advantage of the global kinetic modeling approach is the ability to extract microscopic processivity values that reflect the individual probability of synthesizing each subsequent telomere repeat. Indeed, the results of our global kinetic modeling reveal a saw-tooth modulation of k_f and k_d, that together explain the 'pattern of four' clustering of telomerase product accumulation.

We present a working model for the mechanism of GQ-dependent effects on telomerase repeat addition processivity (Fig. 6). The complex rearrangements that are necessary for template recycling during multiple rounds of telomere repeat synthesis

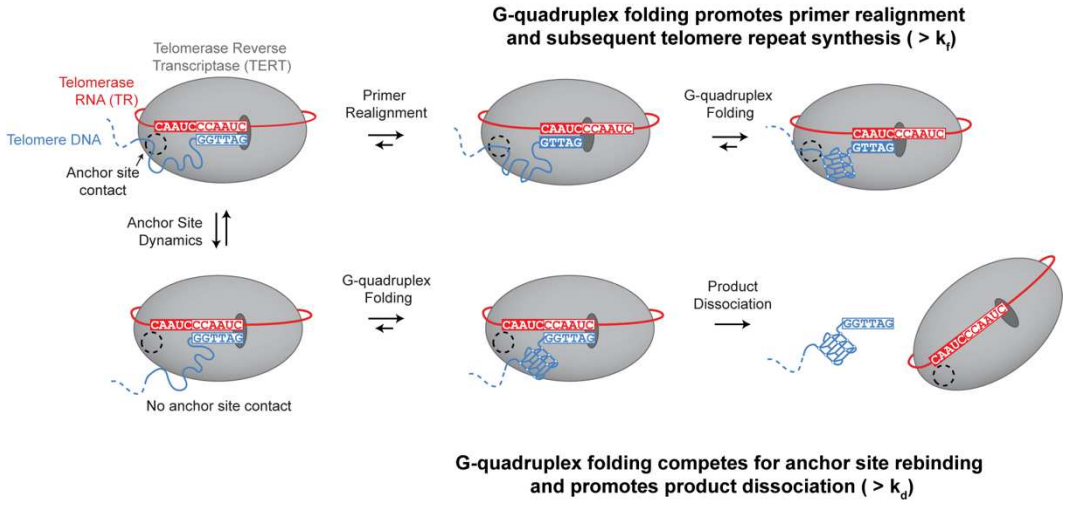


Figure 6. Model of GQ folding within actively extending telomerase complex. At the completion of a telomere repeat, telomerase (with TERT depicted in grey and hTR in red), is annealed to the telomere primer (blue). The anchor site (dashed circle) is either engaged with the telomere DNA (upper pathway) or disengaged (lower pathway). If the anchor site contacts are maintained and primer realignment occurs (upper, middle carton), there should be sufficient DNA to allow for a GQ to form within the actively extending enzyme (top, right cartoon). The formation of a GQ at this stage may bias the enzyme complex towards another round of telomere repeat addition. If the anchor site contacts are broken (lower, left cartoon) while the 3' end of the telomere is bound in the active site, there may be competition between GQ formation and anchor site interactions (bottom left and middle cartoons). If primer realignment occurs at this stage, product dissociation will occur (bottom, right cartoon).

require multiple points of contact between telomerase and its DNA substrate. Minimally, the DNA must engage the enzyme by interacting with the RNA template in the active site and/or making distal contacts with a separate 'anchor site' [27-30]. Upon completion of a telomere repeat, the 3' end of the DNA product must dissociate from the template and dynamically realign with the downstream region of the RNA to prime the next round of repeat synthesis (Fig. 6) [50]. This large-scale translocation step

represents a vulnerable stage of telomerase catalysis, requiring a stable anchor-site interaction to prevent product dissociation. In principle, GQ folding within the DNA product can compete with the anchor-site contact to promote dissociation (Fig. 6, top pathway). Alternately, GQ folding can bias the positioning of the 3' end of the primer to favor realignment for a subsequent round of repeat synthesis while allowing for anchor site contacts to remain intact (Fig. 6, bottom pathway). Which pathway the enzyme takes will depend on if anchor site interactions are maintained and the register of the 3' end of the telomere during the formation of a GQ. In the schematic model depicted in Fig. 6, a GQ formed when the most recently synthesized telomere is bound in the active site will disrupt anchor site interactions and promote product dissociation, while a GQ formed when the 3' telomere end is annealed to the template priming site, will facilitate another round of repeat addition. The latter outcome is mechanistically similar to DNA hairpin induced translocation models proposed for diverse telomerase systems [53,54].

The finding of GQ-dependent effects on telomerase catalysis in vitro begs the question of whether a similar phenomenon occurs in the cell. While direct demonstration of GQ folding in vivo is challenging, recent experimental advances using GQ specific antibodies suggest that GQ folding can occur in a temporally and spatially regulated manner in vivo [12-14]. Additionally, studies of ciliate telomerase have suggested that GQ folding may contribute to product dissociation during active telomere elongation [32,33]. The recently reported cryo-EM structures of the human and *Tetrahymena* telomerase enzymes provide a platform for investigating the potential

for GQ folding within the actively extending telomerase complex [55,56]. Interestingly, when analyzing the EM density of the human telomerase complex, we noticed a structural pocket immediately proximal to the path of the nascent DNA emerging from the active site (Fig. S7). This pocket is flanked by two evolutionarily conserved features that are essential for telomerase processivity: the TEN domain and the TR p6.1 stem loop. The volume of this pocket is remarkably compatible with the dimensions of a single GQ fold, and raises the possibility that the nascent DNA products has ample space to assume a GQ structure within the confines of telomerase RNP complex. In the *Tetrahymena* structure, the C-terminal domain of the processivity factor, Teb1, is positioned as a lid to the DNA exit pocket, forming an enclosed cavity that is also sufficient to accommodate a single GQ fold. We propose the hypothesis that the Teb1-related shelterin component POT1 may occupy a similar position in the human telomerase complex. Thus, the anticipated binding of POT1 to the DNA as it is threaded out of this DNA exit channel is not mutually exclusive with the possibility of GQ folding within this protected cavity [46,56]. While speculative, the notion of GQ folding within the DNA exit channel is consistent with the amount of DNA protected by exonuclease cleavage as well as previous smFRET measurements that showed a similar amount of the DNA substrate is in close proximity to the template [50,57]. (refs). Future structural and biochemical experiments are required to investigate the influence of the shelterin proteins POT1-TPP1 on the folding properties of the nascent telomere DNA product, as well as the in vivo significance of GQ folding during telomerase catalysis.

MATERIALS AND METHODS

Preparation of RNAs

Synthetic, dye-labeled RNA fragments

Synthetic PK hTR fragment 32-62 was ordered rom Dharmacon with an internal aminoallyl uridine (5-N-U) on position U42. A single tube of RNA was resuspended in nuclease-free H2O and ethanol precipitated in the presence of 300mM NaOAc pH 5.2 to prevent introduction of reactive amines. The pellet was resuspended in 100 ul 0.1M NaHCO3 and was used to solubilize a single mono-reactive Cy3-dye pack (Amersham). The solution was incubated for 2 h at 37°C. The solution was then ethanol precipitated, again with NaOAc. The RNA was then deprotected by resuspension in deprotection buffer (100mM acetic acid, pH 3.6) followed by vortexing and centrifugation for 10 seconds each. The solution was then heated at 60°C for 30 min followed by ethanol precipitation. The pellet was resuspended in 60 ul 0.1M TEAA, pH 7.5, and was HPLC purified on a reverse phase C18 column (Agilent technologies).

Ligation of RNA fragments

A splinted ligation reaction containing 800 pmol of Cy3-labeled hTR 32-62 fragment, 1600 pmol of in vitro transcribed unlabeled hTR 63-195, 1600 pmol DNA splint and 0.5x ligase buffer (NEB) was set up in a 200ul volume and incubated for 95C for 5 min followed by 30C for 10 min. 200ul ligation mix (1.5x ligase buffer, 8000 units of T4 DNA ligase (NEB), 2 mM ATP and 1U/ul RNasin Plus) was then

added to the reaction mixture and incubated overnight at 30°C. 10 units of Turbo DNase was added and incubated for 15min at 37°C. The RNA was then phenol-chloroform extracted and ethanol precipitated prior to PAGE purification.

In vitro transcription

Unlabeled CR4/5 (hTR 239-328) and PK (32-195) was in vitro transcribed using homemade T7 polymerase in 1x RNA polymerase buffer (40mM Tris-HCl pH 7.9, 6mM $MgCl_2$, 1mM DTT, 2mM spermidine) with 1.5 mM NTPs, 22 mM additional $MgCl_2$, 90mM additional DTT and 40 units RNasin Plus. The reaction was incubated overnight at 37°C followed by the addition of 10 units of Turbo DNase for 15min at 37C. The RNA was phenol-chloroform extracted and ethanol precipitated prior to PAGE purification. PK 63-195 was in vitro transcribed as above but with 1mM of each NTP and an additional 5mM GMP in order to preferentially make a 5' monophosphate fragment for use in splinted ligation.

Telomerase expression

Human telomerase was reconstituted using the Promega TnT Quick Coupled Transcription/Translation system. In Lo-bind tubes (Eppendorf), 200ul of TnT quick mix was combined with 5ug of pNFLAG hTERT plasmid as well as 1uM of the in vitro transcribed PK and CR4/5 fragments except for when using Cy3-labeled PK, which was added at 0.1uM. The reaction was incubated for 3 h at 30°C. 5 ul of 0.5 M EDTA, pH 8.0, was added to quench excess Mg^{2+} present in the lysate.

Telomerase purification

Directly after reconstitution, telomerase was purified using the N-terminal FLAG tag on hTERT. To pull down the enzyme, Sigma Anti-FLAG M2-agarose resin was used. 50 ul bead slurry was first washed three times with wash buffer (50 mM Tris-HCl, pH 8.3, 3 mM $MgCl_2$, 2 mM DTT, 100 mM NaCl) with 30 sec spins at 5000rpm at 4°C after each wash. The beads were then blocked in blocking buffer (50mM Tris-HCl pH 8.3, 3mM $MgCl_2$, 2mM DTT, 500ug/ml BSA, 50ug/ml glycogen, 100ug/ml yeast tRNA) for 15 min under gentle agitation at 4°C followed by a 30 sec spin and removal of the supernatant. This blocking step was performed twice. After blocking, the beads were resuspended in 200ul fresh blocking buffer and added to the telomerase lysate. The beads and lysate were incubated for 2 h at 4°C under gentle agitation. The beads were then spun down for 30 sec at 5000 rpm and at 4°C and the supernatant was discarded. The beads were then washed three times in wash buffer containing 300mM NaCl (KCl or LiCl for salt dependence experiments) followed by three washes in wash buffer with 100 mM NaCl (KCl or LiCl for salt dependence experiments). A 30 sec spin at 5000 rpm at 4°C was performed between each wash. To elute the enzyme, the beads were incubated in 60 ul elution buffer (50 mM Tris-HCl, pH 8.3, 3 mM $MgCl_2$, 2 mM DTT, 750 ug/ml 3x FLAG peptide, 20% glycerol) under gentle agitation at 4°C for 1 h. After elution, the beads were removed by centrifugation at 13000 rpm through Nanosep MF 0.45um filters. 5ul aliquots were then aliquoted into Lo-bind tubes and flash frozen in liquid nitrogen and stored at -70°C until use.

End-labeling of primers

50pmol of primer was labeled with alpha-^{32}P ATP using T4 PNK in 1x PNK buffer (70 mM Tris-HCl, pH 7.6, 10 mM $MgCl_2$, 5 mM DTT) in 50 ul. The reaction was incubated for 1h at 37°C followed by heat inactivation of T4 PNK at 65°C for 20min. Centrispin columns (Princeton separations) were used to obtain labeled primer.

Primer extension assays

Telomerase activity assays were performed by using 5 ul of purified telomerase in 1x activity buffer (50 mM Tris-HCl pH 8.3, 50 mM KCl (NaCl or LiCl when indicated), 1 mM $MgCl_2$, 2 mM DTT) in a final volume of 15ul. For experiments containing end-labeled primer, each reaction contained 10 uM each of dATP, dTTP and dGTP as well as 50 nM of indicated ^{32}P-primer. For reactions containing radiolabeled dGTP, each reaction contained the indicated dNTP concentrations as well as 50nM unlabeled $(TTAGGG)_3$ primer. Reactions were incubated for 90 minutes at 30C. Reactions were stopped using 200ul 1x TES buffer (10 mM Tris-HCl, pH 7.5, 1 mM EDTA, 0.1% SDS), followed by phenol-chloroform extraction and ethanol precipitation of the DNA. The pellet was resuspended in 1x formamide gel loading buffer and resolved on a 12% denaturing PAGE gel. The gel was then dried and exposed on a phoshorimager screen and scanned using a Typhoon scanner. Band intensities were quantified using SAFA [58]. R1/2 values were calculated by summing each band and all bands below it divided by the total counts for a given

lane (to calculate the fraction left behind (flb). The natural log of (1- flb) was then plotted against repeat number and 0.693 was divided by the value of the slope to give the R1/2 value [37].

Single-molecule experiments

Slide preparation

Quartz slides (Finkenbeiner Inc.) were boiled in water for 20 min to remove parafilm, epoxy and coverslips from previous experiments. The quartz slides were then scrubbed with alconox by hand and rinsed with MilliQ water. The slides were then placed in a solution containing 10% w/v Alconox and sonicated for 20 min. The slides were then rinsed and sonicated in MilliQ water for 5 min. The slides were then sonicated in acetone for 15 min followed by direct transfer to a solution containing 1 M KOH and sonicated for 20 min. The slides were then rinsed with water and thoroughly flame dried using a butane torch (BernzOmatic). After flaming, the slides were left to cool on a drying rack in the fume hood. While the slides were cooling, a slide holder containing methanol was sonicated for 5 min and a silanizing solution containing 100 ml methanol, 5 ml of glacial acetic acid and 1 ml of N-(2-aminoethyl)-3-aminopropyltrimethoxysilane (UCT) was prepared during this time. After the slides had cooled down, the methanol in the slide holder was exchanged for the silanizing solution and the slides were sonicated in this solution for 1 min and then allowed to stand in the solution for an additional 20 min at room temperature. While slides were incubating, 400 mg of mPEG-Succinimidyl Valerate MW 5000

(Laysan Bio, Inc.) was resuspended in 800 ul 0.1M $NaHCO_3$. Also, 2 mg of Biotin-PEG-Succinimidyl Valerate MW 5000 (Laysan Bio, Inc.) was resuspended in 200 ul of 0.1M $NaHCO_3$ and mixed with the other PEG solution. After combining, the PEG solution was briefly sonicated to ensure complete resuspension. The slides were rinsed with MilliQ water and dried with nitrogen gas and then put in a humidor box. 150 ul of the PEG solution was applied to each slide and covered with a coverslip. The slides were incubated in the dark overnight. The next day the coverslip and PEG solution was rinsed off using MilliQ water and dried using nitrogen gas. Channels were assembled using parafilm strips as spacers on the pegylated quartz slide and plasma-cleaned coverslips were added as the upper channel face. For real-time experiments, flow cells were assembled by first drilling two holes at opposite ends of a quartz slide. A cut pipette tip was glued into one hole to act as a reservoir, and intramedic® PE100 polyethylene tubing glued to the other to provide a vacuum for buffer exchange. The channel was cut from a piece of double sided tape which was sandwiched between the pegylated quartz slide and a coverslip.

Enzyme immobilization

Channels were incubated with 30 ul 10 mg/ml BSA (NEB) for at least 30 min. 60 ul of 0.2 mg/ml MPS (streptavidin) was then added and incubated for 5 min. The channel was then washed with 150 ul T50 buffer (10 mM Tris-HCl, pH 8, 50 mM NaCl, or LiCl for Li^+ experiments)). 2.5 ul of Cy3-telomerase, 0.5 nM primer-Cy5 and Ix imaging buffer (50 mM Tris-HCl, pH 8.3, 50 mM KCl, or LiCl for Li^+

experiments, 1 mM $MgCl_2$, 0.8 % glucose, 0.5 mg/ml BSA) in a total volume of 10 ul was incubated for 1h on the benchtop in the dark. Before addition of primer to the enzyme, a 10x primer stock was diluted in 1x T50 with LiCl and then heated for 5 min at 95°C. The 10x primer was snap cooled on ice and 1 ul was subsequently added to the enzyme. The imaging buffer was first saturated with Trolox (triplet state quencher) and then filtered through a 0.22 um filter and brought to pH 8.3. The enzyme/primer complex was then diluted with 40 ul of Ix imaging buffer and then incubated on the slide for 10 minutes. The slide was put on the TIRF objective, the density of immobilized molecules was assessed and 100 ul of imaging buffer + 1 ul gloxy solution (200 ug/ml catalase and 100 mg/ml glucose oxidase in T50) was flowed through the channel.

Primer extension by surface-immobilized telomerase

After the telomerase complexes were stalled on the slide surface, the indicated dNTPs were added as described in the text at a concentration of 200 uM. Reactions were allowed to proceed for the indicated amount of time. For real-time experiments, dNTPs were added to the slide during active data acquisition.

Data acquisition and analysis

Imaging fields containing 300-700 molecules were imaged at 2 frames/second for real-time smFRET experiments and 10 frames/second for standard smFRET experiments. Green laser intensity was set to 3mW for real-time experiments and to

15mW for standard experiments. 20 two-second movies were collected at each indicated time-point for the standard FRET experiments and 15 minute movies were collected for the real-time experiments. Individual traces were parsed out using custom written IDL software to correct for dye-cross talk and background. Individual traces were then filtered in MATLAB and molecules that did not contain and acceptor dye were discarded. FRET intensities were calculated using the equation $I_A/(I_A+yI_D)$ where I_A is the acceptor intensity, I_D is the donor intensity and y is the gamma correction factor. The first two seconds of individual FRET traces were binned into histograms.

Kinetic analysis

For kinetic analysis, the primer extension assay described above was modified by chasing with 20 μM cold $(TTAGGG)_3$ primer after 20 minutes of initiating the reaction, to prevent hot primer and hot product rebinding within the subsequent 70-minute time course. The intensities of the RAP bands on the gel electrophoretogram were converted to absolute concentrations based on the intensity of the initial hot primer band (50 nM). Multiple exposures were analyzed to ensure intense and weak bands remained within the linear range of the phosphorimager. These data were analyzed globally according to the sequential model (Figure 6B) using DynaFit [59,60]. Further details are given in the supplementary Methods.

REFERENCES

1. de Lange, T. Shelterin: the protein complex that shapes and safeguards human telomeres. *Genes Dev* **19**, 2100-10 (2005).
2. de Lange, T. How shelterin solves the telomere end-protection problem. *Cold Spring Harb Symp Quant Biol* **75**, 167-77 (2010).
3. Palm, W. & de Lange, T. How shelterin protects mammalian telomeres. *Annu Rev Genet* **42**, 301-34 (2008).
4. McElligott, R. & Wellinger, R.J. The terminal DNA structure of mammalian chromosomes. *EMBO J* **16**, 3705-14 (1997).
5. Maizels, N. & Gray, L.T. The G4 genome. *PLoS Genet* **9**, e1003468 (2013).
6. Li, J., Correia, J.J., Wang, L., Trent, J.O. & Chaires, J.B. Not so crystal clear: the structure of the human telomere G-quadruplex in solution differs from that present in a crystal. *Nucleic Acids Res* **33**, 4649-59 (2005).
7. Burge, S., Parkinson, G.N., Hazel, P., Todd, A.K. & Neidle, S. Quadruplex DNA: sequence, topology and structure. *Nucleic Acids Res* **34**, 5402-15 (2006).
8. Saharia, A. et al. Flap endonuclease 1 contributes to telomere stability. *Curr Biol* **18**, 496-500 (2008).
9. Leon-Ortiz, A.M., Svendsen, J. & Boulton, S.J. Metabolism of DNA secondary structures at the eukaryotic replication fork. *DNA Repair (Amst)* **19**, 152-62 (2014).

10. London, T.B. et al. FANCJ is a structure-specific DNA helicase associated with the maintenance of genomic G/C tracts. *J Biol Chem* **283**, 36132-9 (2008).

11. Balasubramanian, S., Hurley, L.H. & Neidle, S. Targeting G-quadruplexes in gene promoters: a novel anticancer strategy? *Nat Rev Drug Discov* **10**, 261-75 (2011).

12. Schaffitzel, C. et al. In vitro generated antibodies specific for telomeric guanine-quadruplex DNA react with Stylonychia lemnae macronuclei. *Proc Natl Acad Sci U S A* **98**, 8572-7 (2001).

13. Lam, E.Y.N., Beraldi, D., Tannahill, D. & Balasubramanian, S. G-quadruplex structures are stable and detectable in human genomic DNA. *Nature Communications* **4**(2013).

14. Rodriguez, R. et al. Small-molecule-induced DNA damage identifies alternative DNA structures in human genes. *Nature Chemical Biology* **8**, 301-310 (2012).

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

16. 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).

17. Allsopp, R.C. et al. Telomere length predicts replicative capacity of human fibroblasts. *Proc Natl Acad Sci U S A* **89**, 10114-8 (1992).

18. Alder, J.K. et al. Ancestral mutation in telomerase causes defects in repeat addition processivity and manifests as familial pulmonary fibrosis. *PLoS Genet* **7**, e1001352 (2011).

19. Parry, E.M., Alder, J.K., Qi, X., Chen, J.J. & Armanios, M. Syndrome complex of bone marrow failure and pulmonary fibrosis predicts germline defects in telomerase. *Blood* **117**, 5607-11 (2011).

20. Gramatges, M.M., Qi, X., Sasa, G.S., Chen, J.J. & Bertuch, A.A. A homozygous telomerase T-motif variant resulting in markedly reduced repeat addition processivity in siblings with Hoyeraal Hreidarsson syndrome. *Blood* **121**, 3586-93 (2013).

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

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

23. Feng, J. et al. The RNA component of human telomerase. *Science* **269**, 1236-41 (1995).

24. Nakamura, T.M. et al. Telomerase catalytic subunit homologs from fission yeast and human. *Science* **277**, 955-9 (1997).

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

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

27. 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).

28. 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).

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

30. Robart, A.R. & Collins, K. Human telomerase domain interactions capture DNA for TEN domain-dependent processive elongation. *Mol Cell* **42**, 308-18 (2011).

31. Hwang, H. et al. Telomeric overhang length determines structural dynamics and accessibility to telomerase and ALT-associated proteins. *Structure* **22**, 842-53 (2014).

32. Zahler, A.M., Williamson, J.R., Cech, T.R. & Prescott, D.M. Inhibition of telomerase by G-quartet DNA structures. *Nature* **350**, 718-20 (1991).

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

34. Traut, T.W. Physiological concentrations of purines and pyrimidines. *Mol Cell Biochem* **140**, 1-22 (1994).

35. Sun, D., Lopez-Guajardo, C.C., Quada, J., Hurley, L.H. & Von Hoff, D.D. Regulation of catalytic activity and processivity of human telomerase. *Biochemistry* **38**, 4037-44 (1999).

36. Chen, J.L. & Greider, C.W. Determinants in mammalian telomerase RNA that mediate enzyme processivity and cross-species incompatibility. *EMBO J* **22**, 304-14 (2003).

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

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

39. Hammond, P.W. & Cech, T.R. dGTP-dependent processivity and possible template switching of euplotes telomerase. *Nucleic Acids Res* **25**, 3698-704 (1997).

40. Maine, I.P., Chen, S.F. & Windle, B. Effect of dGTP concentration on human and CHO telomerase. *Biochemistry* **38**, 15325-32 (1999).

41. Hardy, C.D., Schultz, C.S. & Collins, K. Requirements for the dGTP-dependent repeat addition processivity of recombinant Tetrahymena telomerase. *J Biol Chem* **276**, 4863-71 (2001).

42. Wu, R.A., Tam, J. & Collins, K. DNA-binding determinants and cellular thresholds for human telomerase repeat addition processivity. *EMBO J* **36**, 1908-1927 (2017).

43. Chen, Y.N., Podlevsky, J.D., Logeswaran, D. & Chen, J.J.L. A single nucleotide incorporation step limits human telomerase repeat addition activity. *Embo Journal* **37**(2018).

44. Lei, M., Zaug, A.J., Podell, E.R. & Cech, T.R. Switching human telomerase on and off with hPOT1 protein in vitro. *J Biol Chem* **280**, 20449-56 (2005).

45. Moye, A.L. et al. Telomeric G-quadruplexes are a substrate and site of localization for human telomerase. *Nat Commun* **6**, 7643 (2015).

46. Zaug, A.J., Podell, E.R. & Cech, T.R. Human POT1 disrupts telomeric G-quadruplexes allowing telomerase extension in vitro. *Proc Natl Acad Sci U S A* **102**, 10864-9 (2005).

47. Sen, D. & Gilbert, W. A sodium-potassium switch in the formation of four-stranded G4-DNA. *Nature* **344**, 410-4 (1990).

48. Klenow, H. & Henningsen, I. Effect of monovalent cations on the activity of the DNA polymerase of Escherichia coli B. *Eur J Biochem* **9**, 133-41 (1969).

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

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

51. Hwang, H., Opresko, P. & Myong, S. Single-molecule real-time detection of telomerase extension activity. *Sci Rep* **4**, 6391 (2014).

52. Parks, J.W., Kappel, K., Das, R. & Stone, M.D. Single-molecule FRET-Rosetta reveals RNA structural rearrangements during human telomerase catalysis. *RNA* **23**, 175-188 (2017).

53. Yang, W. & Lee, Y.S. A DNA-hairpin model for repeat-addition processivity in telomere synthesis. *Nat Struct Mol Biol* **22**, 844-7 (2015).

54. Jarstfer, M.B. & Cech, T.R. Effects of nucleotide analogues on Euplotes aediculatus telomerase processivity: evidence for product-assisted translocation. *Biochemistry* **41**, 151-61 (2002).

55. Nguyen, T.H.D. et al. Cryo-EM structure of substrate-bound human telomerase holoenzyme. *Nature* **557**, 190-195 (2018).

56. Jiang, J. et al. Structure of Telomerase with Telomeric DNA. *Cell* **173**, 1179-1190 e13 (2018).

57. Wu, R.A. & Collins, K. Human telomerase specialization for repeat synthesis by unique handling of primer-template duplex. *EMBO J* **33**, 921-35 (2014).

58. Laederach, A. et al. Semiautomated and rapid quantification of nucleic acid footprinting and structure mapping experiments. *Nat Protoc* **3**, 1395-401 (2008).

59. Kuzmic, P. Program DYNAFIT for the analysis of enzyme kinetic data: application to HIV proteinase. *Anal Biochem* **237**, 260-73 (1996).

60. Kuzmic, P. DynaFit--a software package for enzymology. *Methods Enzymol* **467**, 247-80 (2009).

SUPPLEMENTARY METHODS

Kinetic Analysis

The intensities of RAP bands resolved by gel electrophoresis were analyzed globally using DynaFit according to the mechanism shown in Figure 6A [1,2]. All rate constants were floated, except k_{a0} and k_{d0} whose values, within limits, had little effect on the analysis. The value of k_{a0} is beyond the time resolution of the measurement (Fig. S6A) and was fixed at 1 nM^{-1} min^{-1} (with a lower limit of 0.1 nM^{-1} min^{-1}), while k_{d0} was fixed at ≤ 0.1 min^{-1}. The association rate constant for the products to the telomerase enzyme was assumed to have the same k_{a0} value as the primer. While this is unlikely to be strictly true, rebinding of products to the telomerase enzyme is likely to be slower, at least in the early stages of the reaction, when the hot primer remains in >5-fold excess over the sum of the dissociated products. Following a chase with 20 μM cold primer, binding of hot primer and labeled products is negligible and allowed for a more robust analysis without any assumptions regarding the precise values of k_a. The cold chase data were fitted using the "incubate" command to model the preincubation of telomerase with the hot primer.

When modeling more than 10 sequential steps, DynaFit took many minutes to converge to a solution. The analysis could be accelerated by modeling the data in batches that overlapped by 2 repeats. For example, the estimated rate constants for the first 8 steps, obtained from fitting to the first batch of 10 steps, were set as fixed parameters in the analysis of the second batch of data which extended to 18 steps. This simplification is valid because k_f is irreversible so that the steps are decoupled

from each other. Batch wise analysis also allowed Monte-Carlo analysis to be performed on a practical time scale, to check the confidence of the fit and for any covariance between parameters (Fig. S5).

We also attempted to analyze time courses for primer extension over 90 minutes without a cold chase (Fig. S3) in which substrate depletion and product rebinding could be a complicating factor. The initial hot primer showed an initial rapid drop in concentration, followed by a slightly curved decay (Fig. S6A). The amplitude of the initial drop in primer concentration (5.3 nM) matched the initial burst in the sum of the first 3 dominant repeat bands (B, C and D) and provides a measure of the active telomerase concentration [3]. The small deviation of the primer decay profile from linearity indicates that the primer is depleted to an extent that k_{a0}[primer] is no longer >> $k_{d0} + k_{f0}$ and/or the DNA products are rebinding competitively with the primer. The effect is equivalent to deviation of the initial linear rate in classical steady-state enzyme kinetics due to substrate depletion and/or product inhibition. However, at the 20-minute time point, equivalent to the preincubation period in the cold chase experiment, the deviation is small (< 0.4 nM at a remaining primer concentration of 39 nM). Furthermore, when the complete time course of the first 12 repeats (B to M) were globally fitted using DynaFit, the saw-tooth pattern of the rate constants for sequential steps in the reaction remained (Fig. S6B).

The low intensity intervening bands corresponding to nucleotide addition processivity (NAP) were ignored in the analysis above. As a check, we analyzed a data set containing the five NAP bands between two RAP events and found their

inclusion had little effect (< 10% change) on the rate constant estimates, k_f and k_d returned for the RAP bands. The intensity of the NAP bands were insufficient to extract any meaningful rate constants for nucleotide addition and dissociation.

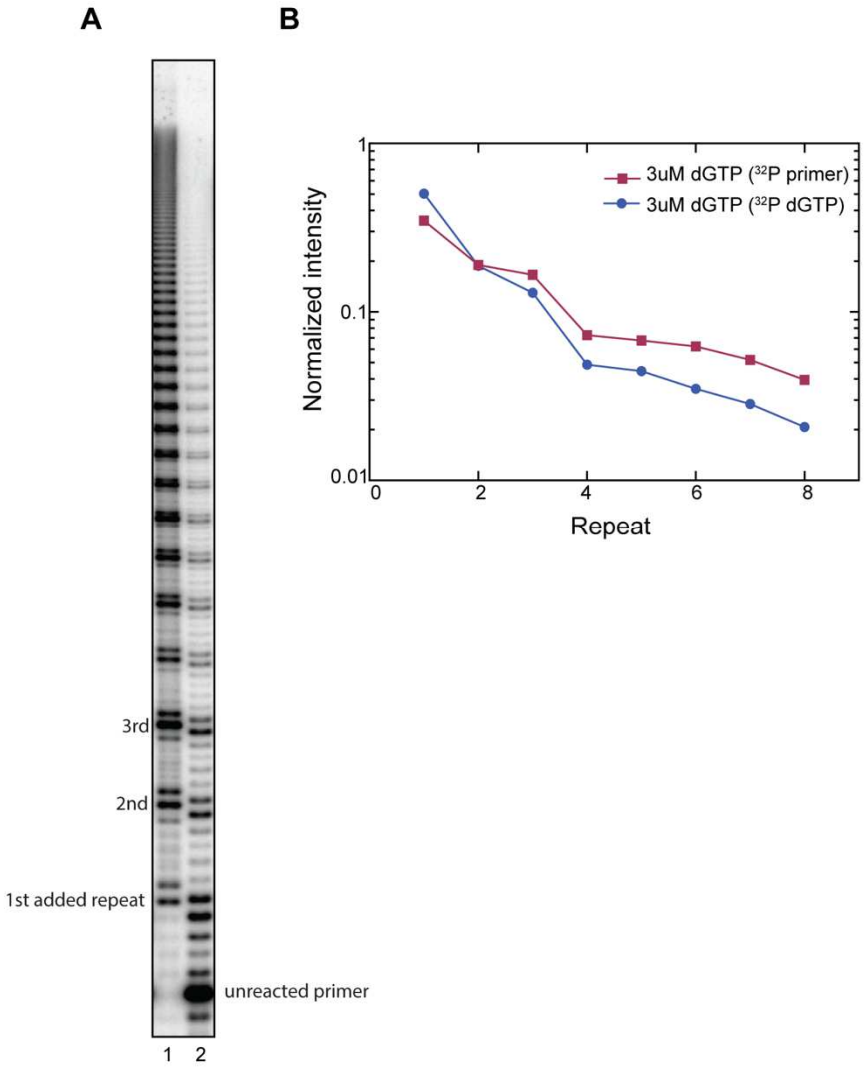


Fig. S1. Using radiolabeled dGTP or radiolabeled primer result in similar product accumulation profiles. (**a**) Telomerase primer extension assay with 3uM dGTP and 500uM dATP and dTTP. Repeats added to the primer are indicated to the left. Lane 1 is performed with 0.3uM radiolabeled dGTP and 2.7uM cold dGTP. Lane 2 is performed with 50nM radiolabeled $(TTAGGG)_3$ primer and 3uM cold dGTP. All assay conditions except for the identity of the radiolabeled nucleic acid are constant between the two samples. **b**) Normalized intensity plotted vs repeat number. Fractional intensity of each band compared to total lane counts is plotted against repeat number.

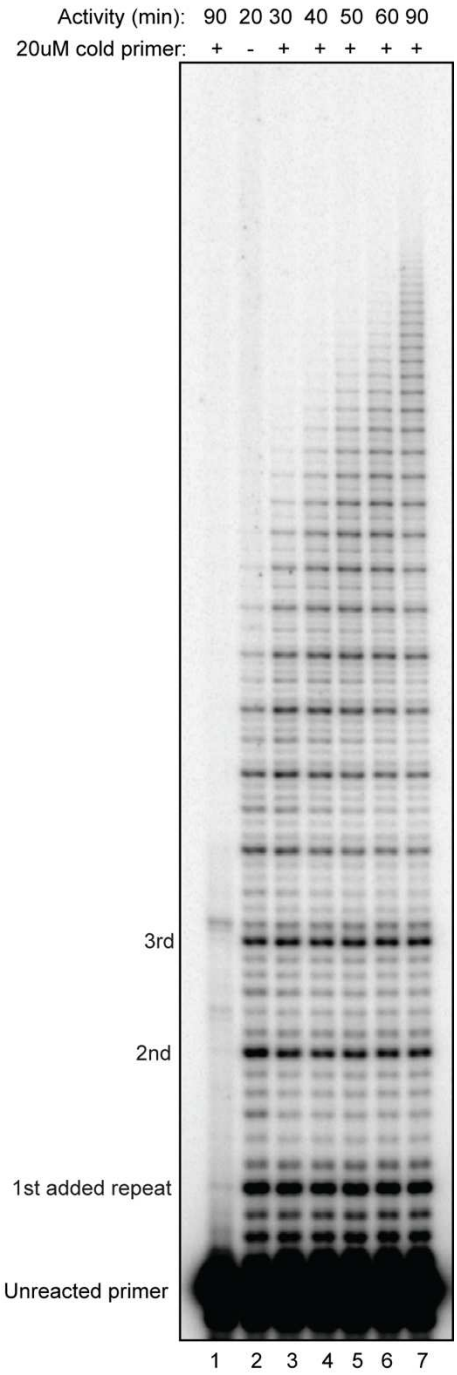


Fig. S2. Telomerase is processive in the presence of 50nM primer. Telomerase primer extension assay with 20uM chase primer $(TTAGGG)_3$ added after 20 minutes. Repeats added to the primer are indicated to the left.

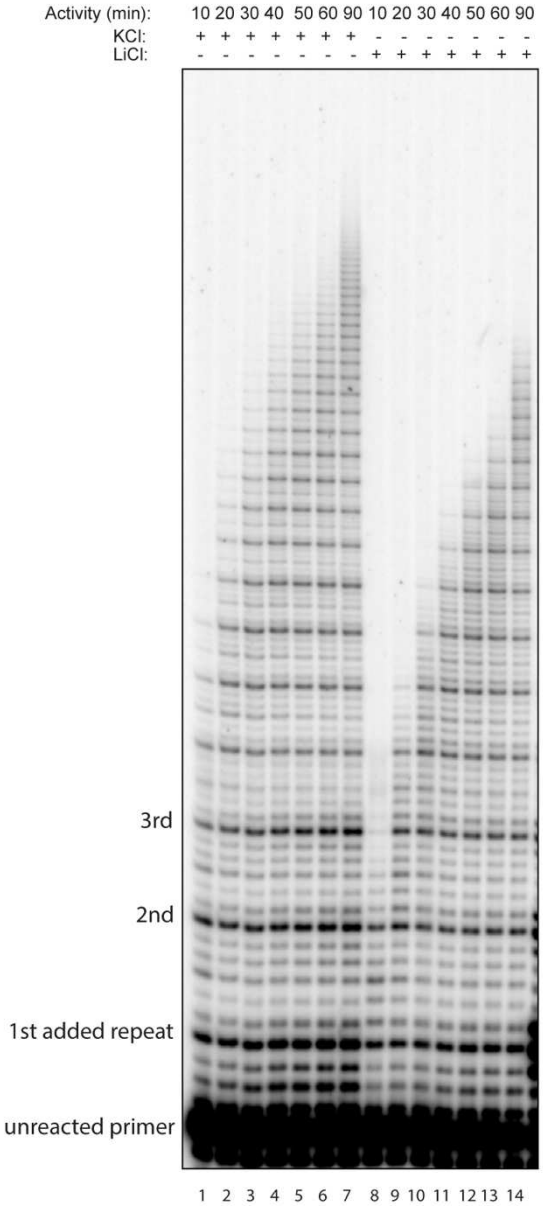


Fig. S3. Telomerase activity is slower in Li^+. Telomerase primer extension assay with samples taken at the indicated time points. Repeats added to the primer are indicated to the left.

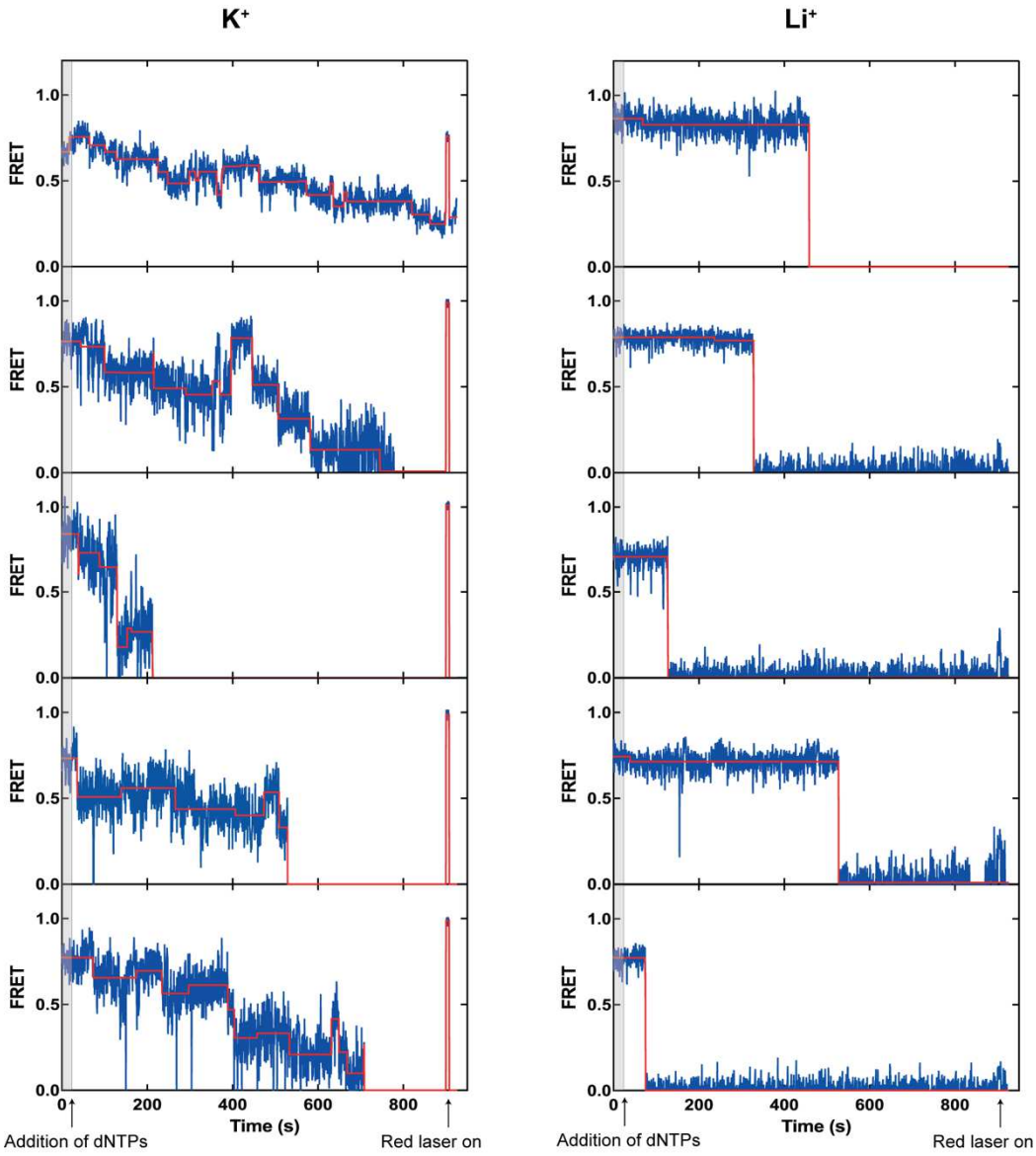


Fig. S4. Real-time smFRET traces. Representative FRET traces are shown for experiments performed in KCl (left panels) and LiCl (right panels). Steps (red) were fit to each trace using MATLAB. Red laser was turned on at 900 (s) to check for Cy5 photobleaching. dNTPs were added after 25 seconds (grey bar).

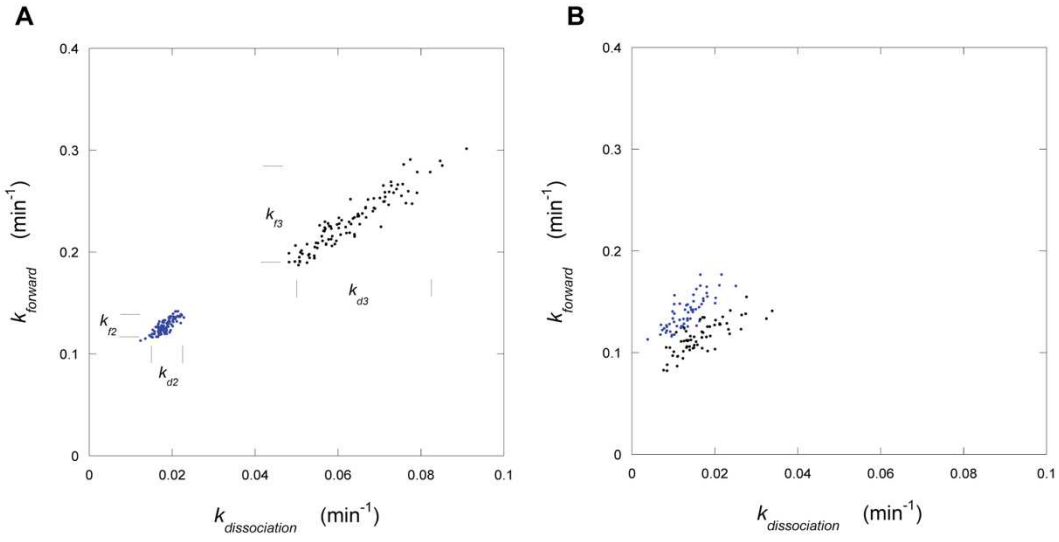


Fig. S5. Example Monte-Carlo analysis to illustrate confidence of fitting and covariance of parameters for the scheme shown in Figure 6A [1,2]. The plots show the rate constants for the dissociation and forward elongation for the repeat bands, C and D (Figure 6A) in the presence of (A) K^+ and (B) Li^+. The analysis shows that k_{f3} is about 2-fold larger than k_{f2} and k_{d3} is about 3 times larger than k_{d2} in the presence of K^+, leading to reduced microscopic processivity ($k_f/(k_f+k_d)$), whereas the differences in the presence of Li+ are marginally significant. The markers in (A) indicate the 5 and 95 percentile limits of the estimated rate constants based on over 500 iterations. This analysis indicates that the saw-tooth pattern of consecutive rate constants and the processivity observed in K^+ is significant compared with the near-featureless pattern in Li^+ (Figure 6C). The analysis also shows that k_f and k_d for each repeat band are linearly covariant, leading to an elliptical distribution. Consequently, the microscopic processivity is better defined than the individual rate constants. No covariance is seen between pairs of other rate constants (e.g. k_{f3} versus k_{d2}).

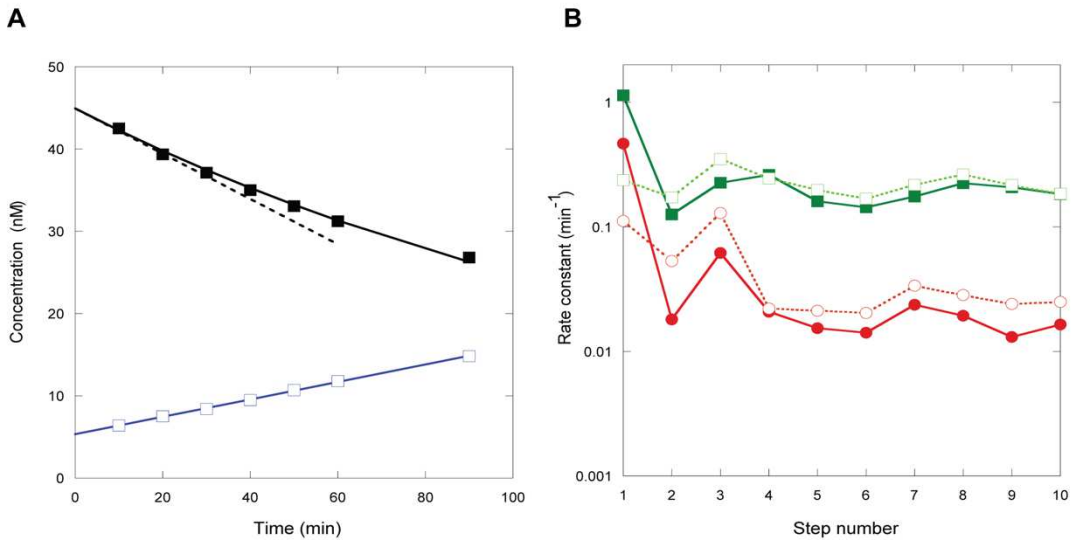


Fig. S6. Reproducibility of kinetic analyses. (A) Primer extension assay performed in the absence of a cold chase, starting with 50 nM hot primer. Note the drop in primer concentration to 45 nM which is complete within the first time point (black solid symbols). This indicates that $k_{on} = k_{a0}$[primer] + k_{d0} + k_{f0} is rapid on the time scale of this experiment and is dominated by k_{a0}[primer] based on modeling (Fig 6D; $k_{f0} \approx 0.1$ min^{-1}) and literature data ($k_{d0} < 0.0006$ min^{-1} [4]. The solid black line shows an exponential fit to the data, and the dashed line shows the equivalent initial rate, to illustrate the small deviation from linearity of the early time points. The initial rapid drop in [primer] is matched by a 5 nM burst in products formation, which is dominated by the first 3 repeat products (blue open symbols = sum B + B$^{\#}$ + C + C$^{\#}$ + D + D$^{\#}$). (B) The rate constants returned by global fitting using DynaFit for primer extension assays carried out in the presence of K^+, with a 20 μM cold primer chase (solid symbols and line = experiment of Figure 6D) compared with a similar experiment (Fig. S3) in the absence of a cold chase (open symbols, dashed line) in which product rebinding may occur during the 90-minute incubation. These data indicate that the saw-tooth profile of rate constants (green squares = k_f and red circles = k_d) is a robust feature of telomerase that is observed using a different experimental design and a different sample preparation.

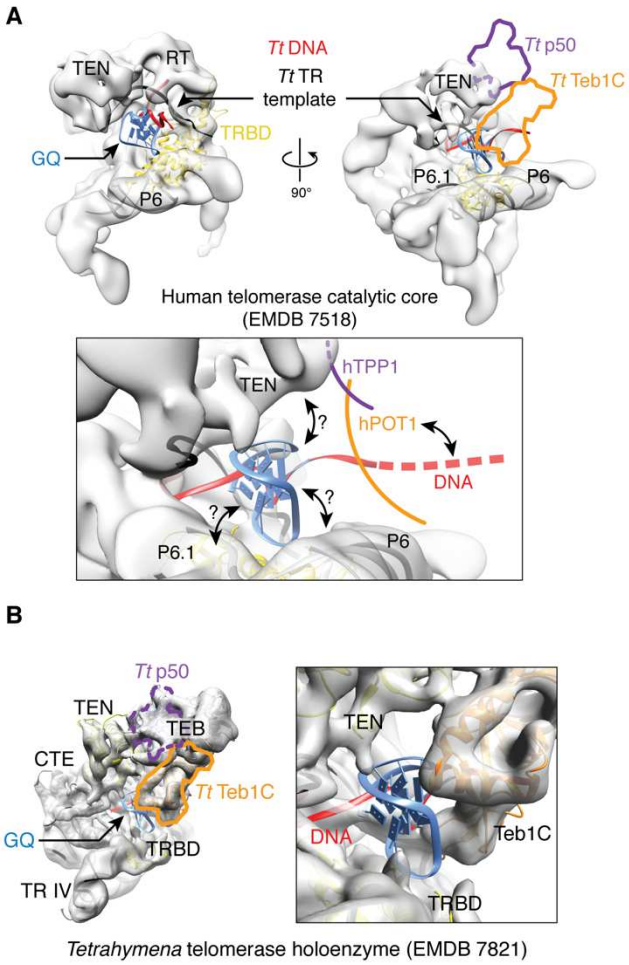


Fig. S7. Model of a GQ on the telomerase enzyme. A. The DNA exit channel in human telomerase provides space to accommodate a GQ. Structural superposition with the Tetrahymena telomerase EM structure places the hPOT1-hTPP1 counterparts TtTeb1-Ttp50 in juxtaposition to the DNA exit upstream of the GQ (outlines at right, see B). Bottom panel, close up of the GQ model. The distance of the GQ to the active site is below two telomeric repeats (12 nt) according to observations presented in this study. In this position the GQ has access to critical elements such as the TEN domain, the TRBD and the hTR CR4/5 to exert its potential impact on telomerase function. Modeled coordinate PDB IDs: GQ, 2HY9; TR template and DNA, 6D6V; TRBD-P6-P6.1, 4O26. Human telomerase catalytic core EM reconstruction, EMDB 7518. B. In Tetrahymena, the modeled GQ pocket is further confined by the C-terminal domain of the TEB complex protein Teb1, illustrating how telomere binding proteins can be excluded from a GQ promoting environment within telomerase holoenzyme complexes. Modeled coordinate PDB IDs: Tetrahymena TERT, TR, p50, Teb1 and DNA, 6D6V; GQ, 2HY9. Tetrahymena telomerase holoenzyme EM reconstruction, EMDB 7821.

Table S1. Oligos used for this study.

$(TTAGGG)_3$ primer	TTAGGGTTAGGGTTAGGG
$(TG)_3$ primer	TGTGTGTTAGGGTTAGGG
$(TG)_6$ primer	TGTGTGTGTGTGTTAGGG
PK 32-62 fragment	GGGCCAUUUU 5-N-U UGUCUAACCCUAACUGAGAA
DNA splint	CAGCGCGCGGGGAGCAAAAGCACGGCGCCTACGC CCTTCTCAGTTAGGGTTAGACAAAAAATGGCCAC CACCCCTCCCAGG
Biotin $(TTAGGG)_3$ primer	Biotin-TTAGGGTTAGGGTTAGGG
Cy5 complementary oligo	5AmMC6-CCCTAACCCTAACCC

REFERENCES

1. Kuzmic, P. Program DYNAFIT for the analysis of enzyme kinetic data: application to HIV proteinase. *Anal Biochem* **237**, 260-73 (1996).
2. Kuzmic, P. DynaFit--a software package for enzymology. *Methods Enzymol* **467**, 247-80 (2009).
3. Hartley, B.S. & Kilby, B.A. The reaction of p-nitrophenyl esters with chymotrypsin and insulin. *Biochem J* **56**, 288-97 (1954).
4. Wallweber, G., Gryaznov, S., Pongracz, K. & Pruzan, R. Interaction of human telomerase with its primer substrate. *Biochemistry* **42**, 589-600 (2003).

CHAPTER 4: Ongoing single molecule FRET studies probing the dynamics of the telomerase enzyme during catalysis

Introduction

Single-molecule techniques are particularly well suited to help tease apart the dynamics that occur during telomerase catalysis and have the added benefit of requiring relatively small amounts of sample material, a feature that is particularly relevant in telomerase research given the challenges associated with producing large amounts of homogeneous enzyme for biochemical and biophysical experiments. As in the case of other reverse transcriptases such as HIV RT, smFRET has so far emerged as the method of choice for studying the assembly and structural dynamics of telomerase at the single-molecule level (2-8).

While the *Tetrahymena* system has provided valuable insight into telomerase function, studies of human telomerase are understandably of greater biomedical relevance. However, because of the larger size of human telomerase, in particular the RNA component (451 nucleotides compared to 159 in *Tetrahymena)*, human telomerase is more difficult to reconstitute into a functional enzyme *in vitro*. To circumvent this problem, a modified reconstitution protocol using only the two critical hTR domains required for catalysis, the PK/template and STE domains was established (10). This approach simplifies the assembly of human telomerase while still maintaining wild type levels of telomerase activity, and also facilitates reconstitution of human telomerase with dye-labeled RNA, an advance that has led to several exciting studies focused on nucleic acid dynamics during human telomerase

catalysis. Studies focused on dynamics of the protein component of telomerase have lagged behind those of nucleic acid dynamics however, due to challenges in site-specific protein labeling. During my graduate career, I have focused on identifying suitable labeling approaches for site-specific TERT labeling in order to study coordinated dynamics of TERT and TR during catalysis. While there is still work to be done before dynamics can be properly measured for TERT using smFRET, in this chapter I outline the progress that I have made towards reaching this goal. Lastly, I discuss preliminary work on the study of real-time telomerase nucleotide incorporation kinetics in collaboration with PacBio and the Puglisi lab at Stanford.

Human telomerase reconstitution and purification

To study human telomerase dynamics using smFRET, we require the incorporation of FRET-compatible fluorophores into the telomerase complex. Our current strategy uses the well characterized Cy3/Cy5 dye pair. We label hTR with Cy3 at U42, a previously determined static site near the active site to serve as a reference point for other potentially dynamic regions of telomerase. Because of the use of dye-labeled RNAs, which have to be synthetically made, we cannot perform our telomerase reconstitutions in cell culture. Instead, we use the Promega optimized transcription and translation system in Rabbit Reticulocyte Lysate (RRL). To this lysate we add FLAG-hTERT plasmid as well as the two in vitro transcribed (or synthesized) required RNA components (as described above in the introduction). Following telomerase assembly, we perform a FLAG-IP to purify the enzyme.

Because TERT is FLAG-tagged, this purification scheme also pulls down large amounts of free TERT, a problem that is discussed in multiple sections below.

Site specific labeling of hTERT

The protein component of telomerase, TERT, is composed of four conserved domains. Two of these domains, the Telomerase Essential N-terminal (TEN) domain and the C-terminal extension (CTE) have been proposed to undergo conformational changes during telomerase catalysis (8, 11-13). However, no direct measurements of TERT dynamics have been achieved due to difficulties in site-specific TERT labeling. TERT contains 29 cysteine residues, which complicates the use of standard maleimide labeling due to the need to delete out all but one cysteine in order to ensure site-specific labeling. While this still remains an avenue to try, it has long been presumed to result in an inactive enzyme, in addition to being fairly labor intensive. This approach should be pursued however if other labeling approaches fail to work.

During a large portion of my graduate work I have focused on inserting small tags into various regions of TERT to serve as sites for enzymatic labeling. Specifically, I have been using the 11 amino acid ybbR tag, an optimized, small peptide derived from the *B. Subtilis* ybbR ORF (14). The ybbR tag gets enzymatically labeled by Sfp phosphopantetheinyl transferase (Sfp) when in the presence of dye-conjugated CoA. Sfp transfers the dye from CoA onto a conserved serine residue on the ybbR tag (**Fig. 1**) (9, 15). The ybbR tag can be added to the termini of protein

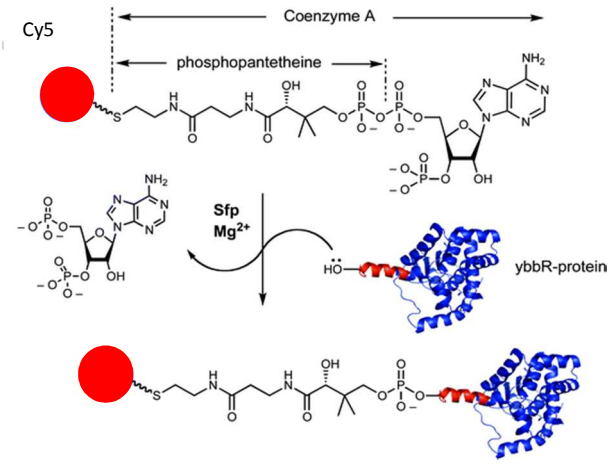


Figure 1. Sfp labeling. A Cy5 dye (red dot) is conjugated to a Coenzyme A (CoA). In the presence of Mg^{2+}, Sfp transfers the Cy5 dye along with the phosphopantetheine linker onto a serine residue on the ybbR tag (red helix). Figure adapted from Wong et al. 2008 (9).

constructs but is also small enough to fit into internal loops or disordered regions. This enables the analysis of individual protein domain dynamics within the context of a larger protein. Our current labeling approach involves labeling hTERT after reconstitution but still within the RRL. After telomerase assembly, we add home-made Sfp enzyme, CoA-Cy5 and 10mM $MgCl_2$. The labeling reaction is then allowed to proceed for one hour at 30C. A standard FLAG-IP is then performed to purify the enzyme and to remove Sfp and free dye.

During my graduate work, I have focused on engineering ybbR tags into both the TEN domain and the CTE for subsequent use in smFRET studies. Using a human homology model of TERT based on secondary structure prediction and previously solved TERT domains from other organisms (1), I have attempted to find viable sites within TERT to insert the ybbR tag. Out of 9 sites tried within the TEN domain and 15 tried in the CTE, only 4 are viable in the TEN domain and 1 in the CTE (**Table 1**

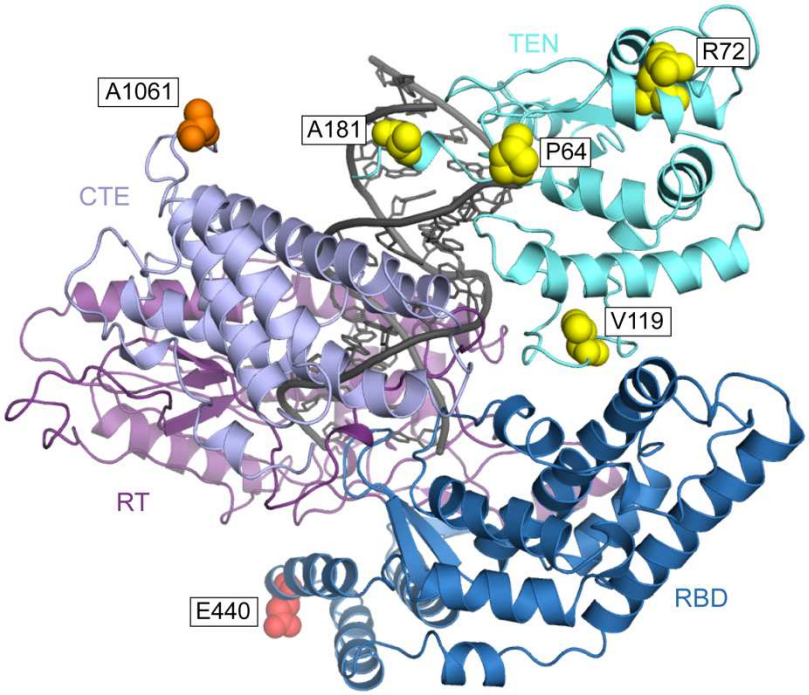


Figure 2. Current ybbR label sites on TERT. The six current, viable ybbR label sites are indicated by spheres in yellow, orange and red for the TEN, CTE and RBD, respectively. The TERT structure is constructed using secondary structure prediction as well as available TERT domain structures from other organisms with the human sequence threaded into it (1). The TEN, RBD, RT and CTE are shown in cyan, dark blue, purple and light violet respectively.

and Fig. 2). I have additionally found a viable label site in the RNA Binding Domain (RBD) with the initial purpose of serving as a static reference site. In total, I have designed six TERT constructs with ybbR tags inserted that when reconstituted and labeled, form a functional telomerase enzyme (**Fig. 2**). All sites within TERT that have been tested are shown in Table 1.

Table 1

Amino acid after which the tag is added	Domain	Tags tried	Viable?
H17	TEN	ybbR	no
R37	TEN	ybbR	no
D43	TEN	ybbR	no
P64	TEN	ybbR, A1	Yes, for ybbR tag only
R72	TEN	ybbR, A1	Yes, for ybbR tag only
D105	TEN	ybbR	no
V119	TEN	ybbR, A1	yes
S165	TEN	ybbR	no
A181	TEN	ybbR, A1	yes
V435	RBD	ybbR	no
P438	RBD	ybbR	no
E440	RBD	ybbR, A1	Yes, for ybbR tag only
K965	CTE	ybbR	no
Q1018	CTE	ybbR	no
L1019	CTE	ybbR	no
P1020	CTE	ybbR	no
F1021	CTE	ybbR	no
H1022	CTE	ybbR	no
G1057	CTE	ybbR, S6	no
A1061	CTE	ybbR, S6	Yes, for ybbR tag only
R1086	CTE	ybbR, S6	no
H1085	CTE	ybbR	no

R1084	CTE	ybbR	no
T1088	CTE	ybbR, S6	no
T1111	CTE	ybbR, S6	no
T1113	CTE	ybbR, S6	no
T1129	CTE	ybbR, S6	no

Table 1. Sites used for labeling tag insertion within TERT. Each site that has been tested for ybbR, S6 or A1 incorporation is shown. Viable indicates comparable telomerase activity to that of WT enzyme.

In addition to the ybbR tag, we tried using the 12 amino acid tags, S6 and A1, which are derived from the PCP (Peptidyl Carrier Protein) and ACP (Acyl Carrier Protein) proteins, respectively (16). The initial idea for using these additional tags was for the purpose of inserting one tag in the CTE and one in the TEN domain to investigate coordinated dynamics between these two domains during catalysis. The benefit of using a combination of these tags is that they can be used for orthogonal labeling by the Sfp and AcpS enzymes since AcpS labels only the A1 tag while Sfp labels all three tags but with a preference for ybbR and S6 (16). So, in theory, one could label the A1 tag first with AcpS followed by labeling of ybbR or S6 by Sfp. However, neither the S6 or the A1 tag proved to be good labeling tags due to diminished telomerase activity when inserted at many of the same sites that had been proved to be viable for the ybbR tag (data not shown). Most likely this is due to a loss of solubility after insertion of these other tags at many of the insertion sites. So, in terms of small protein tags for use in enzymatic labeling, the ybbR tag seems to be

the best candidate for telomerase studies. Additionally, orthogonal labeling of multiple protein domains still remains an area to be optimized.

The six viable ybbR constructs identified to date have been successfully reconstituted with Cy3-labeled hTR, Sfp labeled and purified. However, when analyzed during preliminary smFRET experiments (**Fig. 3**), we noticed that the labeling efficiency was very low (1-20%), and that it varied depending on the site of the ybbR tag insertion. We additionally noticed a high background of Cy5-labeled TERT sticking to the slide surface. This indicates that while TERT itself can be efficiently labeled by Sfp, once reconstituted into a functional enzyme, something is preventing efficient labeling, most likely steric hindrance by the RNA. The fact that we see such a high variability in labeled enzyme depending on the label site further supports this hypothesis since different label sites will be more or less occluded by hTR. To get around this problem, I attempted to label TERT first followed by assembly with hTR. However, this strategy did not positively impact the amount of active, labeled enzyme (data not shown), most likely due to telomerase assembly being more efficient when TERT is translated and folded in the presence of hTR. While labeling efficiency remains a problem to be solved for properly studying TERT dynamics, I obtained preliminary smFRET data for each of the TEN-ybbR sites to a static label site on hTR (U42) (**Fig. 4**). When immobilized and stalled on the slide surface using a biotinylated telomere primer ($(TG)_6TTAGGG$), most label sites display single peak FRET distributions with the exception of A181 which shows two

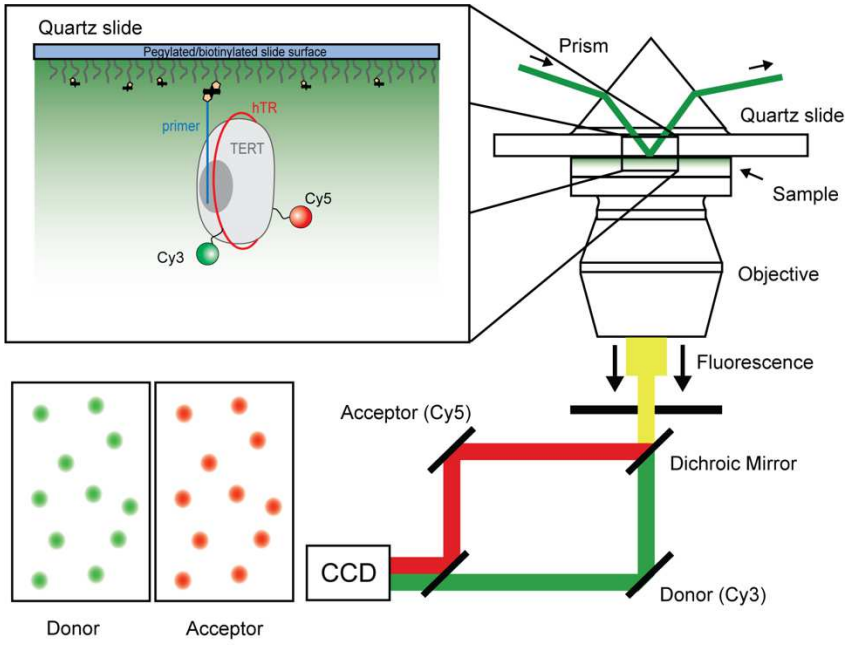


Figure 3. smFRET experimental setup. smFRET measurements are taken using a Total Internal Reflection Fluorescence (TIRF) microscope. Telomerase (TERT in grey and hTR in red) is immobilized onto a quartz slide onto a telomere primer (blue) via biotin/streptavidin linkage. TERT is labeled with Cy5 (red) at the indicated ybbR insertion site and hTR is conjugated to a Cy3 dye (green). The appropriate excitation laser is used to illuminate the sample at an angle, which causes total internal reflection of the incoming light due to the difference in the index of refraction of the quartz slide and the sample volume. In this way, the intensity of the evanescent field of light that propagates into the sample volume decays exponentially and dramatically enhances the signal-to-noise ratio of the measurement. The emitted fluorescence of several hundred molecules in a single field of view is split using dichroic mirrors into separate acceptor and donor channels and captured by the EMCCD camera. Automated software routines are then used to analyze the raw movie data, and the FRET ratio can be calculated according to the simple expression $FRET = I_A / (I_D + I_A)$, where I_A and I_D are the background corrected intensities of the acceptor and donor dyes, respectively.

separate peaks, suggesting that this region of the TEN domain might exist in two different conformations. Additionally, the CTE label site (A1061) shows a single low FRET population (data not shown). While monitoring changes in conformation of

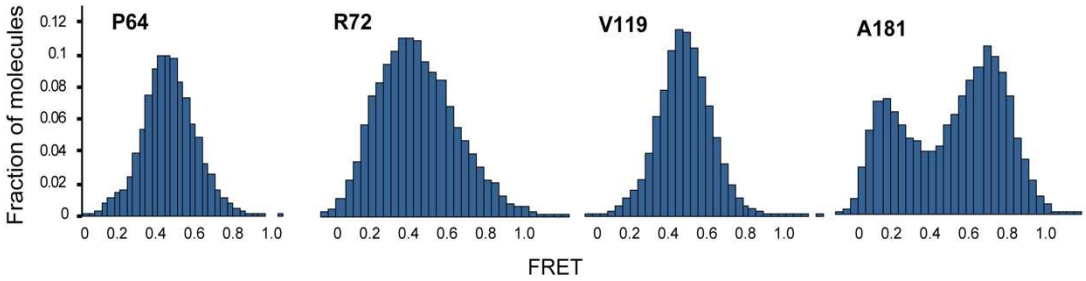


Figure 4. FRET distribution for ybbR-TEN enzymes. Each enzyme was immobilized using biotinylated $(TG)_6TTAGGG$ primer and has a Cy3 dye on U42 of hTR and a Cy5 dye at the indicated ybbR insertion site. Fraction of molecules is shown on the y-axis with FRET value on the X-axis.

various TERT domains during activity has been hard to achieve due to the combination of low labeling efficiency and telomerase partially eluting off of the primer when in the presence of dNTPs, the FRET measurements of the stalled enzyme might still be useful for modeling the proper conformation of the TEN domain within the telomerase holoenzyme. While potentially informative, these initial measurements should be taken with a grain of salt since we cannot exclude the possibility that the small amount of enzyme that gets labeled represents misfolded or inactive enzyme. To get around this problem, we optimized a purification scheme to obtain stocks of pure active enzyme (See **Activity-dependent purification of telomerase below** below).

For future studies of TERT dynamics, there are multiple approaches that have yet to be tried. One of these is to label the telomerase enzyme outside of the RRL. For example, one could label the enzyme during the FLAG-IP step. However, since the FLAG tag is attached to the N-terminus of TERT, this might preclude some of the

TEN domain sites from being efficiently labeled. Varying the time, temperature and amount of Sfp during the labeling reaction also remains to be fully explored.

Activity-dependent purification of human telomerase

To exclude the possibility of artifacts being introduced during smFRET measurements due to the low labeling efficiency combined with the excess Cy5-labeled TERT stuck on the slide surface, we optimized an approach originally proposed by Scott Cohen and Roger Reddel(17) to purify and immobilize only active, properly assembled telomerase enzymes (**Fig. 5**). This way, even if the labeling efficiency is poor, we can be confident that any labeled enzyme represents properly folded and active telomerase. Instead of using a FLAG-IP for purification which pulls down a lot of excess free TERT in addition to assembled telomerase, we used beads tethered to a telomere primer terminating in TTAGGG (which is the highest affinity binding site of telomerase to the telomere primer (18)). After washing enzymes bound to these beads (to get rid of excess free TERT), we add dTTP and dATP to elute the enzyme off the primer. This method works due to telomerase having a drastically lower affinity for the telomere primer once TTA is incorporated (TTAGGG**TTA**). This will cause active telomerase to become free in solution. To prevent free TERT from reanneling to bead-bound primers, we add a capture oligo that perfectly anneals to the full hTR template, thereby providing higher affinity for the enzyme than the primer bound to the beads. To obtain active, unbound enzyme,

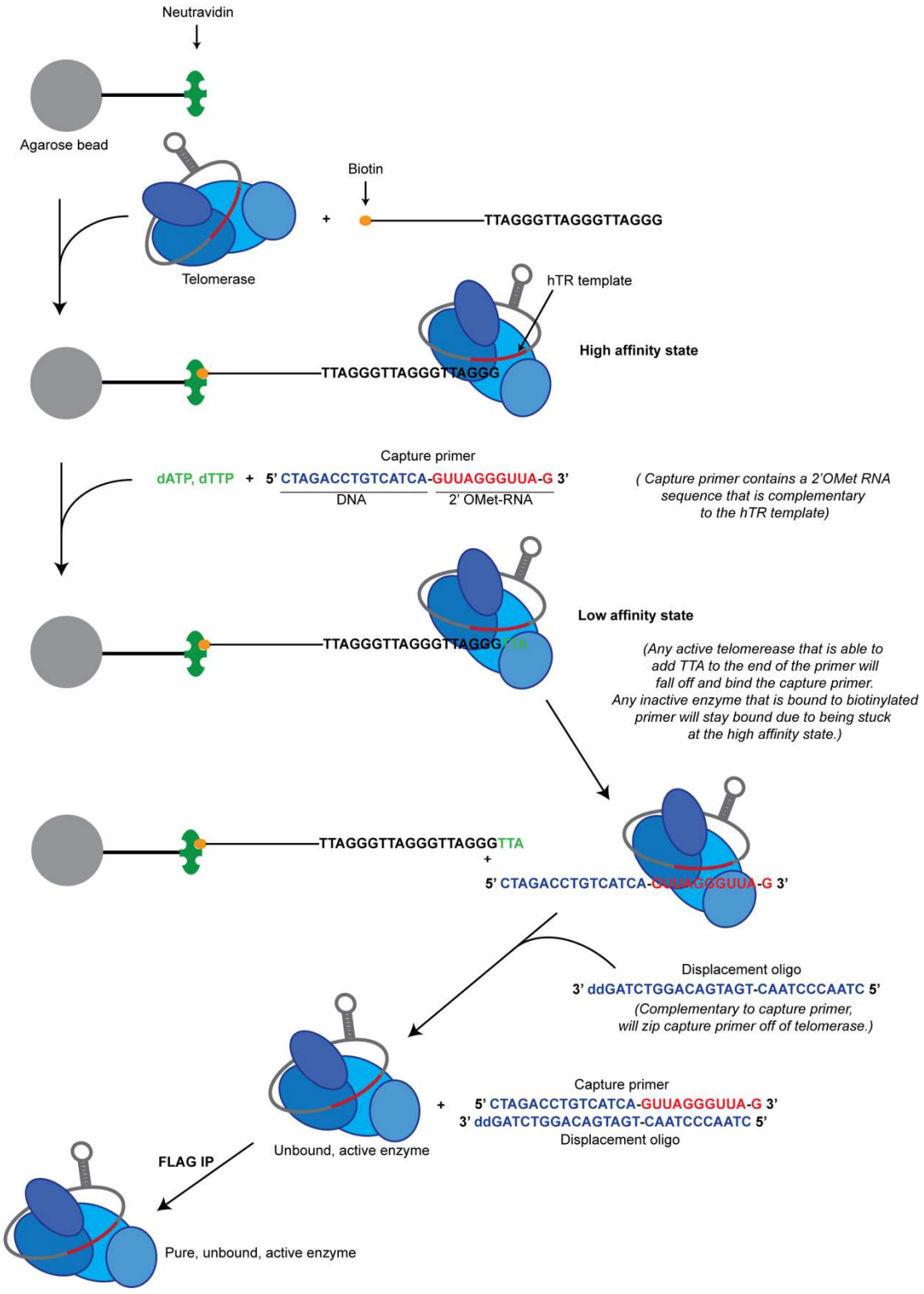

Neutravidin
Agarose bead
Biotin
TTAGGGTTAGGGTTAGGG
Telomerase
hTR template
High affinity state
TTAGGGTTAGGGTTAGGG
Capture primer
dATP, dTTP + 5' CTAGACCTGTCATCA-GUUAGGGUUA-G 3'
DNA
2' OMet-RNA
(Capture primer contains a 2'OMet RNA sequence that is complementary to the hTR template)
Low affinity state
TTAGGGTTAGGGTTAGGGTTA
(Any active telomerease that is able to add TTA to the end of the primer will fall off and bind the capture primer. Any inactive enzyme that is bound to biotinylated primer will stay bound due to being stuck at the high affinity state.)
TTAGGGTTAGGGTTAGGGTTA
+
5' CTAGACCTGTCATCA-GUUAGGGUUA-G 3'
Displacement oligo
3' ddGATCTGGACAGTAGT-CAATCCCAATC 5'
(Complementary to capture primer, will zip capture primer off of telomerase.)
Capture primer
5' CTAGACCTGTCATCA-GUUAGGGUUA-G 3'
3' ddGATCTGGACAGTAGT-CAATCCCAATC 5'
Displacement oligo
Unbound, active enzyme
FLAG IP
Pure, unbound, active enzyme

Figure 5. Activity-dependent purification of telomerase. Telomerase (blue) is pre-bound to a biotinylated telomere primer and the telomerase/primer complex is then immobilized on neutravidin resin. After addition of dATP and dTTP as well as a capture primer, active telomerase is eluted off the primer and is bound by the capture primer. The capture primer contains 5' sequence which is complementary to a displacement oligo, which is then added to strip telomerase off the capture primer. Telomerase (which contains an N-terminal FLAG tag) is then purified away from the capture primer/displacement oligo dimer through a FLAG-IP step.

we then add a displacement oligo that is perfectly complementary to the capture oligo to release telomerase from the capture oligo. A FLAG IP step is then performed to purify away the active enzyme from the high concentration of capture primer/displacement oligo dimers. While this approach yields a population of pure and active telomerase, the multiple purification steps result in a very low concentration of enzyme. When combining Sfp labeling with the activity-dependent elution protocol, a prohibitively small fraction of labeled enzyme is available to study, rendering this approach unsuitable for our current smFRET setup. However, this purification scheme has proven useful for other studies (see **Real-time telomerase dynamics** below).

Future optimization of the activity-dependent elution of telomerase should focus on increasing the yield of the pure enzyme. One approach to further explore is using neutravidin resin to deplete the capture primer/displacement oligo dimer that is left after the elution rather than a FLAG IP step which further reduces the final enzyme concentration. By attaching a biotin moiety to the displacement oligo, the

excess nucleic acid could theoretically be bound by neutravidin resin and removed from the pure enzyme in solution. Initial experiments using this approach have yielded promising results (data not shown) but the low binding capacity of the neutravidin resin used resulted in multiple resin batches having to be used, which further diluted the enzyme. Using this approach with high-capacity neutravidin resin or magnetic beads could be tried to improve the yield of this method.

Effects of MgCl2 on telomerase activity

An interesting finding discovered during the TERT labeling optimization was that the amount of Mg^{2+} present in the RRL during Sfp labeling is inhibitory for telomerase activity. Sfp requires $MgCl_2$ for proper function (10mM), so we initially added 10mM $MgCl_2$ to the RRL during labeling as per labeling instructions (15). However, we noticed very low yields of enzyme when initially performing these studies (data not shown). To determine the cause of this low yield, we added each labeling component to telomerase individually and subsequently measured activity. To our surprise, the addition of 10mM $MgCl_2$ to the lysate almost completely abolished telomerase activity (**Fig. 6**). These results indicate that the concentration of Mg^{2+} in the RRL is already fairly high (although we cannot know the exact concentration due to proprietary reasons). To test if the normal concentration of Mg^{2+} present in the lysate is disruptive for telomerase activity, we performed an EDTA titration assay to determine the optimal Mg^{2+} concentration for telomerase activity

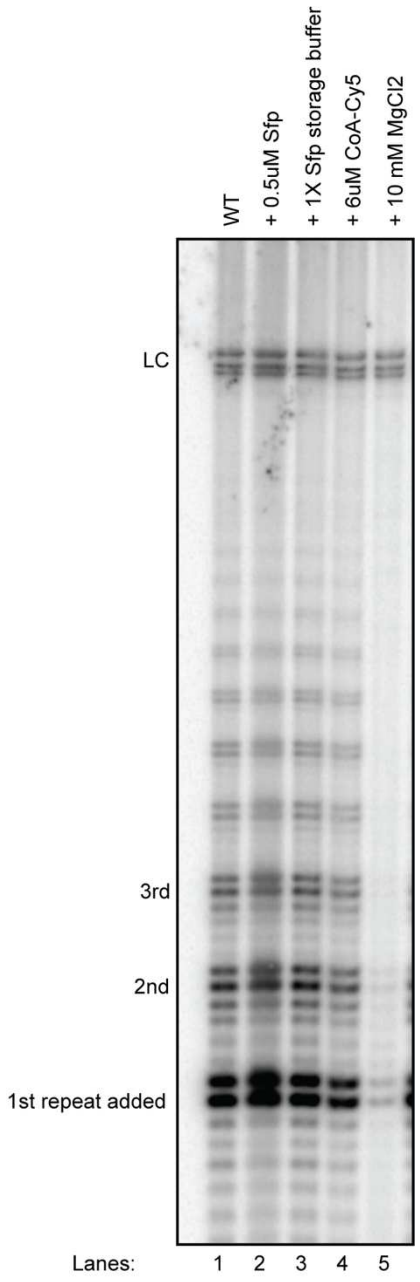


Figure 6. Excess $MgCl_2$ disrupts telomerase activity. A primer extension assay with 0.3uM P32-dGTP and 500uM dATP and dTTP was performed with 1uM unlabeled $(TTAGGG)_3$ primer. 5ul of telomerase lysate was used per reaction and incubated at 30C for 90 minutes. A 159 nt RNA was used as a loading control (LC).

performed in RRL. To our surprise, adding an additional 10mM EDTA to the lysate after the reconstitution, (or 20mM after labeling) resulted in significantly higher telomerase activity (**Fig. 7**) indicating that removing 10mM of the Mg^{2+} already present in the lysate is beneficial for telomerase activity. It yet remains to be

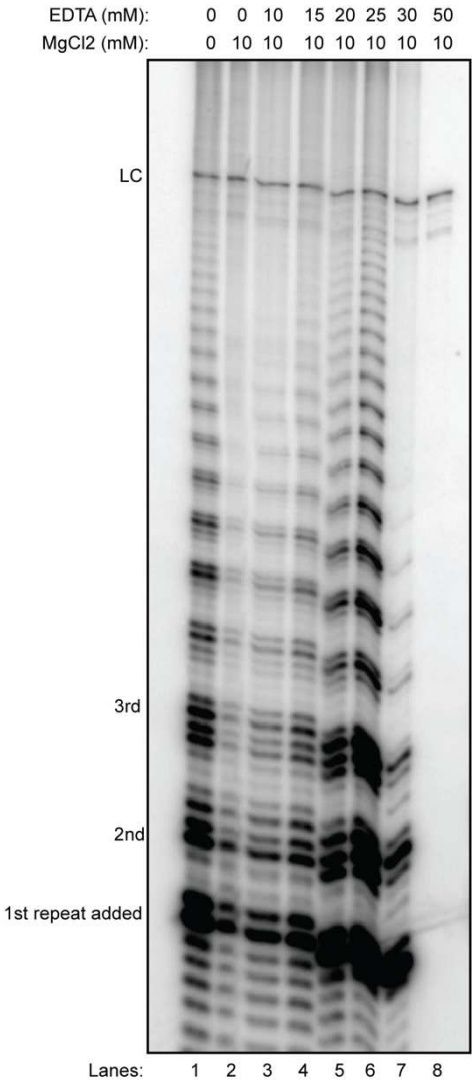


Figure 7. Reducing $MgCl_2$ in the RRL improves telomerase activity. A primer extension assay with 0.3uM P32-dGTP and 500uM dATP and dTTP was performed with 1uM unlabeled $(TTAGGG)_3$ primer. 5ul of telomerase lysate was used per reaction and incubated at 30C for 90 minutes. A 159 nt RNA was used as a loading control (LC).

determined if this effect of Mg^{2+} on telomerase is affecting folding/assembly of the enzyme or if the excess Mg^{2+} is inhibitory for primer extension. This can easily be tested by adding EDTA to half of the lysate before the purification step and then assaying both batches of enzyme after the IP. If Mg^{2+} affects folding/assembly, then there will be a noticeable increase in activity for the enzyme that was treated with EDTA pre-purification. If Mg^{2+} is inhibitory for primer extension then both batches will be equally active provided they are treated with the same Mg^{2+} concentrations during the primer extension assay. Regardless, if assaying telomerase activity directly out of the RRL, 10mM EDTA should be added for maximum telomerase activity.

Effects of POT1/TPP1 on telomerase activity and dynamics

While studying TERT dynamics during catalysis, we also attempted to investigate the effects of POT1/TPP1 binding on telomerase dynamics. To do this, we added 500nM of homemade POT1/TPP1 heterodimer to Cy3-U42/Cy5-TEN-ybbR labeled telomerase prior to surface immobilization. We then immobilized the telomerase/POT1/TPP1 complex onto a biotinylated telomere primer and investigated any differences in FRET distribution. We did this for all four TEN-ybbR constructs in the presence or absence of POT1/TPP1 (**Fig. 8**). We observed a general shift towards lower FRET for all TEN-ybbR mutants that can be explained by two main theories. Either POT1/TPP1 induces a conformational change in the TEN domain when bound to the telomere, causing it to move further from the Cy3 dye, or POT1 outcompetes telomerase for binding to the telomere and the low FRET observed is a result of

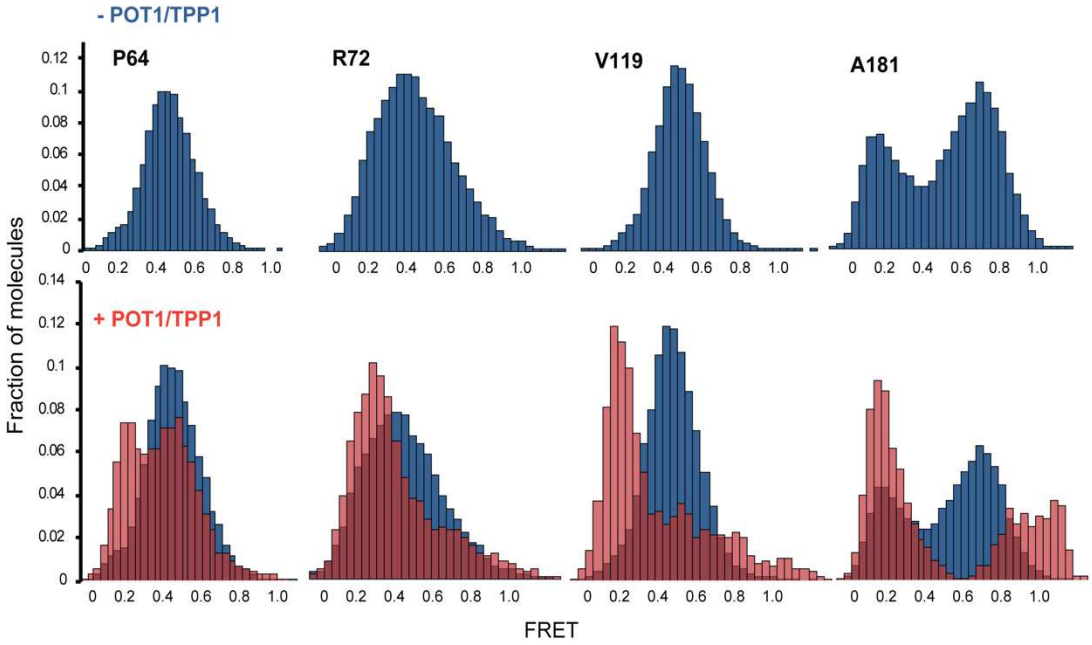


Figure 8. POT1/TPP1 induces a shift in TEN-ybbR mutant FRET populations. Each enzyme has a Cy3 dye on U42 of hTR and a Cy5 dye on at the indicated ybbR insertion site. Fraction of molecules is shown on the y-axis with FRET value on the X-axis. Histograms for data generated in the absence of POT1/TPP1 are shown in blue and ones generated in the presence of POT1/TPP1 are shown in red.

telomerase staying associated with POT1/TPP1 but is itself not engaged with the primer. We observed a general shift towards lower FRET for all TEN-ybbR mutants that can be explained by two main theories. Either POT1/TPP1 induces a conformational change in the TEN domain when bound to the telomere, causing it to move further from the Cy3 dye, or POT1 outcompetes telomerase for binding to the telomere and the low FRET observed is a result of telomerase staying associated with POT1/TPP1 but is itself not engaged with the primer. This might cause a distortion of the TEN domain, situating it further from the Cy3 dye, resulting in the low FRET values observed.

To better study the effects of POT1/TPP1 on telomerase conformation and dynamics, several issues should first be resolved. First, the order of addition of telomerase, POT1/TPP1 and primer to the slide surface should be optimized. Due to the vast excess of POT1/TPP1 to both telomerase and primer in our preliminary experiments, we are most likely preventing telomerase from accessing much of the immobilized primer. An alternative strategy to try should be to immobilize telomerase to the primer first, followed by addition of POT1/TPP1. Allowing enough access for POT1/TPP1 to bind might be an issue however due to the limited amount of space for POT1/TPP1 to bind between telomerase and the slide surface. This problem could be overcome by adding random sequence to the 5' end of the biotinylated primer. Another thought to keep in mind is that there might be a preferential order of binding of POT1/TPP1 to telomerase and the telomere in vivo. Trying different order of addition approaches would be best to ensure that a physiologically relevant conformation is observed. Another strategy to try is to immobilize telomerase/POT1/TPP1 with a mutant telomere primer that forces POT1 to bind to the 5' end of the primer, thereby preventing competition for the 3' end of the telomere with telomerase. This approach has been used successfully for monitoring POT1/TPP1 effects on telomerase activity in vitro (19-21).

In addition to smFRET experiments, we additionally performed bulk biochemical assays on all four TEN-ybbR mutants in the presence and absence of POT1/TPP1. For all but one of the ybbR mutants, we observed an expected increase

in processivity in the presence of POT1/TPP1 (**Fig. 9**) (19). For the V119-ybbR mutant however, we did not observe this same boost in processivity, but rather saw a decrease in overall activity. This either indicates that introducing a ybbR tag after

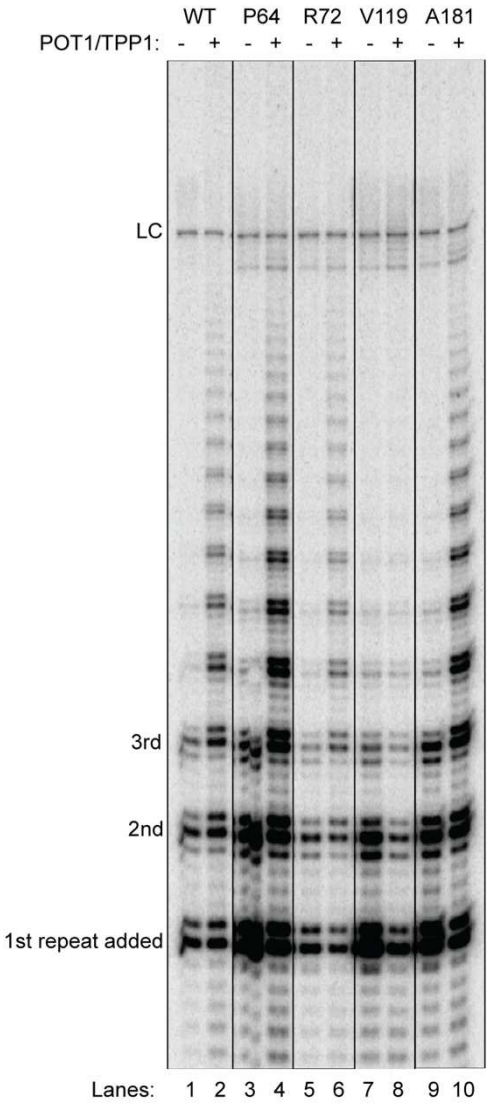


Figure 9. POT1/TPP1 effects on TEN-ybbR mutants.
A primer extension assay with 0.3uM P32-dGTP and 500uM dATP and dTTP was performed with 50 nM unlabeled $(TTAGGG)_3$ primer. 500nM POT1/TPP1 was added to 5ul of purified telomerase. A total reaction volume of 15ul was used per sample and incubated at 30C for 90 minutes. A 159 nt RNA was used as a loading control (LC).

V119 perturbs the telomerase/TPP1 interphase or, if POT1/TPP1 induces a conformational change in the TEN domain, then in this altered conformation the position of the ybbR tag may be disruptive for primer extension. Regardless, this result might provide additional clues about how telomerase interacts with TPP1, an interaction which is currently not well understood.

Future optimizations of human telomerase smFRET experiments

Currently, the above-mentioned experiments have been performed using standard smFRET and TIRF microscopy to maximize the number of FRETing molecules available to study. However, a caveat of our current method is that it is hard to distinguish a change in FRET as being due to conformational changes of active molecules from FRET distributions from inactive or misfolded enzymes left behind after active molecules have eluted off the biotinylated primer. Another caveat is that an apparent low FRET value can be hard to distinguish from Cy3 bleed through into the Cy5 channel. To account for this, the red laser has to be turned on at some point during data collection to ensure that there is a Cy5 dye present on the enzyme. However, this still does not eliminate the possibility that a very low FRET value is actually due to Cy3 bleed through. All of these caveats can be overcome by performing real-time smFRET experiments to monitor conformational changes during telomerase activity. However, at the current labeling efficiency, this would result in a prohibitively low number of dye-labeled enzymes to analyze since only one field of view can be imaged per experiment.

Additionally, the recent report of a mid-angstrom human telomerase cryo-EM structure should greatly aid in designing good label sites that are minimally occluded by hTR (22). Also, attaching a flexible linker to the ybbR tag may be an option to try to get around the steric hindrance imposed by hTR.

Ongoing single molecule studies of telomerase dynamics and kinetics in real time

While standard smFRET studies using TIRF microscopy are illuminating regarding telomerase dynamics, the inherently low required concentration of labeled substrate prevents the use of physiologically relevant dye-labeled dNTPs to monitor the kinetics of nucleotide incorporation. The recent development of SMRT sequencing by PacBio has aided in the technology to do just that using DNA polymerase with the intention of improving long read sequencing. We have recently set up a collaboration with PacBio as well as the Puglisi lab at Stanford to use PacBio technology to investigate nucleotide incorporation dynamics in real time by telomerase. SMRT sequencing utilizes zero mode waveguide technology (**Fig. 10**) which allows for high local concentration of dye-labeled nucleotides. Tens of thousands of zeptoliter wells are located on a small chip, allowing them all to be illuminated simultaneously by the desired wavelength of light. The zeptoliter volume prevents illumination of anything past the bottom surface of the well due to the well size being smaller than the wavelength of light being excited. This allows for the addition of higher concentrations of dye-labeled substrates since the local concentration at the bottom of each well is minimal. In addition, each nucleotide

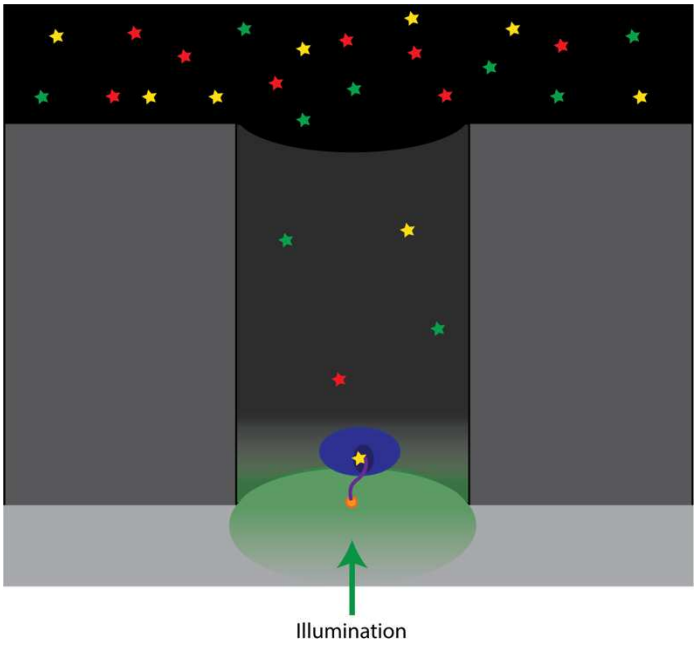


Figure 10. Zero Mode Waveguide technology. Telomerase (blue) is immobilized to the bottom of a well using a biotinylated (orange) telomere primer (purple). Dye-labeled dNTPs (yellow, red and green stars) are free in solution. The chip is illuminated from below, and because the well volume is smaller than the wavelength of the light used (green arrow), only the very bottom of the well is illuminated. When a nucleotide binds in the active site of telomerase, the nucleotide-conjugated fluorophore is excited, producing detectable emission. Because the local concentration of nucleotide in the detectable range at the bottom of each well is so low, this allows for the addition of high overall dye-dNTP concentrations. After nucleotide incorporation, the dye gets cleaved off, preventing a buildup of high signal at the well surface. (Only three dye-dNTPs are depicted since telomerase only used dATP, dTTP and dGTP.)

added is phospho-linked, meaning that the dye gets cleaved after each nucleotide incorporation, thereby preventing a buildup of dye signal within each well.

Before attempting to use this ZMW technology for our telomerase studies, we performed a series of experiments to analyze what parameters and experimental

conditions are optimal for telomerase activity. First, we tested the ability of telomerase to use the PacBio phospho-linked nucleotide analogs during nucleotide incorporation. These analogs contain hexaphosphates instead of the traditional tripohosphates with a dye-moiety conjugated to the 5' phosphate. First, we treated both WT and PacBio dNTPs with Shrimp Alkaline Phosphatase (SAP) to ensure that the PacBio analogs were not contaminated with trace amounts of WT nucleotides due to hydrolysis of the analogs. The phospho-linked analogs are resistant to cleavage by SAP due to having the dye attached to the 5' phosphate, therefore if they are amenable for use by telomerase we would expect no difference in activity in the presence or absence of SAP. Surprisingly, what we see is that the use of the PacBio analogs result in slightly better telomerase activity than WT nucleotides (**Fig. 11**), perhaps because the presence of hexaphosphates may allow for more efficient hydrolysis during nucleotide incorporation compared to triphosphate-containing nucleotides. We additionally performed a time course and nucleotide concentration titration for both WT and PacBio analogs to determine optimal reaction conditions. We see a fair amount of activity after 15 minutes with the PacBio analogs but significantly less with the WT nucleotides (**Fig. 12**). Regardless, this result argues that 15-30 minute long data collections should be sufficient when using the ZMW technology together with the PacBio analogs. Additionally, we see good levels of telomerase activity in the low micromolar range of nucleotides, indicating that these are the concentrations we preferentially want to perform our experiments in (**Fig. 13**).

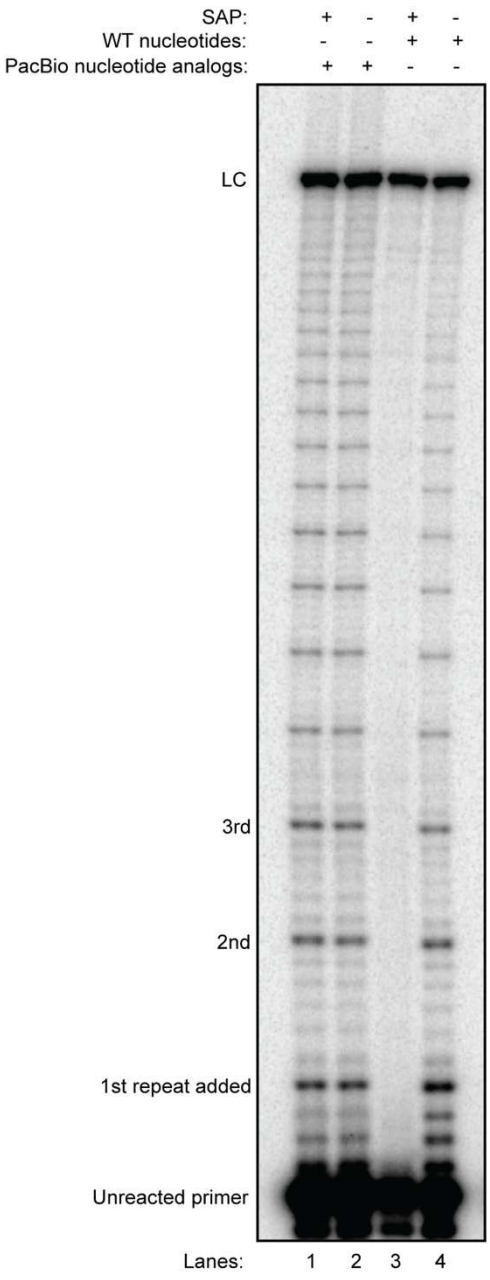


Figure 11. Telomerase is active in presence of PacBio nucleotide analogs. A primer extension assay with 8uM dGTP, dTTP and dATP was performed with 50 nM P32-labeled $(TTAGGG)_3$ primer. A total reaction volume of 15ul was used per sample and incubated at 30C for 90 minutes. A 159 nt RNA was used as a loading control (LC).

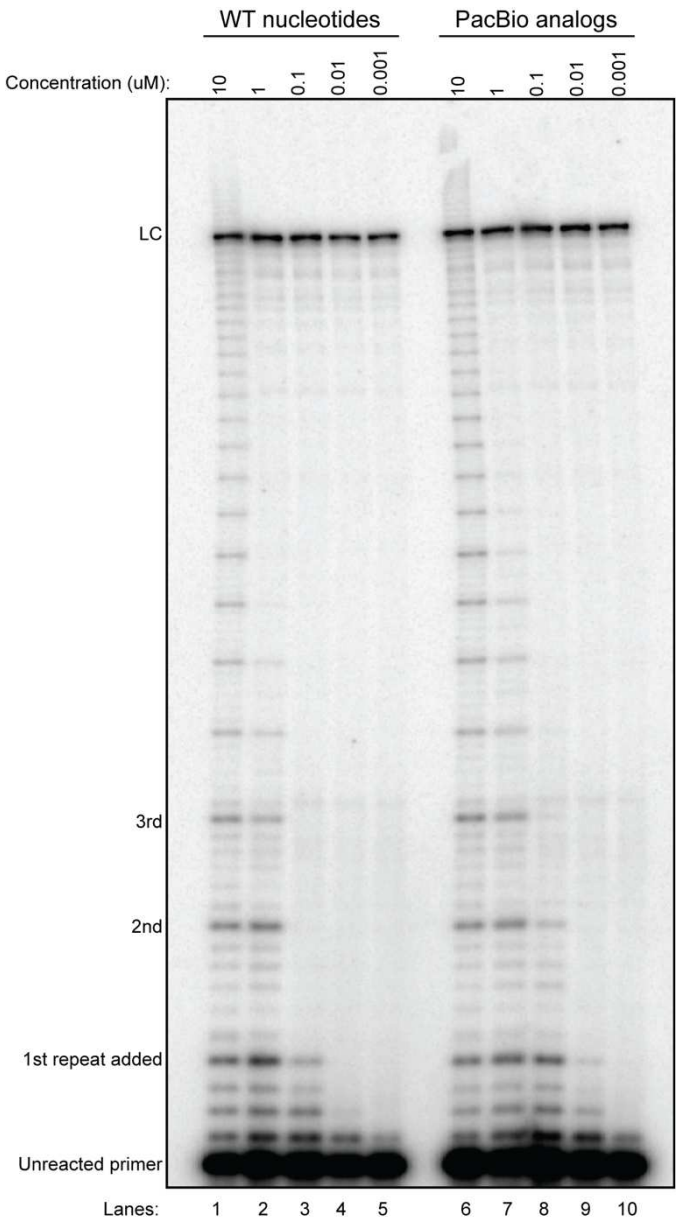


Figure 12. Telomerase nucleotide titration.
A primer extension assay with 8uM dGTP, dTTP and dATP was performed with 50 nM P32-labeled $(TTAGGG)_3$ primer. A total reaction volume of 15ul was used per sample and incubated at 30C for 90 minutes. A 159 nt RNA was used as a loading control (LC).

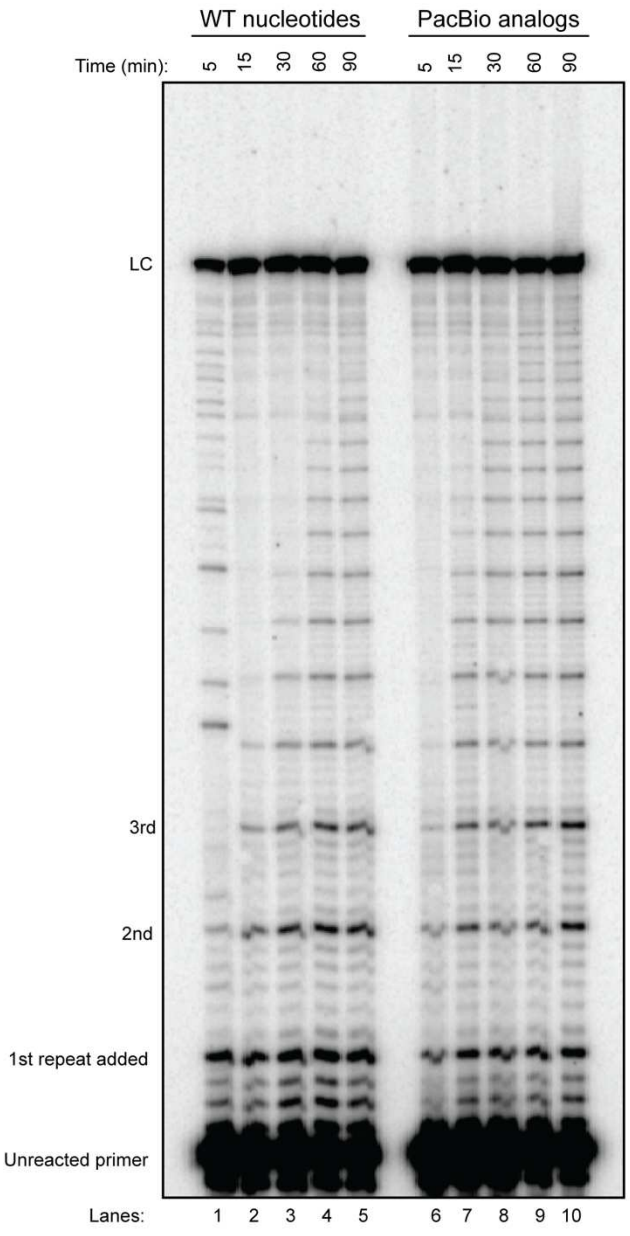


Figure 13. Telomerase activity time course.
A primer extension assay with 8uM dGTP, dTTP and dATP was performed with 50 nM P32-labeled $(TTAGGG)_3$ primer. A total reaction volume of 15ul was used per sample and incubated at 30C for 90 minutes. A 159 nt RNA was used as a loading control (LC). A small amount of nuclease contamination is responsible for the bands in lane 1.

After determining the ideal telomerase reaction conditions, we performed preliminary experiments using the PacBio ZMW technology to study human telomerase kinetics in real time. To do this, we first reconstitute human telomerase with Cy3-labeled hTR (U42) in RRL followed by activity-dependent purification to prevent sticking of free TERT to the ZMW surface. The presence of free TERT on the surface might introduce artifacts due to TERTs ability to also bind free nucleotide. We then immobilize the activity-dependently purified telomerase to the chip surface using biotinylated telomere primers as currently performed for our standard smFRET TIRF experiments. At present, we have performed experiments using 100nM PacBio nucleotide analogs and collected 15-minute long movies. We have not currently been able to increase the nucleotide concentration due to the apparent stickiness of some of the dye analogs (specifically Cy5-dATP) to the chip surface. Since we know that there is limited telomerase activity at 100nM nucleotide concentrations based on our bulk assays, these experimental conditions are not ideal. However, even at these nucleotide concentrations, we have seen intriguing traces (**Fig. 14**). While these traces are telomerase dependent, they are hard to interpret due to the lack of any identifiable telomeric sequence in the data. Additionally, these traces display interesting nucleotide dynamics that seem to indicate that if these traces represent nucleotide bound in the telomerase active site, then there is little to no time when the active site is unbound by nucleotide. The traces seem to indicate that as soon as one nucleotide has been incorporated, another one takes its place in the active site. However, at this

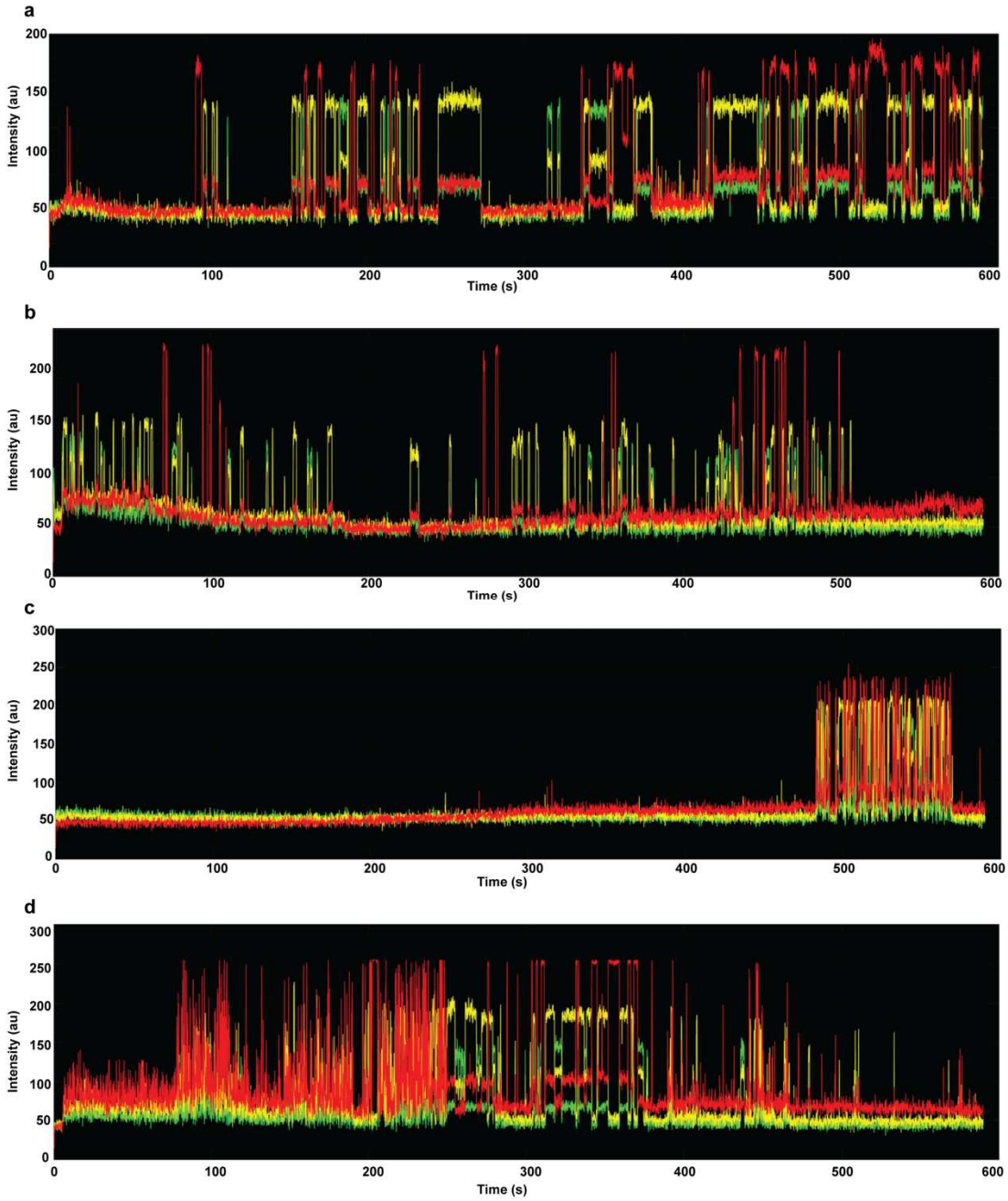

a
b
c
d
Intensity (au)
Time (s)
0
50
100
150
200
250
300
400
500
600

Figure 14. Telomerase ZMW traces. Representative traces generated during initial ZMW experiments. Traces were categorized into four different categorizes as discussed below. Cy5-dATP is represented by red signal, Cy3.5-dGTP is represented by yellow and Cy3-dTTP by green. Intensity is shown on the y-axis with time on the x-axis. (**a**) A representative trace of "slow" behavior. These traces exhibited mutually exclusive signal and displayed very little dark time as well as no observable pattern in terms of sequence. (**b**) A representative trace of "disperse" behavior. These traces exhibited sporadic spikes in intensity and seemed to have no pattern in terms of sequence. **(c**) A representative trace of "fast" behavior. These traces exhibited mutually exclusive signal and displayed very little dark time between peaks but displayed much faster switching between nucleotide signal compared to "slow" traces. Again, these traces show no observable pattern in terms of sequence. **(d**) A representative trace of "mixed" behavior. These traces exhibited a combination of the above three mentioned behaviors.

stage we are not able to distinguish between transient binding in the active site from actual incorporation. To get around this problem we used a ddGTP terminating telomere primer which should only allow for transient binding of nucleotides in the active site. However, with telomerase immobilized with this primer, we saw identical traces as observed with our regularly used primer. These results indicate that the pattern we have observed, while being enzyme dependent, is not activity dependent. An alternative hypothesis is that the traces we observe might be due to FRET between nucleotides bound to more than one location on the enzyme. Additionally, due to our immobilization strategy, there may be multiple telomerase enzymes per well which may complicate the interpretation of our results. At present, we apply excess telomere primer to the bottom of the wells to maximize the amount of enzyme bound per chip, however this might not be the ideal strategy considering our initial results.

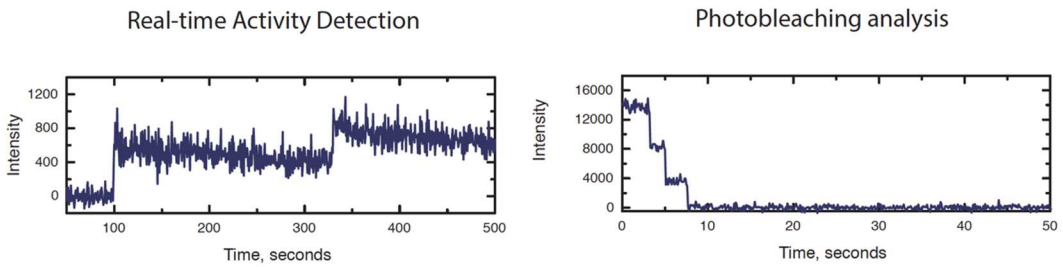


Figure 15. Real-time detection of telomerase activity. Representative traces for experiments performed with a Cy5-detection oligo (CCCATTCCCATTCCC) is shown. Cy5 intensity is represented by blue lines. Telomerase was immobilized on a pegylated quartz slide and data was collected after the addition of dNTPs and detection oligo. Left panel shows the accumulation of Cy5 intensity in discrete steps during activity, indicating binding of successive Cy5-oligos. The right panel shows data under high red laser intensity to promote photobleaching in order to count numbers of Cy5-oligos bound/molecule.

Future ZMW experiments

To get around the problem of both nucleotide-analog stickiness as well as the probability of too many telomerase enzymes being immobilized per well, we have attempted to switch from primer-based immobilization to immobilization using a HaloTag (23). We express recombinant HaloTag-hTERT in RRL (if using dye-labeled RNA) or in 293T cells. A major advantage of using the HaloTag is that it can itself be labeled with biotin, negating the use for primer immobilization. This approach will ensure that enzyme is immobilized to the surface directly, and additionally allows for pre-binding of the telomere primer and enzyme before immobilization, resulting in immobilization of active complexes. Additionally, this immobilization strategy will allow for additional measures of telomerase activity. We have designed experiments using TIRF microscopy where we measure telomerase

activity by using a dye-labeled complementary oligo to 2 and a half telomere repeats (5' CCCTAACCCTAACCC 3'). When telomerase has extended the telomere primer, the complementary oligo can bind, and the more primer extension that has occurred, the more dye-labeled oligo can bind. By measuring the increase in dye intensity combined with counting photobleaching steps to calculate the number of dye-oligos that have bound to the extended telomere, we can estimate the amount of telomerase activity for individual molecules. We have successfully used this strategy using our in house TIRF microscope setup (**Fig. 15**). The next step will be to use this measure of activity together with our ZMW assays. We would not be able to do this with our old immobilization strategy due to the ability of the complementary oligo to bind to free primer immobilized on the surface. However, by immobilizing HaloTag-telomerase, the only primer available to anneal to will be the enzyme-bound primer. Using a combination of this activity assay as well as the addition of PacBio nucleotide analogs, we will be able to correlate traces with active enzyme. We hope that this improvement in our assays will yield insight into the previously unobservable nucleotide incorporation kinetics of telomerase.

References

1. Steczkiewicz K*, et al.* (2011) 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(23):9443-9448.
2. Akiyama BM, Parks JW, & Stone MD (2015) The telomerase essential N-terminal domain promotes DNA synthesis by stabilizing short RNA-DNA hybrids. *Nucleic Acids Res* 43(11):5537-5549.
3. Berman AJ, Akiyama BM, Stone MD, & Cech TR (2011) The RNA accordion model for template positioning by telomerase RNA during telomeric DNA synthesis. *Nature structural & molecular biology* 18(12):1371-1375.
4. Parks JW & Stone MD (2014) Coordinated DNA dynamics during the human telomerase catalytic cycle. *Nat Commun* 5:4146.
5. Shastry S, Steinberg-Neifach O, Lue N, & Stone MD (2018) Direct observation of nucleic acid binding dynamics by the telomerase essential N-terminal domain. *Nucleic Acids Res*.
6. Wu JY, Stone MD, & Zhuang X (2010) A single-molecule assay for telomerase structure-function analysis. *Nucleic Acids Res* 38(3):e16.
7. Hengesbach M, Kim NK, Feigon J, & Stone MD (2012) Single-molecule FRET reveals the folding dynamics of the human telomerase RNA pseudoknot domain. *Angew Chem Int Ed Engl* 51(24):5876-5879.

8. Parks JW, Kappel K, Das R, & Stone MD (2017) Single-molecule FRET-Rosetta reveals RNA structural rearrangements during human telomerase catalysis. *RNA* 23(2):175-188.

9. Wong LS, Thirlway J, & Micklefield J (2008) Direct site-selective covalent protein immobilization catalyzed by a phosphopantetheinyl transferase. *J Am Chem Soc* 130(37):12456-12464.

10. Chen JL & Greider CW (2003) Template boundary definition in mammalian telomerase. *Genes Dev* 17(22):2747-2752.

11. Yang W & Lee YS (2015) A DNA-hairpin model for repeat-addition processivity in telomere synthesis. *Nat Struct Mol Biol* 22(11):844-847.

12. Jacobs SA, Podell ER, & Cech TR (2006) Crystal structure of the essential N-terminal domain of telomerase reverse transcriptase. *Nat Struct Mol Biol* 13(3):218-225.

13. Robart AR & Collins K (2011) Human telomerase domain interactions capture DNA for TEN domain-dependent processive elongation. *Mol Cell* 42(3):308-318.

14. Yin J*, et al.* (2005) Genetically encoded short peptide tag for versatile protein labeling by Sfp phosphopantetheinyl transferase. *Proc Natl Acad Sci U S A* 102(44):15815-15820.

15. Yin J, Lin AJ, Golan DE, & Walsh CT (2006) Site-specific protein labeling by Sfp phosphopantetheinyl transferase. *Nat Protoc* 1(1):280-285.

16. Zhou Z*, et al.* (2007) Genetically encoded short peptide tags for orthogonal protein labeling by Sfp and AcpS phosphopantetheinyl transferases. *ACS Chem Biol* 2(5):337-346.

17. Cohen SB & Reddel RR (2008) A sensitive direct human telomerase activity assay. *Nat Methods* 5(4):355-360.

18. Wallweber G, Gryaznov S, Pongracz K, & Pruzan R (2003) Interaction of human telomerase with its primer substrate. *Biochemistry* 42(2):589-600.

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

20. Lei M, Zaug AJ, Podell ER, & Cech TR (2005) Switching human telomerase on and off with hPOT1 protein in vitro. *J Biol Chem* 280(21):20449-20456.

21. Wang F*, et al.* (2007) The POT1-TPP1 telomere complex is a telomerase processivity factor. *Nature* 445(7127):506-510.

22. Nguyen THD*, et al.* (2018) Cryo-EM structure of substrate-bound human telomerase holoenzyme. *Nature* 557(7704):190-195.

23. Akiyama BM, Gomez A, & Stone MD (2013) 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(30):22141-22149.