UNIVERSITY OF CALIFORNIA

SANTA CRUZ

STRUCTURAL STUDIES OF *TETRAHYMENA THERMOPHILA* TELOMERASE

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

MASTER OF SCIENCE

in

CHEMISTRY

by

John A. Loper

June 2013

The Thesis of John A. Loper is approved:

_____

Assistant Professor Michael D. Stone, Chair

_____

Professor Harry F. Noller

_____

Associate Professor Seth M. Rubin

_____

Tyrus Miller
Vice Provost and Dean of Graduate Studies

Table of Contents

Table of Figures

ABSTRACT

STRUCTURAL STUDIES OF *TETRAHYMENA THERMOPHILA* TELOMERASE

by

JOHN A. LOPER

Eukaryotic organisms contain linear chromosomes that are protected by telomeres; long repeated sequences that act as a protecting buffer from the shortening of chromosomes during each replication cycle, and as a signal that identify the ends from double stranded breaks. Telomeres are maintained in most organisms by a ribonucleoprotein called telomerase. While telomeres and telomerase have been implicated in many human diseases, including cancer, very little is known about the unique enzymatic mechanism of telomerase, making it difficult to develop medicinal therapeutics.

The protein p65 is required for telomerase assembly in *Tetrahymena thermophila*. Using RNase footprinting and single molecule FRET, we show that the C-domain of p65 is not only required, but is sufficient for telomerase assembly, paving the way for understanding possible p65 homologues in other organisms. Preliminary crystals of complexes consisting of the RNA binding domain of the

catalytic subunit and the RNA component of *T. thermophila* were also obtained. Because the catalytic subunit is quite conserved between species, any structure obtained from these crystals will prove valuable in understanding the RNA and protein organization of telomerase.

For a more ambitious goal, a novel method was pursued to solve the structure of telomerase: inserting telomerase RNA into ribosomal RNA, assembling the telomerase complex on the ribosome, using the well known ribosome purification methods and crystallization conditions, and thus using the ribosome as a crystallization chaperone and scaffold. While crystals failed to diffract to a desired resolution, considerable strides were made toward this method that could also be extended to other ribonucleoproteins.

Acknowledgements

My sincere gratitude to everyone that made this thesis possible:

Michael Stone, thesis advisor, for teaching, helping, understanding, and believing in me. I owe you more than I can ever return.

Harry Noller and Seth Rubin, thesis reading committee members, for your time and wisdom in the ways of ribosomes and proteins.

Ben Akiyama, for contributing towards the p65 C-domain studies, specifically the gel shifts, footprinting assays, and TER constructs.

Michael Pearson, for introducing me to the novel ribosome crystallization method. I would have been lost in a sea of failed PCR primers without your guidance.

And most of all, my parents.

Chapter 1

Introduction

1.0 Telomeres

The ends of chromosomes in eukaryotic organisms consist of telomeres; G/T rich 5' to 3' nucleotide repeats that are organism sequence specific, with lengths varying from 350 bp to 50 kb. While mostly double-stranded, telomeres contain a single-stranded G-rich 3' tail that acts as a primer for a unique telomere extending enzyme, telomerase. Telomeres also form g-quadruplexes, T-loops, and recruit a multitude of proteins to form a protective cap (Griffith et al., 1999; Sundquist and Klug, 1989; Williamson et al., 1989).

Linear chromosomes lead to issues with end-protection and end-replication. The ends of chromosomes have the danger of being recognized as double-strand breaks by repair machinery, and may result in end-to-end chromosome fusions through recombination events. These ends are also vulnerable to exonuclease activity. The end-replication problem was first proposed by Olovnikov and Watson in the early 1970s (Olovnikov, 1971; 1973; Watson, 1972). DNA polymerase synthesizes DNA in the 5' to 3' direction and is unable to fill the gap left behind the 5' most Okazaki fragment RNA primer on the lagging strand (Levy et al., 1992). This leads to progressively shortening ends with each replication cycle, limiting the lifespan of the cells.

1.1 Telomerase

Telomerase assists in solving these issues due to linear chromosomes, and was first discovered by Greider and Blackburn in the mid 1980s (Greider and Blackburn, 1985). There are two key elements of telomerase: the catalytic subunit telomerase reverse transcriptase (TERT), and the RNA component (TER) that provides the template for reverse transcription along with other important elements (Greider and Blackburn, 1989). This unique ribonucleoprotein synthesizes tandem repeats found in telomeres by using the template in TER and single stranded telomeric DNA as the primer, compensating for the loss due to the end-replication issue (Blackburn et al., 2006). This repeated sequence recruits the necessary proteins to form the protective cap, distinguishing the ends from double stranded breaks and providing protection from exoucleases (Stewart et al., 2012). The architecture and composition of this cap is also quite varied between organisms. While most organisms use telomerase for telomere maintenance, some have been found to use recombination or transposition mechanisms (Pardue and DeBaryshe, 2008).

Along with the unique property of providing the template for reverse transcription, telomerase exhibits repeat addition processivity (RAP) that was first observed in *Tetrahymena thermophila* and humans (Fig. 2) (Greider, 1991; Morin, 1989). Mutations in TERT can alter RAP *in vitro*, and while variance has

been observed between telomerase RAP *in vivo* and *in vitro*, it has been suggested that RAP depends on telomere length, enabling cells to rapidly elongate critically short telomeres (Chang et al., 2007; Teixeira et al., 2004). Defects in RAP have been implicated with several diseases. The processivity and fidelity of RAP also varies between organisms and growth conditions (Collins, 2009).

1.1.1 Telomerase RNA (TER)

While TER varies greatly in sequence and size between organisms, from about 150 to 1500 nucleotides, it shares similar elements other than the template (Fig. 1). A template boundary element (TBE) defines the limit of reverse transcription and prevents read through that would add additional unwanted nucleotides to each repeat (Box et al., 2008; Chen, 2003; Miller, 2000; Tzfati, 2000). A pseudoknot that is stabilized by a triple-helix is suspected to have a role in orienting the primer-template duplex in the active site of TERT (Gilley, 1999; Qiao and Cech, 2008; Shefer et al., 2007; Theimer et al., 2005). There is also a stem terminal element (STE) that stimulates activity through interaction with TERT (Mitchell and Collins, 2000).

1.1.2 Telomerase Reverse Transcriptase (TERT)

The catalytic reverse transcriptase subunit TERT specifically binds to TER and contains multiple ssDNA telomeric anchor sites (Lingner, 1997; Meyerson et al., 1997). Around 1000 amino acids and, unlike TER, mostly conserved from humans to *T. thermophila*, TERT consists of four major domains: essential N-terminal domain (TEN), RNA binding domain (RBD), reverse transcriptase domain (RT), and the C-terminal extension (CTE). TERT was shown to be the catalytic subunit in *T. thermophila* by reconstituting TER and TERT *in vitro* using rabbit reticulocyte lysates (Collins and Gandhi, 1998).

A crystal structure of the flour beetle *Tribolium castaneum* TERT bound to a RNA-DNA hairpin, mimicking the primer-template duplex, gave insight into the molecular contacts near the active site (Mitchell et al., 2010). However, given the complex telomerase and telomere maintenance in insects, no formal proof exists that this structure represents the catalytic TERT subunit, but still provides the basis for structural comparisons of TERTs in other organisms (Osanai et al., 2006; Sasaki and Fujiwara, 2003). A ring-like structure found in the beetle TERT contains a thumb, palm, and fingers motifs, similar to other nucleic acid polymerases and the human immunodeficiency virus (HIV) reverse transcriptase. The active site contains three aspartic acids that form a catalytic

triad for nucleotide addition via a two-metal ion mechanism (Lingner, 1997; Nakamura, 1997).

The TEN domain of TERT acts as an anchor site for ssDNA binding upstream of the primer-template duplex and like RBD, directly interacts with TER (Jacobs et al., 2006; O'Connor, 2005; Robart and Collins, 2011). Along with the crystal structure of *T. thermophila* TEN being solved by Jacobs et al., photocrosslinking by Robart showed Trp187 on the surface of TEN is involved in DNA binding. TEN has been implicated as a key domain for RAP, likely by capturing and maintaining telomerase association with ssDNA during RAP (Collins, 2011; Zaug et al., 2008).

1.2 Telomerase and Human Disease

Telomerase is active in germ lines and stem cells, inactive in somatic cells, and improper activation could lead to cancer; at least 85% of human cancer cells express TERT to achieve cellular immortalization, making telomerase an attractive target for anti-cancer therapeutics (Kim et al., 1994; Phatak and Burger, 2009). The first telomerase related disease identified was dyskeratosis congenita (Collins et al., 1999). Symptoms first appear during childhood and include developmental delay, premature hair loss, and organ failure, with bone

marrow failure as the main cause of mortality. Other diseases include aplastic anemia, Hoyeraal-Hreidarsson syndrome, idiopathic pulmonary fibrosis, and liver disease, all of which are related to mutations in the telomerase holoenzyme and significantly shortened telomeres (Armanios, 2009). Additionally, telomeres can reach a critically short length that leads to cell senescence and has been implicated in aging (Jaskelioff et al., 2010).

1.3 *Tetrahymena thermophila* Telomerase

T. thermophila possesses the smallest known TER, 159 nucleotides, and contains all the conserved motifs found between organisms, making it an ideal target for characterizing the catalytic mechanism of telomerase (Greider and Blackburn, 1989). Assembly of the telomerase holoenzyme begins with TER transcription by RNA polymerase III, providing the 3' polyuridine tail that promotes binding of p65, a protein that is required for telomerase accumulation *in vivo* (O'Connor and Collins, 2006; Witkin, 2004). P65 bends stem IV to promote TERT binding and form the "catalytic core", the minimal enzyme components for in vitro catalysis, and also induces conformational changes in stem loop IV that promotes catalytic activity (Fig. 1) (Akiyama et al., 2012; Singh et al., 2012; Stone et al., 2007).

Other associated proteins – p19, p45, p50, p75, and Teb1 – are required for telomere maintenance *in vivo* (Min and Collins, 2009; Witkin, 2004; Witkin et al., 2007). Addition of these proteins to the catalytic core greatly enhances RAP (Fig. 1, Fig. 2). Using cryo-EM, most of the holoenzyme was solved at low resolution and surprisingly demonstrated that p50 was the hub protein that bridged p19, p45, p75, and Teb1 to the catalytic core, suggesting that p50 is a critical component *in vivo* (Jiang et al., 2013).

While there have been many strides in the telomerase field, the exact enzymatic mechanism of telomerase is largely unknown. How does telomerase perform the unique enzymatic activity of repeat addition processivity, and how relevant is this property in organisms? While biochemical studies may help elucidate this question, high-resolution structures are needed of the components of telomerase and of the holoenzyme itself. By studying the relatively simple *T. thermophila* telomerase activity, assembly, and structure, it may yield information that translates to other organisms, specifically human, and assist in the development for telomerase and telomere therapeutics.

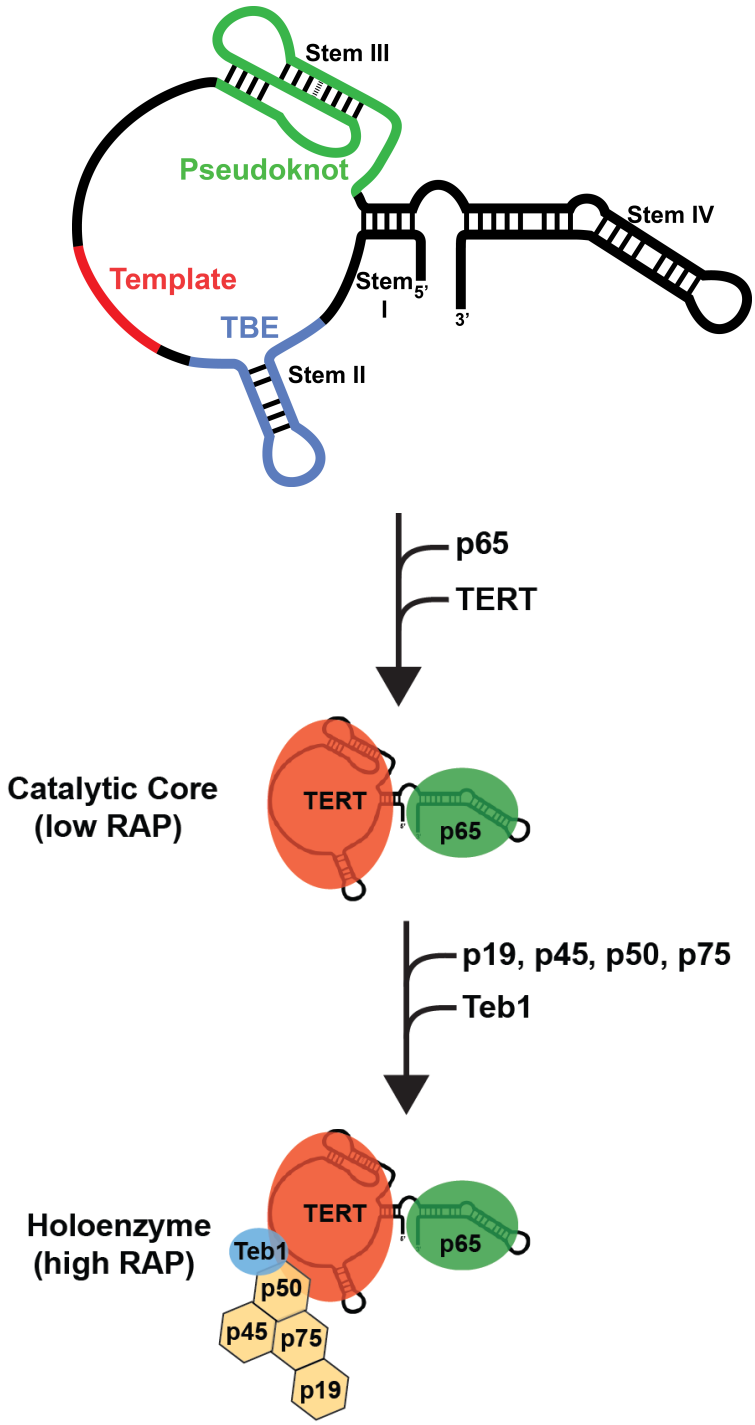


Figure 1. *Tetrahymena thermophila* telomerase RNA secondary structure and holoenzyme assembly.

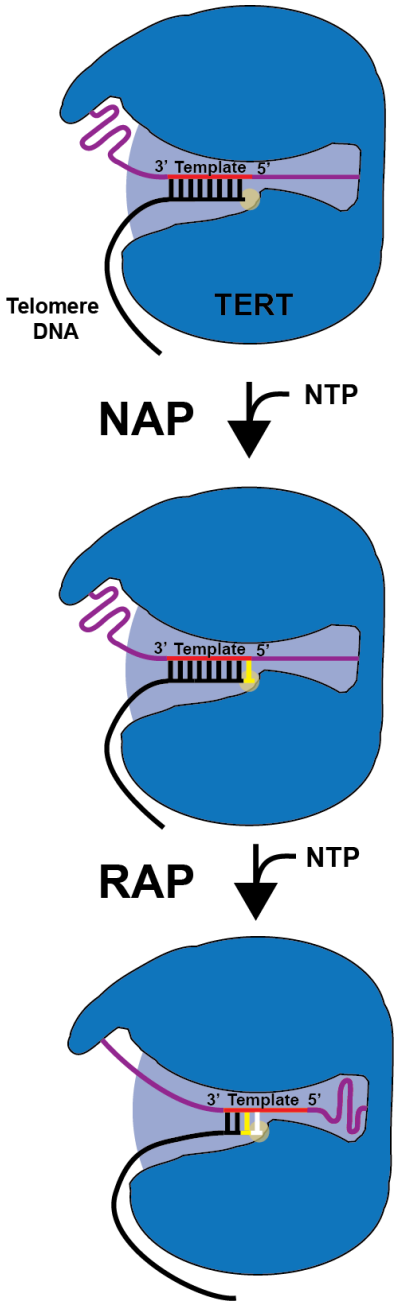


Figure 2. Schematic of nucleotide addition processivity (NAP) and repeat addition processivity (RAP), with TER RNA in purple, RNA template in red, telomere DNA in black, active site as a yellow circle, last nucleotide added in the template repeat in yellow, and first nucleotide added in the template repeat in white.

1.4 References

Akiyama, B.M., Loper, J., Najarro, K., and Stone, M.D. (2012). The C-terminal domain of Tetrahymena thermophila telomerase holoenzyme protein p65 induces multiple structural changes in telomerase RNA. Rna *18*, 653–660.

Armanios, M. (2009). Syndromes of Telomere Shortening. Annual Review of Genomics and Human Genetics *10*, 45–61.

Blackburn, E.H., Greider, C.W., and Szostak, J.W. (2006). Telomeres and telomerase: the path from maize, Tetrahymena and yeast to human cancer and aging. Nat Med *12*, 1133–1138.

Box, J.A., Bunch, J.T., Zappulla, D.C., Glynn, E.F., and Baumann, P. (2008). A flexible template boundary element in the RNA subunit of fission yeast telomerase. Journal of Biological Chemistry *283*, 24224–24233.

Chang, M., Arneric, M., and Lingner, J. (2007). Telomerase repeat addition processivity is increased at critically short telomeres in a Tel1-dependent manner in Saccharomyces cerevisiae. Genes & Development *21*, 2485–2494.

Chen, J.L. (2003). Template boundary definition in mammalian telomerase. Genes & Development *17*, 2747–2752.

Collins, K. (2009). Forms and Functions of Telomerase RNA. In Springer Series in Biophysics, (Springer Berlin Heidelberg), pp. 285–301.

Collins, K. (2011). Single-stranded DNA repeat synthesis by telomerase. Curr Opin Chem Biol *15*, 643–648.

Collins, K., and Gandhi, L. (1998). The reverse transcriptase component of the Tetrahymena telomerase ribonucleoprotein complex. Proc. Natl. Acad. Sci. *15*, 8485–8490.

Collins, K., Mitchell, J.R., and Wood, E. (1999). A telomerase component is defective in the human disease dyskeratosis congenita. Nature *402*, 551–555.

Gilley, D. (1999). The telomerase RNA pseudoknot is critical for the stable assembly of a catalytically active ribonucleoprotein. Proc. Natl. Acad. Sci. *96*, 6621–6625.

Greider, C.W. (1991). Telomerase is processive. Mol. Cell. Biol. *11*, 4572–4580.

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

Greider, C.W., and Blackburn, E.H. (1985). Identification of a specific telomere terminal transferase activity in tetrahymena extracts. Cell *43*, 405–413.

Griffith, J.D., Comeau, L., Rosenfield, S., and Stansel, R.M. (1999). Mammalian telomeres end in a large duplex loop. Cell *97*, 503–514.

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

Jaskelioff, M., Muller, F.L., Paik, J.-H., Thomas, E., Jiang, S., Adams, A.C., Sahin, E., Kost-Alimova, M., Protopopov, A., Cadiñanos, J., et al. (2010). Telomerase reactivation reverses tissue degeneration in aged telomerase-deficient mice. Nature *469*, 102–106.

Jiang, J., Miracco, E.J., Hong, K., Eckert, B., Chan, H., Cash, D.D., Min, B., Zhou, Z.H., Collins, K., and Feigon, J. (2013). The architecture of Tetrahymena telomerase holoenzyme. Nature *496*, 187–192.

Kim, N., Piatyszek, M., Prowse, K., Harley, C., West, M., Ho, P., Coviello, G., Wright, W., Weinrich, S., and Shay, J. (1994). Specific association of human telomerase activity with immortal cells and cancer. Science *266*, 2011–2015.

Levy, M.Z., Allsopp, R.C., Futcher, A.B., Greider, C.W., and Harley, C.B. (1992). Telomere end-replication problem and cell aging. J. Mol. Biol. *225*, 951–960.

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

Meyerson, M., Counter, C.M., Eaton, E.N., Ellisen, L.W., Steiner, P., Caddle, S.D., Ziaugra, L., Beijersbergen, R.L., Davidoff, M.J., Liu, Q., et al. (1997). hEST2, the Putative Human Telomerase Catalytic Subunit Gene, Is Up-Regulated in Tumor Cells and during Immortalization. Cell *90*, 785–795.

Miller, M.C. (2000). Template definition by Tetrahymena telomerase reverse transcriptase. The EMBO Journal *19*, 4412–4422.

Min, B., and Collins, K. (2009). An RPA-related sequence-specific DNA-binding subunit of telomerase holoenzyme is required for elongation processivity and telomere maintenance. Mol. Cell *36*, 609–619.

Mitchell, J.R., and Collins, K. (2000). Human Telomerase Activation Requires Two Independent Interactions between Telomerase RNA and Telomerase Reverse Transcriptase. Mol. Cell *6*, 361–371.

Mitchell, M., Gillis, A., Futahashi, M., Fujiwara, H., and Skordalakes, E. (2010). Structural basis for telomerase catalytic subunit TERT binding to RNA template and telomeric DNA. Nat. Struct. Mol. Biol. *17*, 513–518.

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

Nakamura, T.M. (1997). Telomerase Catalytic Subunit Homologs from Fission Yeast and Human. Science *277*, 955–959.

O'Connor, C.M. (2005). Two Purified Domains of Telomerase Reverse Transcriptase Reconstitute Sequence-specific Interactions with RNA. J. Biol. Chem. *280*, 17533–17539.

O'Connor, C.M., and Collins, K. (2006). A Novel RNA Binding Domain in Tetrahymena Telomerase p65 Initiates Hierarchical Assembly of Telomerase Holoenzyme. Mol. Cell. Biol. *26*, 2029–2036.

Olovnikov, A.M. (1971). Principle of marginotomy in template synthesis of polynucleotides. Dokl Akad Nauk SSSR *201*, 1496–1499–1499.

Olovnikov, A.M. (1973). A theory of marginotomy. Journal of Theoretical Biology *41*, 181–190.

Osanai, M., Kojima, K.K., Futahashi, R., Yaguchi, S., and Fujiwara, H. (2006). Identification and characterization of the telomerase reverse transcriptase of Bombyx mori (silkworm) and Tribolium castaneum (flour beetle). Gene *376*, 281–289.

Pardue, M.-L., and DeBaryshe, P.G. (2008). Drosophila Telomeres: A Variation on the Telomerase Theme. Fly *2*, 101–110.

Phatak, P., and Burger, A.M. (2009). Telomerase and its potential for therapeutic intervention. British Journal of Pharmacology *152*, 1003–1011.

Qiao, F., and Cech, T.R. (2008). Triple-helix structure in telomerase RNA contributes to catalysis. Nat. Struct. Mol. Biol. *15*, 634–640.

Robart, A.R., and Collins, K. (2011). Human Telomerase Domain Interactions Capture DNA for TEN Domain-Dependent Processive Elongation. Mol. Cell *42*, 308–318.

Sasaki, T., and Fujiwara, H. (2003). Detection and distribution patterns of telomerase activity in insects. European Journal of Biochemistry *267*, 3025–3031.

Shefer, K., Brown, Y., Gorkovoy, V., Nussbaum, T., Ulyanov, N.B., and Tzfati, Y. (2007). A Triple Helix within a Pseudoknot Is a Conserved and Essential Element of Telomerase RNA. Mol. Cell. Biol. *27*, 2130–2143.

Singh, M., Wang, Z., Koo, B.-K., Patel, A., Cascio, D., Collins, K., and Feigon, J. (2012). Structural Basis for Telomerase RNA Recognition and RNP Assembly by the Holoenzyme La Family Protein p65. Mol. Cell *47*, 16–26.

Stewart, J.A., Chaiken, M.F., Wang, F., and Price, C.M. (2012). Maintaining the end: Roles of telomere proteins in end-protection, telomere replication and length regulation. Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis *730*, 12–19.

Stone, M.D., Mihalusova, M., O'Connor, C.M., Prathapam, R., Collins, K., and Zhuang, X. (2007). Stepwise protein-mediated RNA folding directs assembly of telomerase ribonucleoprotein. Nature *446*, 458–461.

Sundquist, W.I., and Klug, A. (1989). Telomeric DNA dimerizes by formation of guanine tetrads between hairpin loops. Nature *342*, 825–829.

Teixeira, M.T., Arneric, M., Sperisen, P., and Lingner, J. (2004). Telomere Length Homeostasis Is Achieved via a Switch between Telomerase- Extendible and -Nonextendible States. Cell *117*, 323–335.

Theimer, C.A., Blois, C.A., and Feigon, J. (2005). Structure of the Human Telomerase RNA Pseudoknot Reveals Conserved Tertiary Interactions Essential for Function. Mol. Cell *17*, 671–682.

Tzfati, Y. (2000). Template Boundary in a Yeast Telomerase Specified by RNA Structure. Science *288*, 863–867.

Watson, J.D. (1972). Origin of Concatemeric T7DNA. Nature *239*, 197–201.

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

Witkin, K.L. (2004). Holoenzyme proteins required for the physiological assembly and activity of telomerase. Genes & Development *18*, 1107–1118.

Witkin, K.L., Prathapam, R., and Collins, K. (2007). Positive and Negative Regulation of Tetrahymena Telomerase Holoenzyme. Mol. Cell. Biol. *27*, 2074–2083.

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

Chapter 2

The C-terminal Domain of p65 is Required and Sufficient for TER Rearrangement and TERT Recruitment

2.0 Introduction

2.0.1 p65

Ciliates require La domain proteins for assembly of telomerase *in vivo* (Witkin, 2004). Often found in conjunction with RNA recognition motifs (RRM) within N-terminal domains, La motifs bind 3' poly-U tracts of mRNAs produced by RNA polymerase III to assist in their folding and maturation (Bayfield et al., 2010).

In *Tetrahymena thermophila*, La motif protein p65 is required for telomere maintenance in-vivo, increases the binding affinity of the reverse transcriptase component of telomerase (TERT) with telomerase RNA (TER) in-vitro, and directly interacts with stem I and IV of TER (O'Connor and Collins, 2006; Prathapam et al., 2005; Witkin, 2004). Of its four domains, the N and C-terminal domains have little sequence homology with known protein motifs, while the La and RRM domains have been well characterized and likely bind TER in the vicinity of the 3' poly-U tract. P65 is also able to rescue TER mutants and TERT assembly mutants, but not TERT activity mutants (Berman et al., 2010).

A single molecule FRET study demonstrated that p65 is responsible for bending stem IV, allowing the recruitment of TERT, and confirmed the

importance of the evolutionary conserved GA bulge within stem IV, which has also been shown to be flexible by NMR spectroscopy (Richards, 2006; Stone et al., 2007). This raises the question of the exact mechanism of how p65 remodels TER and which domains are directly involved.

2.0.2 FRET

Many researchers have implemented the use of Förster resonance energy transfer (FRET) to study the motions and interactions of biomolecules in real time. It consists of attaching donor and acceptor fluorescent dyes to the molecule or molecules of interest and measuring efficiency of energy transfer between dyes. This energy transfer occurs due to overlapping excitation and absorption spectra of the dyes and a dipole-dipole interaction that can be observed by the decreasing fluorescence of the donor dye and increasing fluorescence of the acceptor dye.

The transfer efficiency is calculated using $E = 1/(1 + (R/R_0)^6)$, with E as transfer efficiency, R as the distance between dyes, and R_0 as the Förster radius. The Forster radius is the distance between dyes when the transfer of energy is 50% efficient, typically 30-80 Å. Transfer efficiency is affected by the orientation of the dyes to one another, which can be alleviated by using flexible dye linkers,

but still adds uncertainty to measured values (cite?). Because of this uncertainty, converting measured FRET values directly to distance can be misleading.

Single molecule FRET (smFRET) allows measurements of single molecules in real time (Ha et al., 1996). This opens the possibility of observing short-lived states and heterogeneity of the molecule in question. Histograms can also be easily built from many individual molecules that represent the distribution of FRET states. A common smFRET strategy is to immobilize molecules of interest to the surface of a microscope slide and excite by total internal reflection fluorescence (TIRF) so that only molecules attached to the surface fluoresce (Fig. 1). These immobilized molecules are observed using a microscope with a fast CCD camera, and processed with software to track and record the real time FRET states of hundreds of single molecules at once. The efficiency of energy transfer is quantified using I_A / (I_A + I_D), with I_A being the intensity of the acceptor dye and I_D being the intensity of the donor dye. For the purpose of this study, dyes Cy3 and Cy5 were used, acting as donor (green fluorescence) and acceptor (red fluorescence), respectively. For example, molecules will appear green when further away, and red when closer together.

2.1 Results

2.1.1 Role of the C-domain of p65 in Remodeling TER

It has been suggested that the C-domain of p65 increases TERT and TER affinity, rather than the La/RRM domains, so an approach was taken to construct four p65 constructs: full length p65 ($p65^{FL}$), delta-N domain p65 ($p65^{\Delta N}$), delta-C domain p65 ($p65^{\Delta C}$), and C-terminal domain p65 ($p65^{CTD}$) (O'Connor and Collins, 2006). These constructs were first examined for their competency of binding TER by native PAGE gel shift assays (Fig. 2). While all constructs bound TER, $p65^{\Delta N}$ and $p65^{\Delta C}$ had reduced affinity, and $p65^{CTD}$ had greatly reduced affinity (Fig. 3 and 4). This suggests that domains other than the C-terminal domain of p65 have an influence on p65's affinity with TER.

Following the assay used in a previous smFRET p65 study, the shift in FRET of each p65 construct binding to TER was measured (Stone et al., 2007). The donor dye was placed at U139 and the acceptor dye at U10 of the native TER sequence (Fig. 5). Dye labeled TER constructs were immobilized and excited using a prism-based TIRF microscope (Fig. 6). The expected FRET shift of $p65^{FL}$ was observed along with a similar shift for $p65^{\Delta N}$ using a slightly higher protein concentration. No FRET shift was observed with $p65^{\Delta C}$, however a similar shift to

$p65^{FL}$ was observed with high concentrations of $p65^{CTD}$, suggesting that not only is the C-domain required, but also is sufficient for remodeling TER (Fig. 7). A FRET shift due to high protein concentrations was discounted by continuing to see no FRET shift with high concentrations of $p65^{\Delta C}$. While the shift in FRET with $p65^{CTD}$ is slightly different from $p65^{FL}$, it could be due to additional remodeling affects by excluded domains. Interestingly when $p65^{\Delta C}$ is added in-trans to $p65^{CTD}$, the required $p65^{CTD}$ concentration to produce the FRET shift is reduced, suggesting $p65^{\Delta C}$ helps stabilize the RNA to promote RNA remodeling by $p65^{CTD}$. These results show the essential role of the C-terminal domain of p65 in the remodeling of TER for the promotion of TERT.

2.1.2 p65 C-terminal Domain Reorganizes Bases Within Stem-loop IV

To further understand the mechanism of p65 remodeling of TER, RNase protection experiments were implemented to identify specific C-domain interaction sites with TER. Due to the ability of full length TER to bind multiple p65 proteins at the concentrations required for protection experiments, a TER construct consisting only of stems I and IV was designed (Fig. 8). This TER construct allowed higher p65 concentrations before higher-order complexes

were formed. This shorter TER construct also provided sufficient resolution and quantification to be run on a single gel.

Using RNase I, a single-stranded RNA nuclease that cleaves at unpaired nucleotides, cleavage was performed on TER with no p65 present. Cleavage products were found at known unpaired nucleotides. While there was cleavage on the top strand of stem IV, no cleavage was observed on the bottom strand. Due to increased cleavage occurring at p65 concentrations that do not appear to form a complex in gel shift assays, a p65 concentration that showed very little to no binding was used as a baseline.

The protection data using increasing concentrations of $p65^{FL}$ shows binding across the GA bulge to the 4-nucleotide linker, and though no cleavage was detected at the poly-U tract, the same result occurs with no protein (Fig. 9). Surprisingly, hypersensitivity to cleavage was found in stem-loop IV. This result is in direct conflict with the results of a recent paper that showed protection in stem-loop IV using full-length TER (Berman et al., 2010). The experiment was repeated using full-length TER and hypersensitivity in stem-loop IV was still observed. This difference is unclear, but could be due to the slightly different binding and cleavage conditions, along with protein production techniques.

$p65^{\Delta N}$ showed a pattern similar to $p65^{FL}$, $p65^{\Delta C}$ increased sensitivity in the top strand of stem IV and reduced sensitivity to stem-loop IV, and $p65^{CTD}$ was also similar to $p65^{FL}$, but with reduced sensitivity of cleavage in stem-loop

IV (Fig. 10). These results confirm the importance of the C-domain of p65 in remodeling stem IV and also reveal that the remodeling exposes the nucleotides of stem-loop IV.

2.2 Discussion

The C-terminal domain of p65 is sufficient and essential for the remodeling of TER to recruit TERT. SmFRET results show the C-domain alone is sufficient to induce a similar TER conformational change as $p65^{FL}$. This can be enhanced with in-trans addition of $p65^{\Delta C}$, while $p65^{\Delta C}$ alone fails to induce any conformational change, suggesting a model that the C-domain is responsible for TER remodeling and the La-RRM domains stabilize this interaction (Akiyama et al., 2012).

NMR spectroscopy structures of stem-loop IV have shown that the nucleotides are highly ordered when not in the presence of p65 (Chen et al., 2006; Richards, 2006). Residues C132 and U138 form a noncanonical base pair and C132, A133, and C134 have a base stacking interaction. These residues have the greatest deprotection and hypersensitivity to RNase ONE cleavage, suggesting that disrupting the interactions of these residues is key to remodeling TER for TERT binding. A result from a recent study revealed that there are

mutations in stem-loop IV that do not affect the binding ability of p65 but prevent the promotion of TERT, further suggesting that disrupting these residue's interactions is key for promoting TERT (Robart et al., 2010).

Our results along with previous studies suggest that the N-domain is not required for the in-vitro activity of p65 (O'Connor and Collins, 2006). The N-domain may be involved in setting up additional contact sites or it may only be required for in-vivo activity, perhaps providing a role other than complex formation.

More recent studies have provided even further insight on the role of p65's C-domain. Along with discovering an atypical RRM within the C-domain, a crystal structure consisting of most of the C-domain in complex with stem IV was solved (Singh et al., 2012). Confirming smFRET results, binding of $p65^{CTD}$ induces a 105° bend in stem IV in a GA bulge dependent manor. This atypical RRM within the C-domain was also found to be structurally homologous to other La protein domains, and may suggest another general mode of RNA binding. An electron microscopy structure of the holoenzyme was also recently solved, demonstrating how p65's C-domain bends stem-loop IV to make direct contact with TERT (Jiang et al., 2013). These two studies further confirm the importance of p65, especially its C-domain, in the hierarchical assembly of *T. thermophila* telomerase.

2.3 Methods

2.3.1 Dye Labeled RNA Construction

Construction of the dye labeled RNA was carried out as previously described (Akiyama and Stone, 2009). Short synthetic RNA fragments of the 3' and 5' ends of TER containing reactive amine groups at U10 and U139 were purchased (Dharmacon). Fragments with U10 or U139 were labeled with Cy5 and Cy3 fluorescent dyes, respectively, and purified by reverse phase HPLC. The dyes contain a reactive N-hydroxysuccinimide-activated carboxylic acid that can be covalently attached to the amine groups on the RNA. Full length TER was transcribed in-vitro and cut to form the central TER piece by using chimeric 2'-O-me RNA/DNA targeting oligos and RNase H. After PAGE purification, the two dye labeled fragments and central unmodified fragment were assembled using DNA-splinted ligation and PAGE purified once again. A DNA biotin-labeled handle was annealed to the purified RNAs to prepare for smFRET.

2.3.2 Protein Expression and Purification

Protein constructs were expressed from a pET28 vector supplied by the Collins lab using BL21 DE3 *Escherichia coli* cells. $p65^{\Delta N}$ consists of amino acids 109-542, $p65^{\Delta C}$ 1-340, and $p65^{CTD}$ 302-542. After reaching mid-log growth phase at 37°C in 2xYT media, expression was performed using 0.8 mM IPTG at 16°C and cells harvested after 4 hours. The fresh or flash frozen cells were resuspended in buffer A (20 mM Tris pH 8.0, 200 mM NaCl, 1 mM $MgCl_2$, 40 mM imidazole, 1 mM DTT, and 10% glycerol). 1 mM PMSF was added just before passing the cells twice through a cell disruptor. The cell lysate was cleared by centrifugation at ~36,000 x g and passed over Ni Sepharose 6 Fast Flow resin (GE Life Sciences) by gravity flow. The resin was thoroughly washed with buffer A that contained 500 mM NaCl until no protein was detected by Bradford assay. After washing with a column volume of buffer A containing 100 mM NaCl, the protein was eluted with buffer A containing 100 mM NaCl and 500 mM imidazole until no protein was detected. The elution was diluted to 100 mL with buffer B (20 mM Tris pH 8.0, 100 mM NaCl, 1 mM $MgCl_2$, 1 mM DTT, and 10% glycerol), passed over a Q anion exchange column (GE Life Sciences), and protein eluted with a 60 minute buffer B gradient to 1 M NaCl. The protein was concentrated with a spin concentrator (Amicon) and passed over a Superdex 200 size-exclusion column (GE Life Sciences) in buffer C (20 mM Tris pH 8.0, 100

mM NaCl, 1 mM $MgCl_2$, 1 mM DTT, and 10% glycerol) 10 mg at a time maximum. The protein was aliquoted, flash frozen, and stored at -70°C.

2.3.3 FRET

Slide preparation was carried out as previously described (Bokinsky et al., 2006; Ha, 2001). Dye labeled RNA with a biotin labeled DNA handle were immobilized to a streptavidin coated PEGylated quartz slide, p65 constructs flowed over, and left in a buffer containing 20 mM Tris pH 8, 100 mM NaCl, 0.1 mg/mL BSA, 0.1 mg/mL yeast tRNA, 2 mM trolox, 10% glucose, 10% glycerol, 1 ug/mL catalase, and 1 mg/mL glucose oxidase. Dyes were excited using a prism-based TIRF microscope and FRET values measured using an Andor IXON CCD camera. Histograms were generated from a 2 second observation time and 100 ms integration time. Molecules with a 0 FRET state were observed due to bleached acceptor dyes and were excluded from the data.

2.3.4 Electrophoretic Mobility Shift Assays

Mobility shift assays were carried out as previously described (O'Connor, 2005). TER was labeled by transcribing with T7 RNA polymerase (NEB) in the presence of α-^{32}P UTP and PAGE purified. In a buffer containing 20 mM Tris, pH 8.0, 100 mM NaCl, 1 mM MgCl2, 0.1 mg/mL BSA, 0.1 mg/mL yeast tRNA, 1 mM DTT, and 10% glycerol, 100 pM of labeled TER was incubated with the appropriate p65 construct/dilution for 20 minutes at 30°C. Samples were loaded with a buffer containing 0.5X TBE and 4% glycerol onto a 5% 37.5:1 acrylaminde:bis-acrylamide gel. The gel was ran in 0.5X TBE at 200V for 3 hours at 4°C. The gel was dried and imaged with a phosphor screen (Amersham) and Typhoon scanner (GE). Imagequant was used for analysis and Origin for calculating K_d using the equation $F= [(Fmax)(c^n)]/[(K_d{}^n)(c^n)]$, with F as fraction bound, c for concentration, and n for the Hill coefficient.

2.3.5 RNase Protection Assays

TER constructs were transcribed using T7 RNA polymerase (NEB), DNase treated (Ambion), phosphatase treated (NEB), and PAGE purified. The RNA was

end-labeled using phosphonucleotide kinase (NEB) with ^{32}P-labeled γ-ATP (Perkin-Elmer) and PAGE purified. In a buffer containing 20 mM Tris pH 8.0, 100 mM NaCl, 1 mM $MgCl_2$, 0.1 mg/mL yeast tRNA, 1 mM DTT, and 10% glycerol, appropriate p65 constructs were incubated at RT for 15 minutes with 5.0 ng/μL labeled TER and 0.00067 U/μL RNase ONE. A recovery control was added, reaction quenched with a phenol/chloroform mixture, and RNA precipitated with ethanol. In a similar fashion, a reference ladder was made using RNase T1 (Ambion).

The RNA was run on a 10% 19:1 acrylamide:bis-acrylamide denaturing sequencing PAGE gel and imaged with a phosphor screen (Amersham) and Typhoon scanner (GE). Semi-automated footprinting analysis (SAFA) was used for quantification (Das 2005 ref). Band intensities were normalized to the amount of material in the recovery control and lanes from the low-protein (LP) control lanes were compared to the high-protein lanes. The extent of protection for protected residues was quantified using 1-(HP/LP). The extent of hypersensitivity for deprotected residues was quantified using (HP/LP)-1.

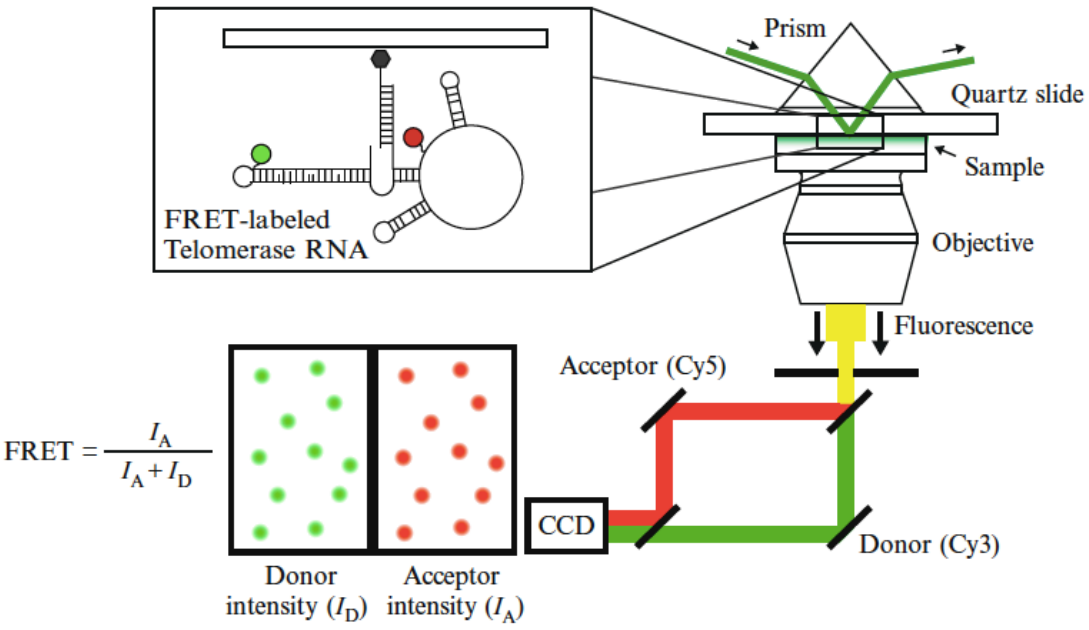
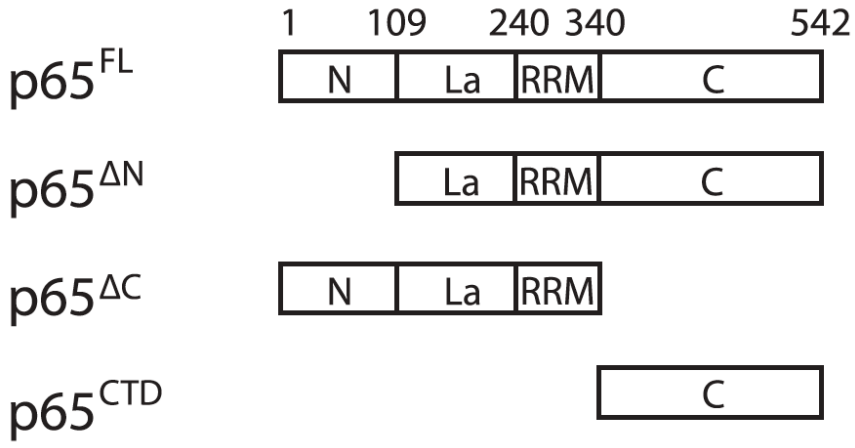


Figure 1. Schematic of the TIRF FRET system used for the study. Note that only the FRET-labeled telomerase RNA attached to the coverslip is excited, drastically reducing background fluorescence. The red and green fluorescent light are separated using dichroic mirrors and captured by the CCD camera side by side.

A



B

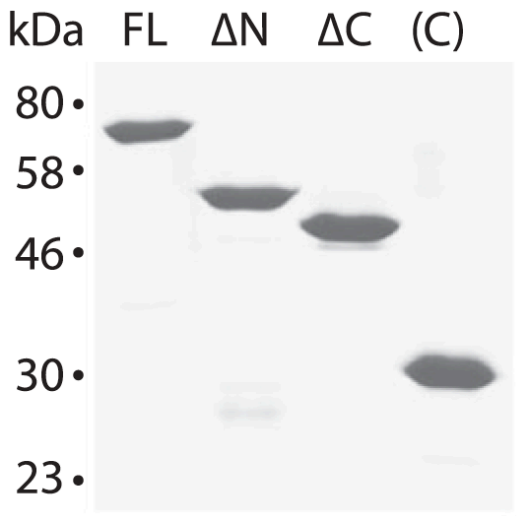


Figure 2. (A) Diagram showing the domain organization of the four p65 constructs used in the study: full length ($p65^{FL}$), lacking the N-terminal domain ($p65^{\Delta N}$), lacking the C-terminal domain ($p65^{\Delta C}$), and the C-terminal domain alone ($p65^{CTD}$). (B) SDS-page gel of the four constructs after expression and purification from BL21 (DE3) *E. coli* cells.

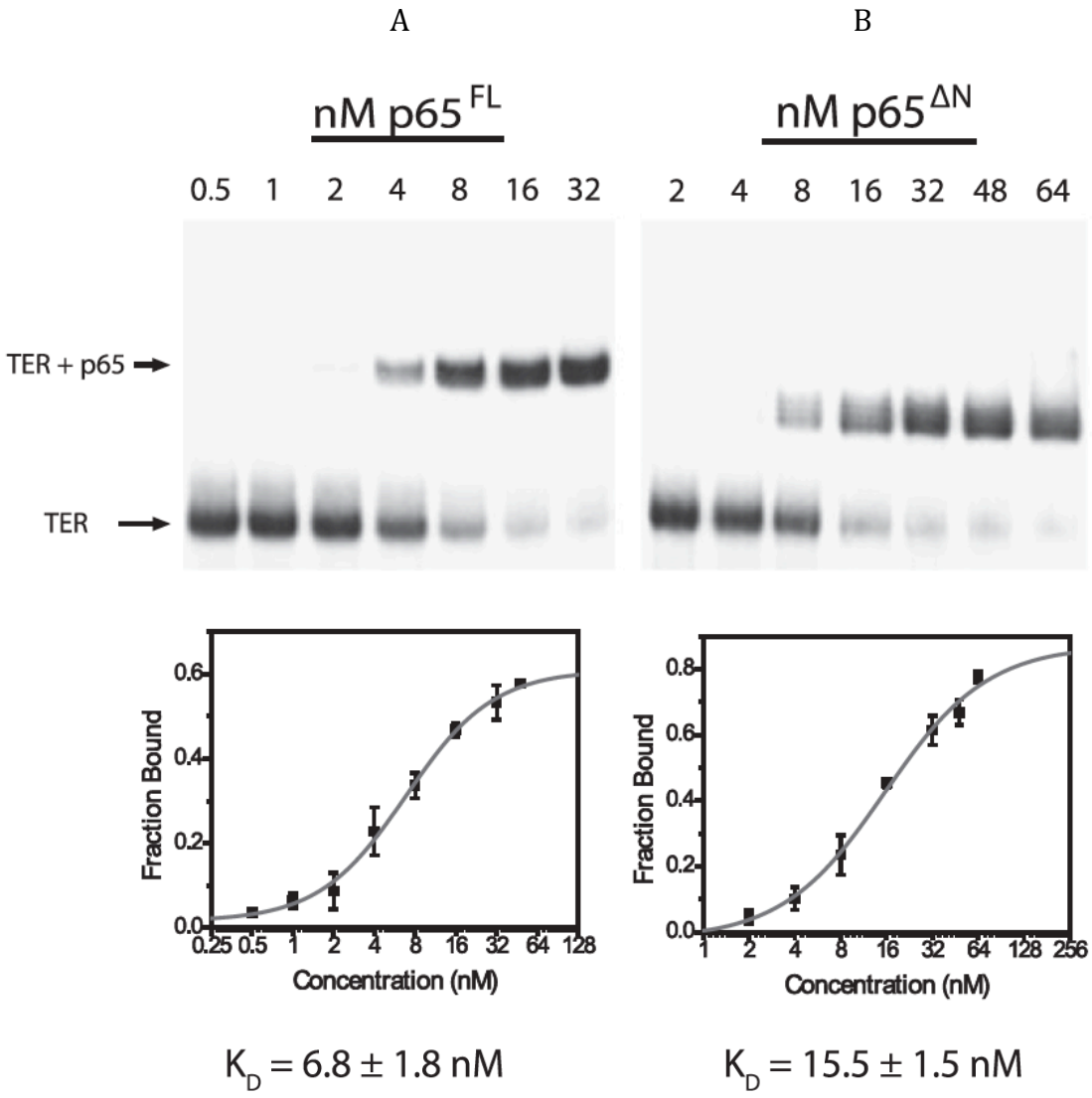


Figure 3. The affinities of increasing concentrations of (A) $p65^{FL}$ and (B) $p65^{\Delta N}$ for binding 100 pM full length TER were measured by gel shift assay.

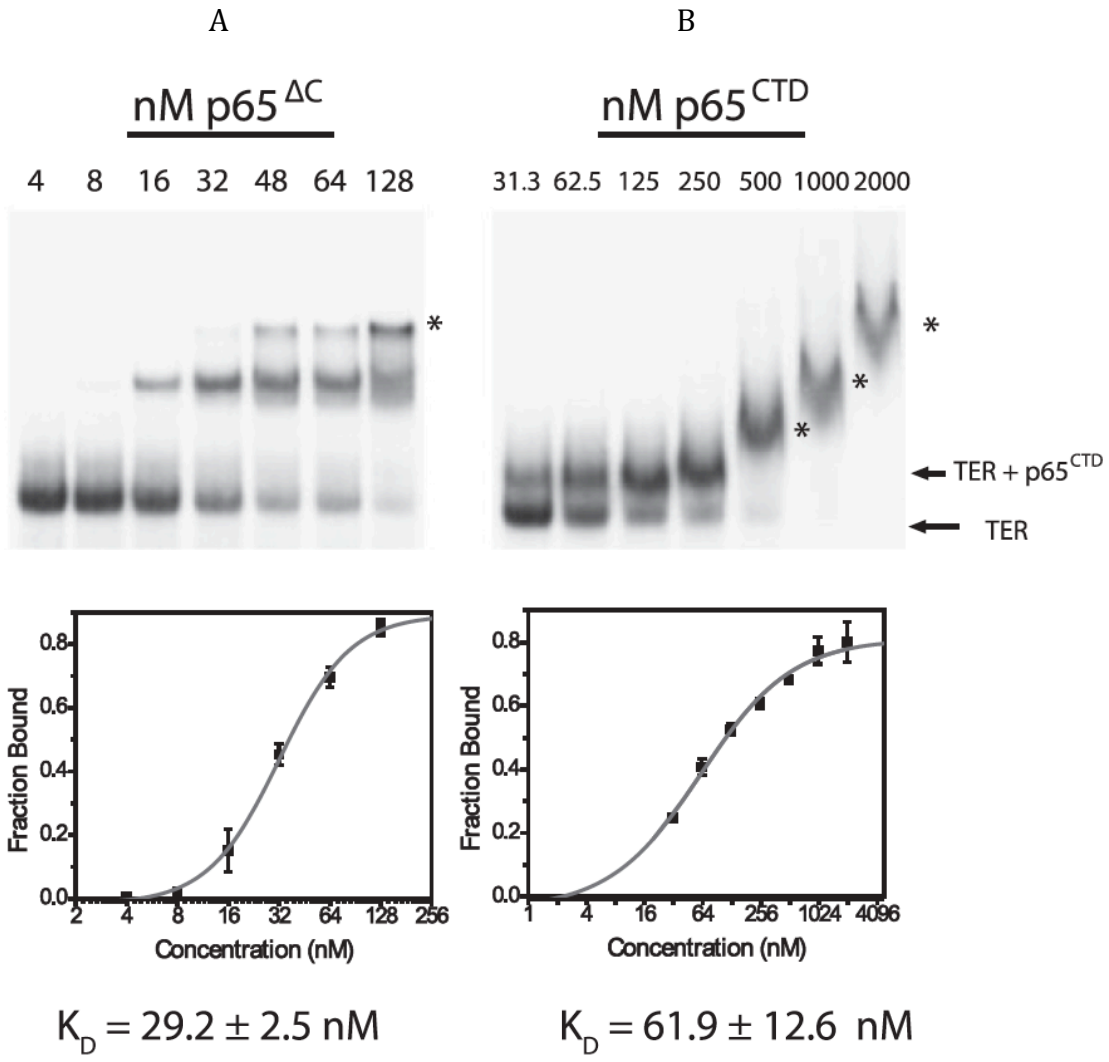


Figure 4. The affinities of increasing concentrations of (A) $p65^{\Delta C}$ and (B) $p65^{CTD}$ for binding 100 pM full length TER were measured by gel shift assay. The asterisks indicate higher order complexes.

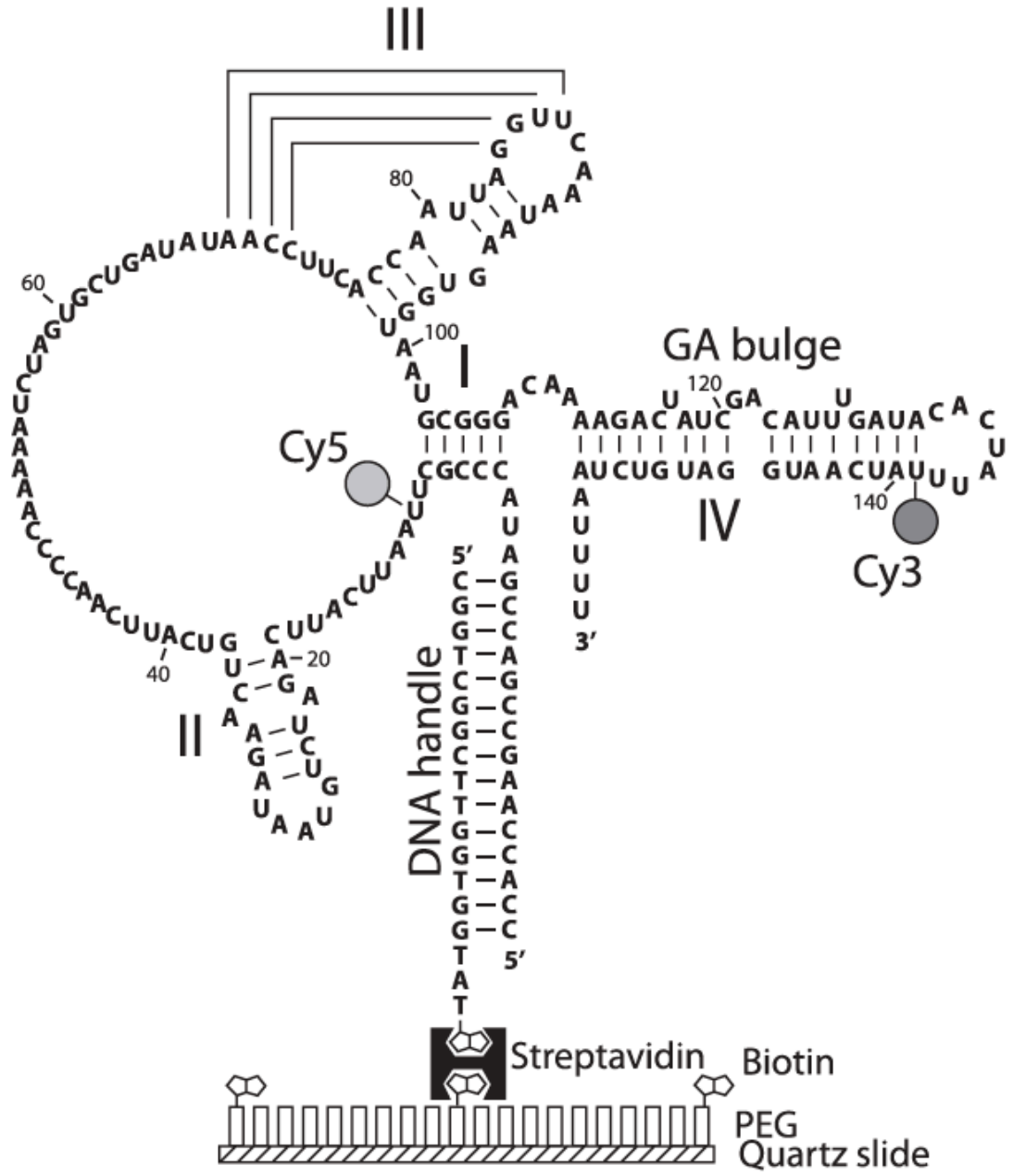


Figure 5. Schematic of the dye labeled TER construct used for the FRET studies. U10 and U139 were labeled with Cy5 and Cy3 fluorescent dyes, respectively. The labeled TER was attached to the biotinylated PEG quartz slide through a biotinylated DNA handle, with streptavidin acting as the linker.

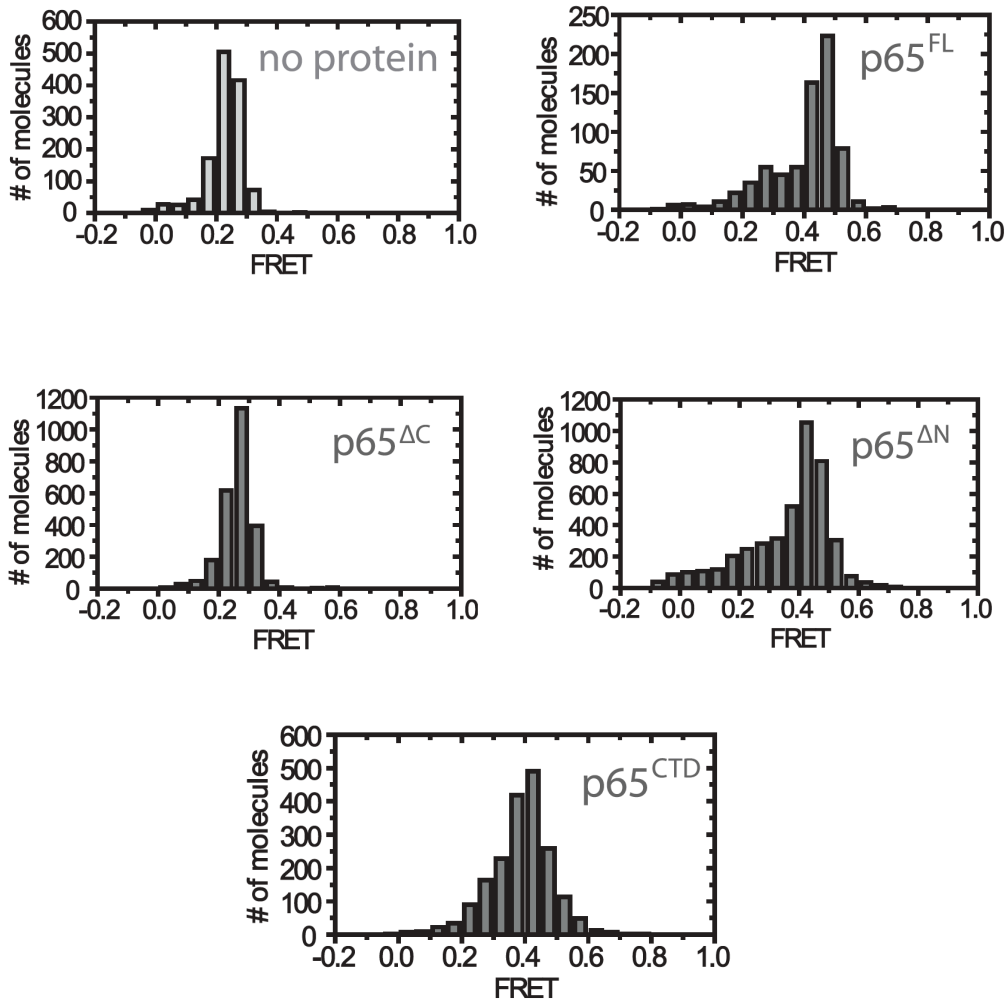
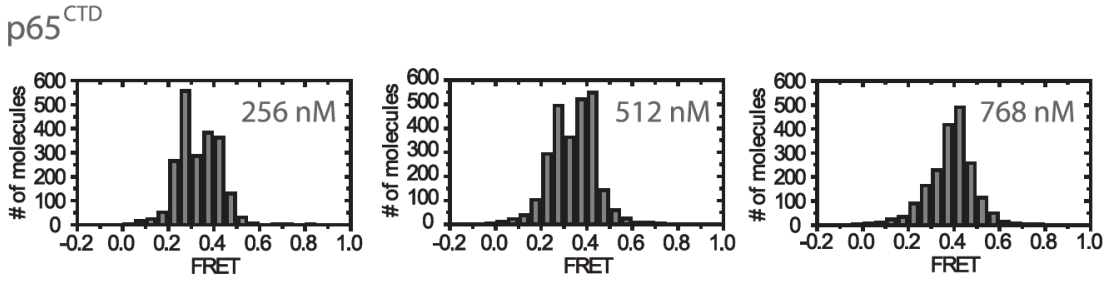


Figure 6. Histograms showing FRET distribution of the p65 constructs. In absence of any protein, the distribution centered around 0.26 FRET. The protein concentrations used were 10 nM $p65^{FL}$, 64 nM $p65^{\Delta C}$, 32 nM $p65^{\Delta N}$, and 750 nM $p65^{CTD}$. Note the similarity of the absence of protein to $p65^{\Delta C}$, and $p65^{FL}$ to $p65^{\Delta N}$. While $p65^{CTD}$ did produce a shift in FRET, it is not as high as $p65^{FL}$ and much broader.

A



B

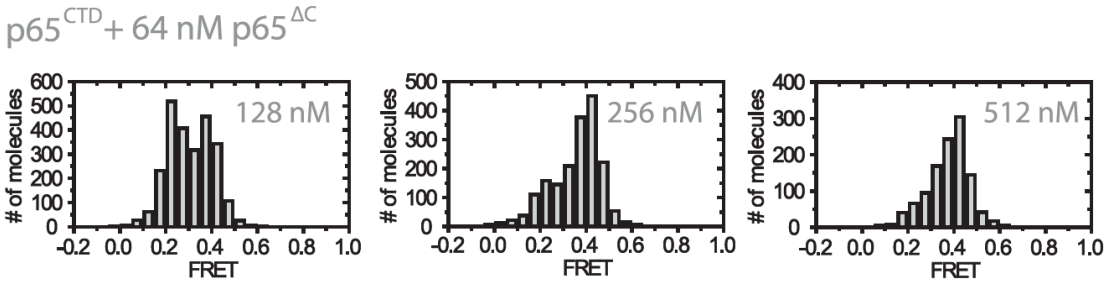
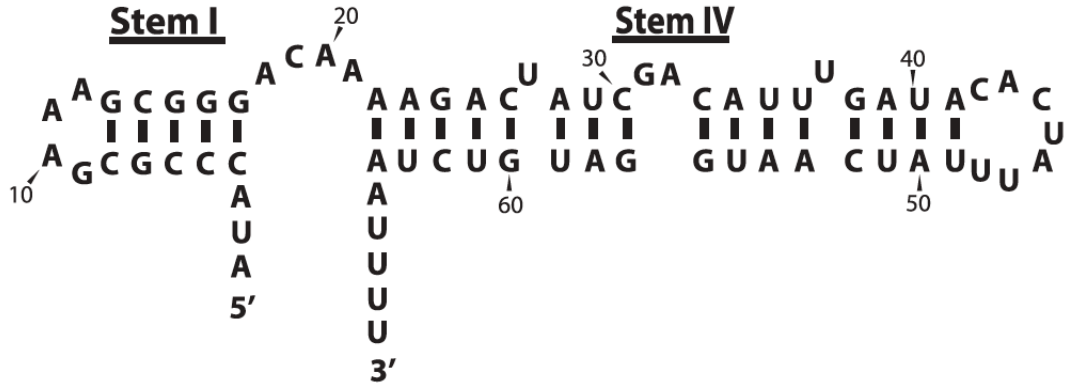


Figure 7. FRET shifts of increasing concentrations of (A) $p65^{CTD}$ and (B) $p65^{CTD}$ with 64 nM $p65^{\Delta C}$. The shift in FRET by $p65^{CTD}$ alone can be achieved at much lower concentrations with the addition of $p65^{\Delta C}$, suggesting $p65^{\Delta C}$ stabilizes the RNA for $p65^{CTD}$ binding.

A



B

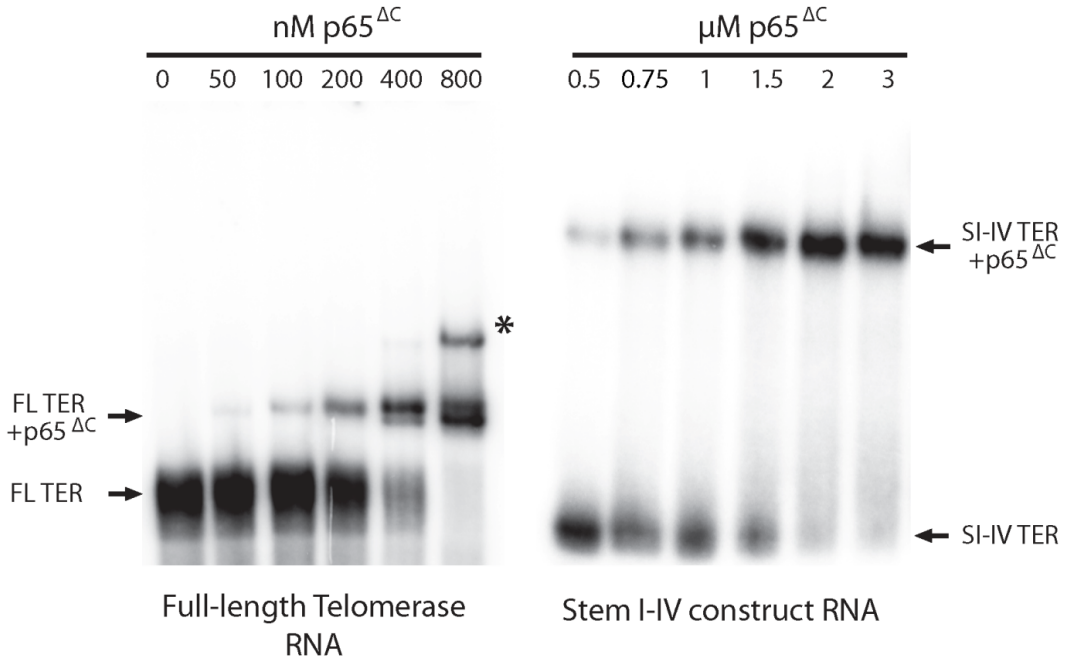


Figure 8. (A) Schematic of the shorter stem I-IV TER construct used in the protection assays. (B) Gel shift assays with TER or stem I-IV TER and increasing concentrations of $p65^{\Delta C}$. The asterisks indicate higher order complexes. The stem I-IV TER construct was able to reach the p65 concentration necessary for protection assays without forming higher order complexes.

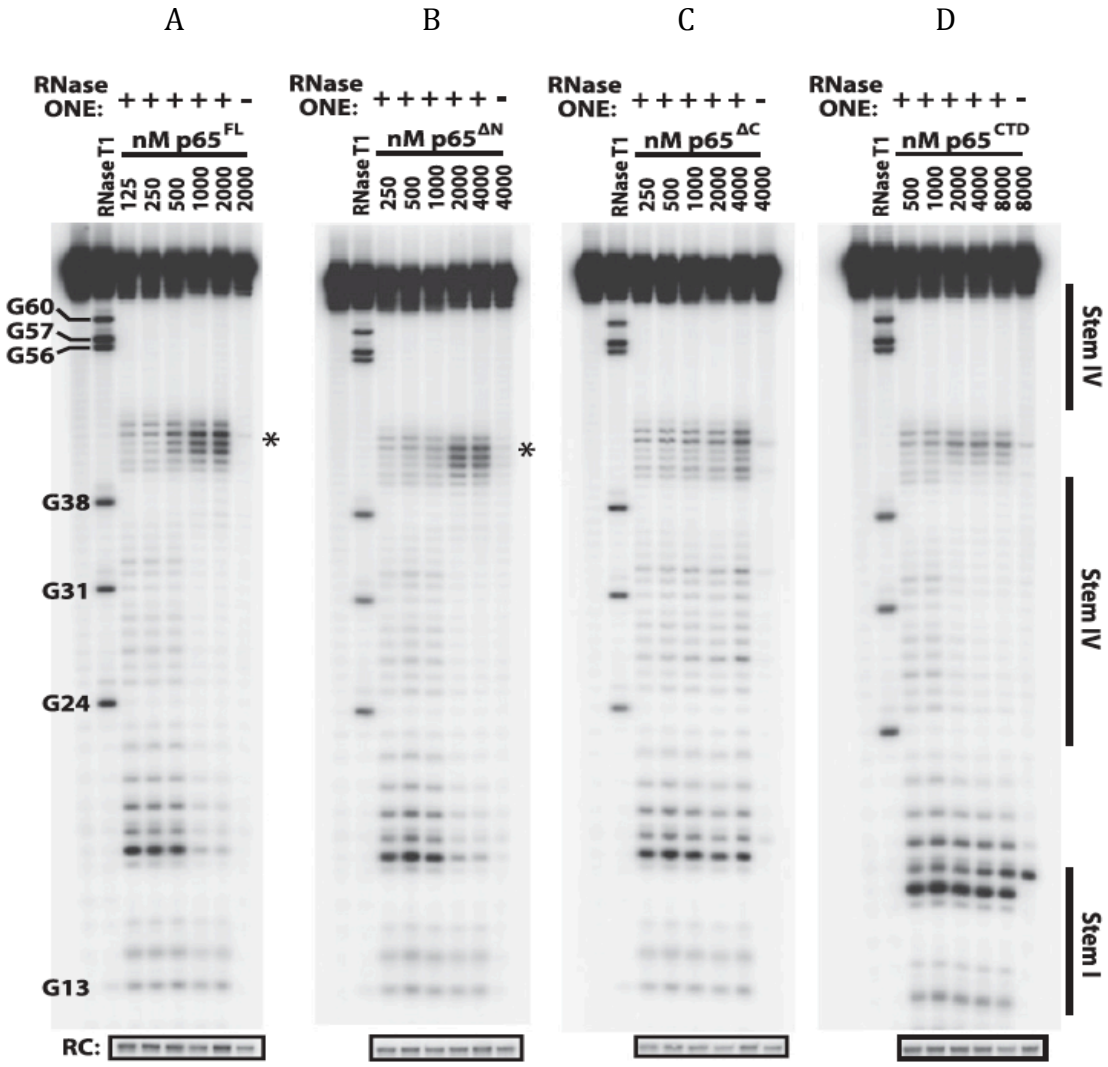


Figure 9. RNase ONE protection gels of stem I-IV construct with increasing concentrations of (A) $p65^{FL}$, (B) $p65^{\Delta N}$, (C) $p65^{\Delta C}$, and (D) $p65^{CTD}$. The RNase T1 digestion lane was used to identify the position of nucleotides. RC denotes recovery control. All constructs but $p65^{\Delta C}$ showed strong protection along stem IV, suggesting that this is the binding site of the C-terminal domain of p65. The asterisks denote the hypersensitivity in stem-loop IV.

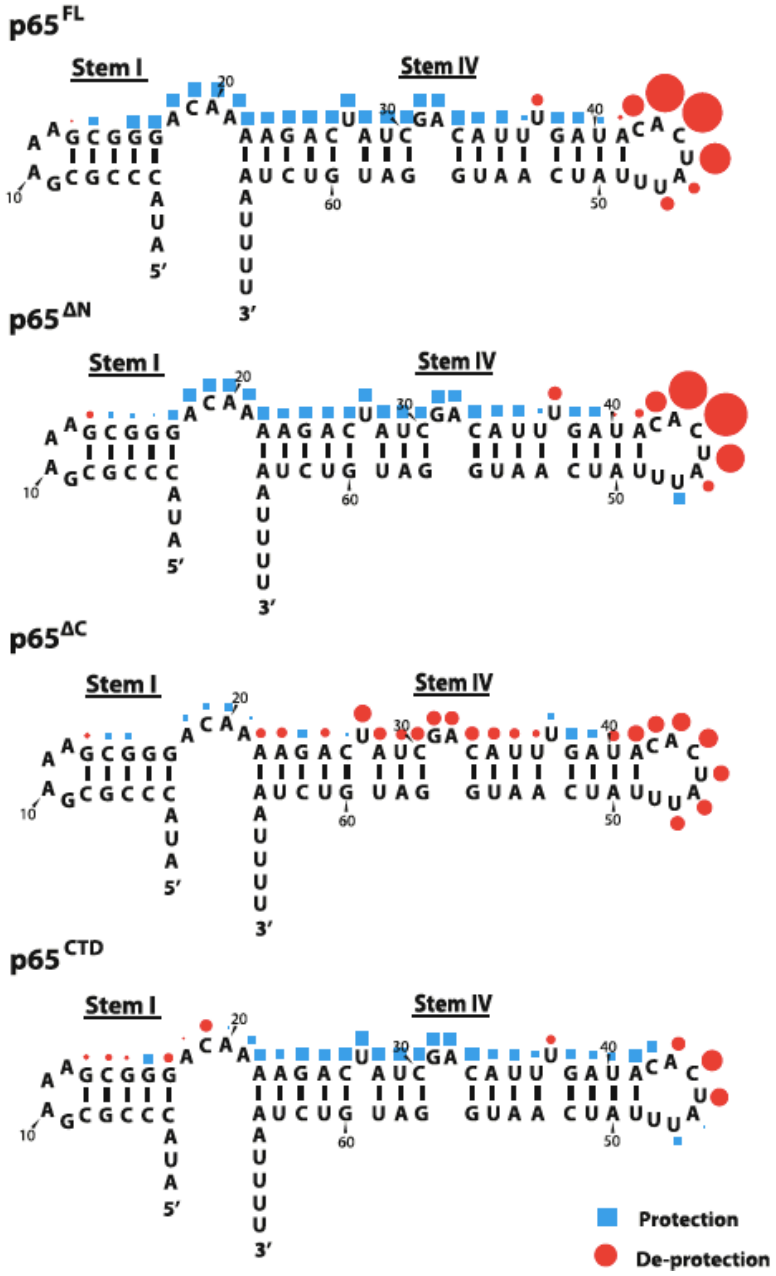


Figure 10. RNase ONE protection results using the semi-automated footprinting analysis software (SAFA) to quantify cleavage at individual nucleotides. The amount of protection or deprotection is scaled to the size of the circles. It is observed that the C-terminal domain binds to stem IV and that binding of p65 destabilizes the nucleotides within stem-loop IV.

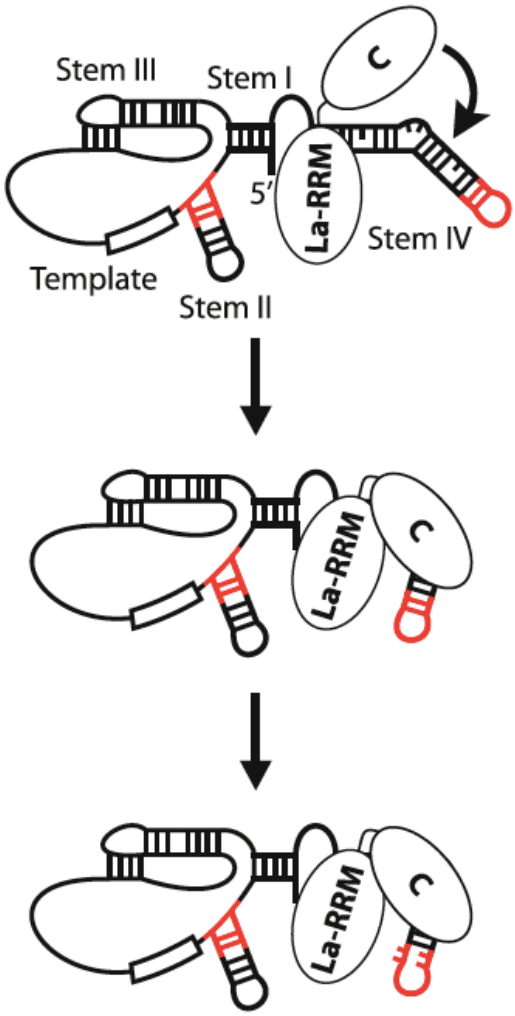


Figure 11. Model of p65 binding and remodeling of *T. thermophila* TER. The La-RRM domain binds to the 3' polyuridine tract of TER, bringing the C-domain near stem IV for binding. Stem IV is bent and orientated in a proper conformation to promote TERT binding. The stem-loop IV bases are also rearranged to further promote TERT binding.

2.4 References

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

Akiyama, B.M., Loper, J., Najarro, K., and Stone, M.D. (2012). The C-terminal domain of Tetrahymena thermophila telomerase holoenzyme protein p65 induces multiple structural changes in telomerase RNA. Rna *18*, 653–660.

Bayfield, M.A., Yang, R., and Maraia, R.J. (2010). Conserved and divergent features of the structure and function of La and La-related proteins (LARPs). Biochimica Et Biophysica Acta (BBA) - Gene Regulatory Mechanisms *1799*, 365–378.

Berman, A.J., Gooding, A.R., and Cech, T.R. (2010). Tetrahymena telomerase protein p65 induces conformational changes throughout telomerase RNA (TER) and rescues telomerase reverse transcriptase and TER assembly mutants. Mol. Cell. Biol. *30*, 4965–4976.

Bokinsky, G., Nivón, L.G., Liu, S., Chai, G., Hong, M., Weeks, K.M., and Zhuang, X. (2006). Two Distinct Binding Modes of a Protein Cofactor with its Target RNA. J. Mol. Biol. *361*, 771–784.

Chen, Y., Fender, J., Legassie, J.D., Jarstfer, M.B., Bryan, T.M., and Varani, G. (2006). Structure of stem-loop IV of Tetrahymena telomerase RNA. The EMBO Journal *25*, 3156–3166.

Ha, T., Enderle, T., Ogletree, D.F., Chemla, D.S., Selvin, P.R., and Weiss, S. (1996). Probing the interaction between two single molecules: fluorescence resonance energy transfer between a single donor and a single acceptor. Proceedings of the National Academy of Sciences *93*, 6264–6268.

Ha, T. (2001). Single-Molecule Fluorescence Resonance Energy Transfer. Methods *25*, 78–86.

Jiang, J., Miracco, E.J., Hong, K., Eckert, B., Chan, H., Cash, D.D., Min, B., Zhou, Z.H., Collins, K., and Feigon, J. (2013). The architecture of Tetrahymena telomerase holoenzyme. Nature *496*, 187–192.

O'Connor, C.M. (2005). Two Purified Domains of Telomerase Reverse Transcriptase Reconstitute Sequence-specific Interactions with RNA. J. Biol. Chem. *280*, 17533–17539.

O'Connor, C.M., and Collins, K. (2006). A Novel RNA Binding Domain in Tetrahymena Telomerase p65 Initiates Hierarchical Assembly of Telomerase Holoenzyme. Mol. Cell. Biol. *26*, 2029–2036.

Prathapam, R., Witkin, K.L., O'Connor, C.M., and Collins, K. (2005). A telomerase holoenzyme protein enhances telomerase RNA assembly with telomerase reverse transcriptase. Nat. Struct. Mol. Biol. *12*, 252–257.

Richards, R.J. (2006). Structural study of elements of Tetrahymena telomerase RNA stem-loop IV domain important for function. Rna *12*, 1475–1485.

Robart, A.R., O'Connor, C.M., and Collins, K. (2010). Ciliate telomerase RNA loop IV nucleotides promote hierarchical RNP assembly and holoenzyme stability. Rna *16*, 563–571.

Singh, M., Wang, Z., Koo, B.-K., Patel, A., Cascio, D., Collins, K., and Feigon, J. (2012). Structural Basis for Telomerase RNA Recognition and RNP Assembly by the Holoenzyme La Family Protein p65. Mol. Cell *47*, 16–26.

Stone, M.D., Mihalusova, M., O'Connor, C.M., Prathapam, R., Collins, K., and Zhuang, X. (2007). Stepwise protein-mediated RNA folding directs assembly of telomerase ribonucleoprotein. Nature *446*, 458–461.

Witkin, K.L. (2004). Holoenzyme proteins required for the physiological assembly and activity of telomerase. Genes & Development *18*, 1107–1118.

Chapter 3

Progress Towards Telomerase Structure Determination by Tethered Ribosome Crystallography

3.0 Introduction

While most of the reverse transcriptase component of telomerase (TERT) structure from the beetle *T*. *castaneum* is known, very few high-resolution structures of the protein components exist for *T. thermophila* telomerase (Gillis et al., 2008; Mitchell et al., 2010). These include the RNA binding domain (RBD) and N-terminal domain (TEN) of TERT, and part of the C-terminal end of p65 (Jacobs et al., 2006; Rouda and Skordalakes, 2007; Singh et al., 2012). For telomerase RNA (TER), the structure of stem-loop IV, stem-loop IV bound to the C-terminal end of p65, and helix II template boundary element (TBE) are known (Chen et al., 2006; Richards, 2006a; 2006b).

There are several methods for obtaining high-resolution structures of macro-molecules. Nuclear magnetic resonance (NMR) is appropriate for non-aggregating, highly soluble proteins less than ~35 kDa. For larger molecules, electron cryo-microscopy (cryo-EM) is useful, but requires extensive averaging due to a low signal to noise ratio, resulting in low-resolution structures. Crystallography, one of the oldest methods for obtaining high-resolution structures, uses the resulting diffraction pattern when a beam of x-rays passes through a crystal formed by the molecule(s) of interest. From this pattern, the electron density can be calculated, and if resolution is high enough, an atomic model can be built into the density.

Crystallography, however, is not a simple process. The molecule of interest requires extensive purification for homogeneity. This purification process becomes more difficult for complex molecules. Crystallization is the most difficult step, requiring extensive screening of buffers, temperatures, and techniques to induce crystallization. This process could also involve the removal of flexible regions and surface residues so that the molecule may aggregate in such a way that a repeating crystal lattice is formed. Sometimes no crystal can be obtained (Chruszcz et al., 2010). If crystals are obtained, they must be cryoprotected with increasing precipitant concentration to remove as much water as possible before freezing. This step also requires screening for optimum cryoprotectant solutions.

Diffraction data is collected by mounting the crystal on a goniometer and passing short bursts of x-rays though the crystal as it rotates. Two types of data are needed to reconstruct electron density: intensity and phase. Intensities are found by diffraction patterns, but phase information is lost, also known as the "phase problem". Isomorphous replacement and anomalous scattering can be used to solve this problem, but if the structure of a similar molecule is known, molecular replacement can be used to find the initial phases and structure refined until calculated intensities match observed intensities.

The ribosome has been studied extensively through the years and its crystallization conditions are well known. This opens the possibility of inserting

telomerase RNA (TER) into ribosomal RNA, and using the ribosome as a purification chaperone and crystallization scaffold. The ribosomes can be easily purified using established methods, and because of its large size relative to the inserted RNA, it will dominate the properties that form the crystal (Fig. 1). The following work details some of the first steps in developing this novel method that can be extended to other RNPs and RNAs.

3.1 Results

3.1.1 *E. coli* Strain Δ7_cm_ts

To simplify purification and ensure homogeneity, an approach was taken to create an *E. coli* strain that contain only ribosomes with the desired RNA insertion. The strain Δ7_cm_ts is derived from a strain that has had all chromosomal ribosomal RNA operons (rrn) deleted, and is instead supplied with ribosomes from a plasmid containing a single rrn (Asai et al., 1999a; 1999b). This plasmid is chloramphenicol resistant and temperature sensitive at 42°C. By transforming a kanamycin plasmid containing a rrn with our insertion into this strain, growing at 42°C, testing for chloramphenicol sensitivity and kanamycin

resistance, a strain is obtained that will produce only ribosomes containing the insertion.

3.1.2 Ribosome Construction

Helix 63 (H63) of the 23S ribosomal RNA was first chosen as the insertion site. Eukaryotic ribosomal RNA contains a large insertion at H63, and before the high-resolution structure of the ribosome was known, a tRNA was inserted to find the location of H63 using cryo-EM (Spahn et al., 1999). A study searching for viable ribosomal RNA insertions identified not only H63, but also H25 as candidates, so H25 was chosen as well (Yokoyama and Suzuki, 2008). The high-resolution structure was also examined to verify an insertion at H63 or H25 would not obviously disrupt the expected ribosome crystal packing (Fig. 1).

A unique cut site, SpeI, was placed within the stem loop at either H63 or H25 in the plasmid pKK3535. Full length *Tetrahymena thermophila* TER was cloned into this cut site, replacing the terminal loop of the helix. Three adenosines were added to the 5' end of TER to pair with the naturally unpaired poly-uridine tail, possibly making the insertion less flexible for diffraction purposes (Fig. 2). Because the termini are naturally unpaired, any structure obtained may not reflect TER's natural folding. Another construct, CIRCTER, was

made for independent verification of correct folding. This is a circular permutation construct that is attached to the ribosomal RNA at stem II. It was previously shown that adding a short linker to the natural termini still results in a functional molecule (Miller, 2002). Each insertion was transferred to plasmid pCsacB7 for kanamycin resistance. After insertion, the plasmid was transformed into Δ7_cm_ts, replacing the original plasmid, creating a strain that produced only ribosomes containing a TER or CIRCTER insertion at H63.

While the strain containing the insertion at H63 was viable, its doubling time was longer than wildtype for both TER and CIRCTER. The TER H25 insertion was first thought to be lethal, but was later determined that it severely compromised cell growth rate, so only ribosomes containing the H63 insertion were used. A CIRCTER H25 construct was not attempted.

3.1.3 Purification and Crystallization of Ribosomes Containing Telomerase RNA Insertions

Cells were grown to stationary phase and ribosomes purified using a sucrose gradient (Fig. 3). Mutant ribosome purity was confirmed by performing PCR with primers that flank the insertion site. Growth was repeated if multiple bands were observed. Both constructs appeared to have slightly weaker

30S/50S association than WT ribosomes, but looked similar. Crystallization screens were set up around known initial conditions for *E. coli* ribosome crystallization, and small rod-like crystals appeared within 3-4 weeks in several conditions, however only one of the crystals of the TER construct diffracted to the poor resolution of ~30-40Å (Fig. 4). After additional initial screening, a new condition for TER was found at 4°C that produced faster growing and larger crystals (Fig. 5). Unfortunately these crystals failed to diffract.

3.1.4 Addition of p65

The *T. thermophila* catalytic core telomerase protein p65 was expressed and purified from *E. coli*. Initial binding studies were performed by adding a 5-fold molar excess of p65 to TER and CIRCTER ribosomes immediately before the final gradient, purified, and analyzed by SDS-page (Fig. 6). These ribosomes were not dissociated and reassociated prior to final purification, and still contain the ribosomal protein S1, which is very similar in size to p65. This made it unclear whether p65 was binding to the telomerase RNA insertions. Running higher percentage and larger gels also failed to resolve the difference. Interestingly, there was a lower band for TER ribosomes that was also observed for p65 alone

after purification from *E. coli*. This band was excised, analyzed by mass spectrometry, and discovered to be p65 with a N-terminal deletion (Fig. 7).

3.2 Discussion

While there was diffraction for ribosomes containing the TER insertion, resolution was too low for any meaningful density. Considering this crystal was produced from an initial screen is exciting; along with screening additional cryoprotection solutions, further well buffer optimizations may improve crystal quality and size. There are, however, significant obstacles to overcome and improvements to be made.

Previous efforts toward this project by Michael Pearson in Harry Noller's lab took a direction of co-expressing the catalytic core proteins p65 and TERT (reverse transcriptase component) during ribosome production. This greatly slowed cell growth, taking a week to reach stationary phase. Like TER RNA alone, crystals were obtained, but did not diffract well enough. Because of the long growth time, protein and RNA quality were unknown, so an approach of adding p65 at the last purification step was taken. A telomerase activity assay was also performed with the ribosomes, but showed unusual reverse transcriptase activity, which could be caused by a multitude of factors. Spliceosomal U2 and

U6 snRNA was also inserted at H63. Crystals were obtained and diffracted to 10.5Å, proving that this could be a viable method and warranted further exploration.

While crystals can be grown with the TER insertion, the folded state of the RNA may be misfolded or heterogeneous. Also, since the insertion likely makes no important crystal contacts and doesn't interfere with crystal packing, insert homogeneity is not necessarily required for crystal growth, leading to poor electron density in the area of insertion. To improve the positioning of inserted RNA, the attached stem can be extended one base pair at a time to rotate the molecule, or other insertion sites can be explored. This may also resolve the unusual reverse transcriptase activity that was previously observed when catalytic core proteins were constituently expressed. One method to find new sites could be the use of Tn5 transposase in a similar fashion to Yokoyama *et al*. This is fairly straightforward and would immediately select for viable cells.

Another approach for crystallization would be the dissociation of the ribosomes and attempting crystallization with only one of the subunits, depending on where the insertion is. This would ease purification by improving yield and relieve the concern of the ribosomes dissociating during or after purification.

Optimization of the Δ7_cm_ts strain may improve ribosome and crystal quality. Along with the ribosomal operons, some tRNA genes were also deleted

from the chromosome during construction of the original strain. These missing tRNA genes are supplied with an additional plasmid and may be causing regulation issues in the cell since they are not expressed using a rRNA or tRNA promotor. When plated, colony sizes of Δ7_cm_ts vary considerably, possibly due to the loss of either of the two plasmids. A solution would be to place the tRNA genes on the rRNA plasmid in the rrn operon.

The result of p65 binding to the RNA insertion was unclear because of the inability to resolve p65 from S1. A solution would be to perform a western blot for p65. It's possible that p65 is binding neither to TER or CIRCTER. What's clear is that the N-terminal deletion clearly binds to the TER ribosomes with decent affinity, suggesting that an N-terminal deletion construct of p65 should be used for further studies, especially since the N-terminal is known to be "floppy" (which citation?). Interestingly, another group experienced similar protease digestion (Jiang et al., 2013).

Difficulties lie with sufficient protein production and crystal growth rates. A significant quantity of protein is required in order to add a 5 molar excess of p65 to the ribosomes before the final purification step. Expressing and purifying this amount greatly hinders steady crystal trials. This can be circumvented by taking a similar approach as the previous work, and introduce a plasmid that will constitutively express p65 lacking the N-terminal. By not also constitutively expressing TERT, near normal growth rate may return. The rate of crystal

growth also hinders steady crystal trials due to the 3-4 week growth rate. While faster growing crystals were found at 4°C, they failed to diffract, but further optimization of well buffer and cryoprotectant conditions may prove fruitful.

Further exploration of crystallization conditions, insertion sites, and additional catalytic core protein constructs may prove this method as an invaluable tool for solving RNP or RNA structures that have failed to produce crystals on their own.

3.3 Methods

3.3.1 Inserting a Unique Cut Site Into H63

To introduce a SpeI cut site at H63, whole plasmid PCR was performed using pkk3535 (Brosius et al., 1981) as the template, and primers 23S_H63_D8_aptamer_F and 23S_H63_D8_aptamer_R. This was purified using a PCR purification silica gel spin column (Qiagen). 8 µL was mixed with 1µL New England Biolabs (NEB) buffer 1 and 1 µL 1:100 dilution exonuclease III (1 unit) (NEB). The mixture was incubated at room temperature for 5 minutes, placed on ice for an additional 5 minutes, and transformed into *E. coli* strain 10-beta (NEB).

Transformants were selected on LB-agar plates containing 50 μg/mL ampicillin at 34°C. The resulting plasmid was digested with SpeI (NEB) and ligated with T4 DNA ligase (NEB), resulting in the plasmid pkk3535_H63_speI. This plasmid contains a single SpeI cut site within the H63 stem loop.

23S_H63_D8_aptamer_F

GAGTAATTTACGTTTTGATACGGTTGCGGACTAGTGGATGGAGCTGAAATCAGTCG

23S_H63_D8_aptamer_R

GCAACCGTATCAAAACGTAAATTACTCGGACTAGTGGACCTCACCTACATATCAGC

3.3.2 Inserting a Unique Cut Site Into H25

Done in the exact fashion as H63, except primers 23S_H25_D8_aptamer_F and 23S_H25_D8_aptamer_R were used, resulting in the plasmid pkk3535_H25_speI. This plasmid contains a single SpeI cut site within the H25 stem loop.

23S_H25_D8_aptamer_F

GAGTAATTTACGTTTTGATACGGTTGCGGACTAGTGCGTGTGACTGCGTACCTTTTG

23S_H25_D8_aptamer_R

GCAACCGTATCAAAACGTAAATTACTCGGACTAGTGCGTGCTCCCACTGCTTGTACG

3.3.3 Inserting Telomerase RNA Into H63

The plasmid pkk3535_H63_speI was digested with SpeI (NEB) and gel purified. *T. thermophila* TER RNA was PCR amplified using primers F_23S_H63_TetRNA and R_23S_H63_TetRNA. The resulting product was purified using a spin column, combined with the previously cut plasmid at a 1:1 molar ratio, and joined using the exonuclease III method as before. This results in the plasmid pkk3535_H63_TER.

F_23S_H63_TetRNA

GCACGCTGATATGTAGGTGAGGTCCAAAATACCCGCTTAATTCATTCAG

R_23S_H63_TetRNA

TCTTCGACTGATTTCAGCTCCATCCAAAAATAAGACATCCATTGATAAATAG

The circular permutation of TER was constructed as with TER, but by using primers F1_23S_H63_circ_TetRNA and R1_23S_H63_circ_TetRNA. This product was diluted 1000 fold and mixed with primers F1_23S_H63_circ_TetRNA and R2_23S_H63_circ_TetRNA. The resulting PCR product was again diluted 1000 fold and mixed with primers F_insert_H63_23S and R_insert_H63_23S. The resulting PCR product was inserted into H63 as before, resulting in the plasmid pkk3535_H63_CIRCTER.

F1_23S_H63_circ_TetRNA

GCACGCTGATATGTAGGTGAGGTCCAGAACTGTCATTCAACCCCAAAAATC

R1_23S_H63_circ_TetRNA

GAATGAATTAAGCGGGTATGGATCCAAAAAAAAATAAGACATCCATTGATAAATA

F_insert_H63_23S

GGTGCCGTAACTTCGGGAGAAGGCACGCTGATATGTAGGTGAGGTCC

R_insert_H63_23S

AAACAGTTGCAGCCAGCTGGTATCTTCGACTGATTTCAGCTCCATCC

3.3.4 Inserting Telomerase RNA Into H25

Done in the exact fashion as with H63, except primers F_23S_H25_TetRNA and R_23S_H25_TetRNA were used, resulting in the plasmid pkk3535_H25_TER. Transformants took ~2-3 days at 34°C to become visible. A CIRCTER construct was not attempted.

F_23S_H63_TetRNA

GCACGCTGATATGTAGGTGAGGTCCAAAATACCCGCTTAATTCATTCAG

R_23S_H63_TetRNA

TCTTCGACTGATTTCAGCTCCATCCAAAAATAAGACATCCATTGATAAATAG

3.3.5 Construction of H63 Mutant Ribosome Strains

Both H63 TER and CIRCTER insertions were transferred to the kanamycin resistant plasmid pCsacB7 (Asai et al., 1999b). This was performed by PCR amplification of the pkk3535 plasmids containing the insertions using primers 23S_H63_Bsu36I_F and 23S_H63_Bsu36I_R and purified as before. The plasmid pCsacB7 was digested using Bsu36I (NEB), gel purified, and combined with the appropriate PCR product using exonuclease III as before, resulting in plasmids pKAN_H63_TER and pKAN_H63_CIRCTER. These plasmids were transformed into the strain Δ7_cm_ts. The plated cells were grown at 42°C, tested for chloramphenicol sensitivity and kanamycin resistance, resulting in strains producing only producing ribosomes containing TER or CIRCTER insertions.

23S_H63_Bsu36I_F

AAGGGTTCCTGTCCAACGTTAATCGG

23S_H63_Bsu36I_R

AGACAGCCTGGCCATCATTACGCCATT

3.3.6 Growth and Purification of Mutant Ribosomes

From a 5mL overnight starter culture, cells were grown in 2xYT at 37°C, and shaken at 225 rpm in 2L Erlenmeyer beveled culture flasks. Once stationary phase was reached, typically ~35 hours with an OD of ~4-5, the cells were pelleted at 5,000 x g at 4°C. The cells were washed twice with buffer A (100 mM KCl, 15 mM $MgCl_2$, 20 mM Tris-Cl pH7.5, 5 mM BME) and flash frozen in liquid nitrogen if necessary. Contamination was tested at this point by performing PCR on 0.5 uL of cells using primers 23S_H63_Bsu36I_F and 23S_H63_Bsu36I_R. Purification was continued if only one band was observed.

The fresh or frozen pellets were resuspended in 35 mL buffer A, and lysed by passing 3 times through a cell disruptor at 15,000 psi. The cell lysate was clarified by centrifugation 3 times at 35,000 x g for 20 minutes in a JA20 rotor. The cleared lysate was layered on top of a 35 mL 40% sucrose cushion containing buffer B (500 mM KCl, 15 mM $MgCl_2$, 20 mM Tris-Cl pH7.5, 5 mM BME) in a Ti45 ultracentrifuge tube. The tube was filled to the neck with buffer B and the ribosomes were pelleted by centrifugation at 39,000 rpm for 20 hours in a Ti45 rotor. The ribosome pellets were gently resuspended into 25 mL buffer with a small stir bar while on ice, and then clarified again by centrifugation at 35,000 x g for 20 minutes in a JA20 rotor. The ribosomes were transferred to Ti60 ultracentrifuge tubes, filled to the neck with buffer B, and pelleted again by

centrifugation in a Ti60 rotor at 49,000 rpm for 3 hours. The pellet was quickly washed twice with buffer A and left upside down at 4°C until the pellet became clear, typically 10-15 minutes. The ribosomes were gently resuspended in 1 mL buffer A as done before and frozen with liquid nitrogen if necessary.

10-35% sucrose gradients were made by carefully layering 19 mL of a buffer A 10% sucrose solution on top of a 19 mL buffer A 35% sucrose solution in SW-28 tubes. These tubes were gently laid on their side for 3 hours at RT the day before and cooled to 4°C overnight. The ribosomes were diluted to 40 mg/mL with buffer A, and 15 mg was loaded onto each 10-35% sucrose gradient and centrifuged for 17 hours at 22,000 rpm. The 70S peaks were conservatively collect by pumping the gradients and monitoring the absorbance at 260 nm. The peaks were pooled in a Ti45 tube and pelleted by centrifugation at 39,000 x g for 20 hours in a Ti45 rotor. The ribosome pellets were gently resuspended in 10 mL buffer A as before. The buffer was exchanged twice with 15 mL of buffer A using a centrifugal concentrator and brought to a final concentration of 15 mg/mL. The ribosomes were aliquoted and flash frozen in liquid nitrogen until used for crystallization.

3.3.7 p65 Expression, Purification, and Binding

From a 5mL overnight starter culture, cells were grown in terrific broth, at 37°C, and shaken at 225 rpm in 2L Erlenmeyer beveled culture flasks. After reaching mid-log growth phase, expression was performed using 0.8 mM IPTG at 16°C and cells harvested after 4 hours. The fresh or flash frozen cells were resuspended in buffer A (20 mM Tris pH 8.0, 200 mM NaCl, 1 mM $MgCl_2$, 40 mM imidazole, 1 mM DTT, and 10% glycerol). 1 mM PMSF was added just before passing the cells twice through a cell disruptor. The cell lysate was cleared by centrifugation at ~36,000 x g and passed over Glutathione HiCap Matrix resin (Qiagen) by gravity flow. The resin was thoroughly washed with buffer A that contained 500 mM NaCl until no protein was detected by Bradford assay. After washing with a column volume of buffer A containing 100 mM NaCl, the protein was eluted with buffer A containing 100 mM NaCl and 50 mM glutathione until no protein was detected. TEV was added to the elution and left overnight at 4°C to remove the GST tag. The cleaved protein was diluted to 100 mL with buffer B (20 mM Tris pH 8.0, 100 mM NaCl, 1 mM $MgCl_2$, 1 mM DTT, and 10% glycerol), passed over a Q anion exchange column (GE Life Sciences), and eluted with a 60 minute buffer B gradient to 1 M NaCl. The elution was once again passed through the glutathione resin to remove any excess GST. The protein was concentrated with a spin concentrator (Amicon) and passed over a Superdex 200 size-

exclusion column (GE Life Sciences) in buffer C (20 mM Tris pH 8.0, 100 mM NaCl, 1 mM $MgCl_2$, 1 mM DTT, and 10% glycerol) 10 mg at a time maximum. The protein was aliquoted, flash frozen, and stored at -70°C. Binding studies were performed by supplementing the protein solution with 15 mM $MgCl_2$ and adding it to 15 mg of ribosomes in 5-fold molar excess for 15 minutes at RT immediately before the sucrose gradient.

3.3.8 Crystallization

Crystallization was performed by sitting drop vapor diffusion using a Crystal Phoenix automatic robot (Art Robbins Instruments) with 96 well plates. The diffracting crystal was obtained using a well buffer containing 3.5% polyethylene glycol 20,000 MW (PEG20K), 13% polypropylene glycol P 400, 0.2 M potassium thiocyanate (KSCN), and 0.1 M tris-acetate pH 7. Ribosomes at 15 mg/mL were combined with well buffer in 2:1, 1:1, and 1:2 ribosome:buffer ratios. Total drop volumes were 0.3 µL or 0.4 µL. After 48 hours at 25°C, the trays were transferred to 16°C, and crystals appeared within 3-4 weeks. Cryoprotection was performed at 16°C by using 3 cryoprotectants containing 0.2 M KSCN, 0.1 M Tris-acetate pH 7, 10 mM Mg acetate, 13% polypropylene glycol P

400, and 7%, 14%, or 21% PEG20K. The PEG20K concentration was increased every 24 hours.

Crystals that formed at 4°C were laid in the exact fashion as before, but with well buffer containing 2.2-5% PEG20K, 17-19% 1,6-hexanediol, 0.2 M KSCN, and 0.1 M tris-acetate pH7. Crystals formed within two weeks and cryoprotection was performed as before, using 3 cryoprotectants containing 5% PEG20K, 10 mM Mg acetate, 0.1 M tris acetate pH7, 0.2 M KSCN, 15%, 20%, or 25% MPD, and 3%, 6%, or 9% PEG200. The MPD and PEG200 concentrations were increased every 24 hours.

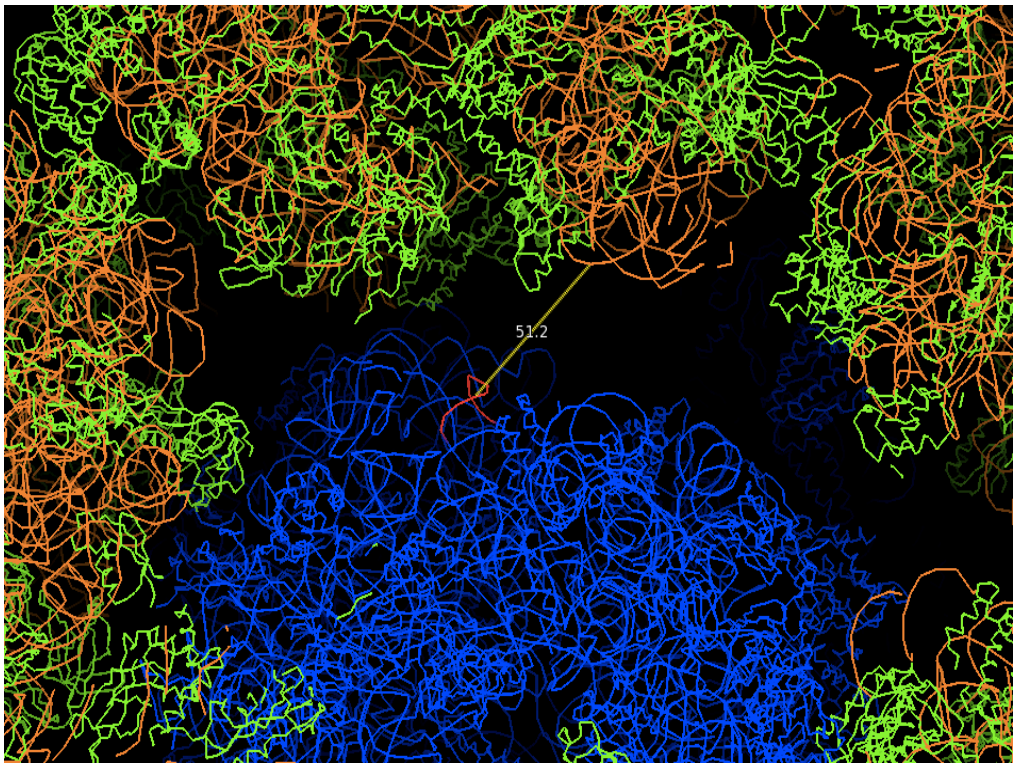


Figure. 1. Visualization of the expected crystal packing of the ribosomes and the available volume at H63 using Pymol and the symexp command. Red highlights H63 of the blue ribosome, while green and orange indicate neighboring ribosomes. PDB file provided by the Noller lab, UC Santa Cruz.

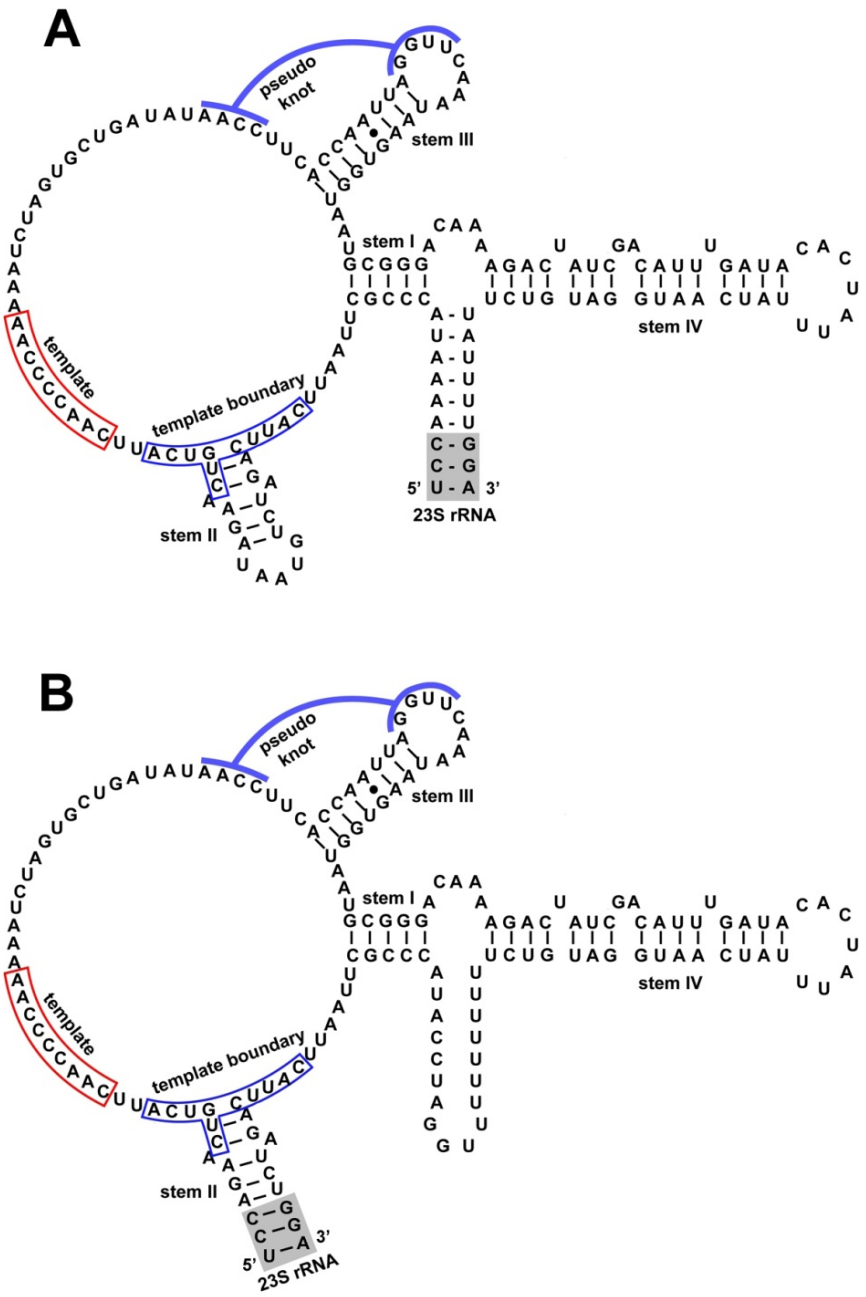


Figure 2. Secondary structure of the two telomerase constructs, TER and CIRCTER. (A) TER is attached through a stem introduced at the natural termini and (B) CIRCTER is a circular permutation attached at helix II.

Figure 3. 10-35% sucrose gradient profile of H63 TER ribosomes grown from Δ7_cm_ts *E. coli*. CIRCTER ribosomes appeared identical. The ribosomes used for crystallization were conservatively collected, as noted by the red dashes on the trace. Other than increased 70S dissociation, gradient profiles appeared similar to WT ribosomes.

Figure 4. TER ribosome crystals that formed at 16°C and successfully diffracted to ~30-40Å. Crystallization conditions were 3.5% PEG20K, 13% PPG 400, 0.2 M KSCN, and 0.1 M tris-acetate pH 7. Trays were left at 25°C for 48 hours before placed at 16°C, and crystals appeared within 3-4 weeks.

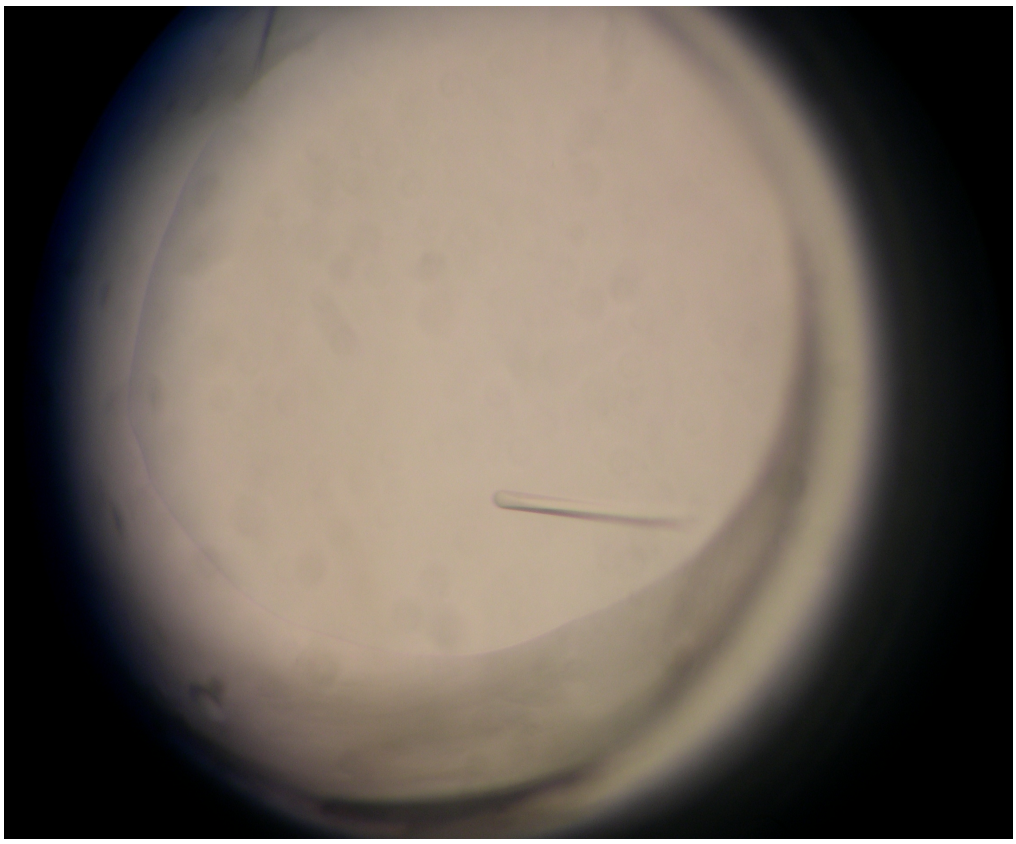


Figure 5. TER ribosome crystals that formed at 4°C, but did not diffract. Crystallization conditions were 2.2-5% PEG20K, 17-19% 1,6-hexanediol, 0.2 M KSCN, and 0.1 M tris-acetate pH7. Trays were left at 25°C for 48 hours before placed at 4°C, and crystals appeared within 2 weeks.

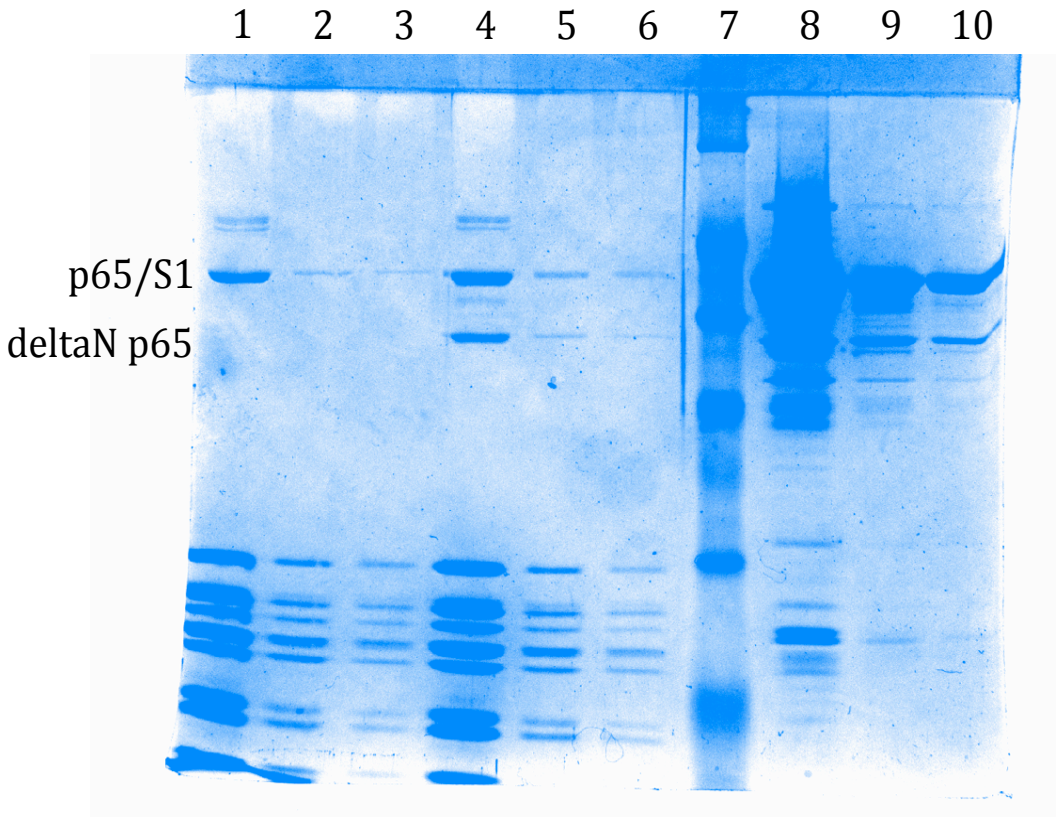
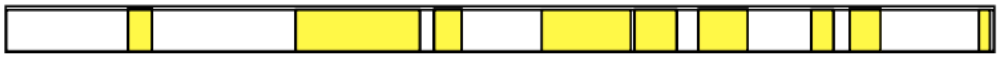


Figure 6. SDS-page gel of ribosomes that were incubated with 5-fold molar excess p65 immediately before purification by sucrose gradient. Lanes 1-3: CIRCTER ribosomes. Lanes 4-6: TER ribosomes. Lanes 8-10: p65 expressed and purified from *E. coli*. The ribosomal protein S1 could not be resolved well enough from p65 to clearly identify binding. Only TER appears to bind the N-terminal truncation of p65.

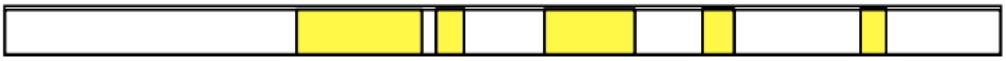
(A)



Protein:

MDEYLENTNL EELEQECFME DYQHEDVVEQ ENHQVDANDI YENQQMNDES
QLNQDVKISQ QKEQAVEMIE EQQQNNQDKF KQFQDCMAHI TELNFKRNYQ
NLTEQSSSNN VVAEELDIKE SLKLQMEYYF CDTNLTHDSY LRGIISKSPK
NCVDIKVFLK FNKIQQILKQ IQDKQIVSTY GIENQSQKKN HKNYKNQNAT
FSKKDLIHLI RDSLKESKIL KVKMDSLKVK RRFPFNLEQA LKNSKQRTLY
IDFLPPKCSK QTLVSIFGNF RIININLPLQ KNSQLCQGFA FIEFFSEEEA
NQALITKNSS IPKELILLTE KKIGQGSIRI ITYKKWQEEK QSFKELSKNQ
NEQKNKNMNQ SRKASDEFVS IDVEIKQNCL IKIINIPQGT LKAEVVLAVR
HLGYEFYCDY IDENSNQINS NKISLSTQQQ NTAQCSNIQI ENNLIQQDQH
PQLNDLLKEG QAMIRFQNSD EQRLAIQKLL NHNNNKLQIE IRGQICDVIS
TIPEDEEKNY WNYIKFKKNE FRKFFFMKKQ QKKQNITQNY NK

(B)



Protein:

MDEYLENTNL EELEQECFME DYQHEDVVEQ ENHQVDANDI YENQQMNDES
QLNQDVKISQ QKEQAVEMIE EQQQNNQDKF KQFQDCMAHI TELNFKRNYQ
NLTEQSSSNN VVAEELDIKE SLKLQMEYYF CDTNLTHDSY LRGIISKSPK
NCVDIKVFLK FNKIQQILKQ IQDKQIVSTY GIENQSQKKN HKNYKNQNAT
FSKKDLIHLI RDSLKESKIL KVKMDSLKVK RRFPFNLEQA LKNSKQRTLY
IDFLPPKCSK QTLVSIFGNF RIININLPLQ KNSQLCQGFA FIEFFSEEEA
NQALITKNSS IPKELILLTE KKIGQGSIRI ITYKKWQEEK QSFKELSKNQ
NEQKNKNMNQ SRKASDEFVS IDVEIKQNCL IKIINIPQGT LKAEVVLAVR
HLGYEFYCDY IDENSNQINS NKISLSTQQQ NTAQCSNIQI ENNLIQQDQH
PQLNDLLKEG QAMIRFQNSD EQRLAIQKLL NHNNNKLQIE IRGQICDVIS
TIPEDEEKNY WNYIKFKKNE FRKFFFMKKQ QKKQNITQNY NK

Figure 7. Mass spectrometry analysis of the gel excised "deltaN p65" band from the previous SDS-page gel. (A) No protease digestion (B) Chymotrypsin digestion. The C-terminal fragment found with no protease digestion suggests our protein contains an N-terminal deletion.

3.4 References

Asai, T., Zaporojets, D., Squires, C., and Squires, C.L. (1999a). An Escherichia coli strain with all chromosomal rRNA operons inactivated: complete exchange of rRNA genes between bacteria. Proc. Natl. Acad. Sci. U.S.a. *96*, 1971–1976.

Asai, T., Condon, C., Voulgaris, J., Zaporojets, D., Shen, B., Al-Omar, M., Squires, C., and Squires, C.L. (1999b). Construction and Initial Characterization of Escherichia coli Strains with Few or No Intact Chromosomal rRNA Operons. The Journal of Bacteriology *181*, 3803.

Brosius, J., Ullrich, A., Raker, M.A., Gray, A., Dull, T.J., Gutell, R.R., and Noller, H.F. (1981). Construction and fine mapping of recombinant plasmids containing the rrnB ribosomal RNA operon of E. coli. Plasmid *6*, 112–118.

Chen, Y., Fender, J., Legassie, J.D., Jarstfer, M.B., Bryan, T.M., and Varani, G. (2006). Structure of stem-loop IV of Tetrahymena telomerase RNA. The EMBO Journal *25*, 3156–3166.

Chruszcz, M., Domagalski, M., Osinski, T., Wlodawer, A., and Minor, W. (2010). Unmet challenges of structural genomics. Curr. Opin. Struct. Biol. *20*, 587–597.

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

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

Jiang, J., Miracco, E.J., Hong, K., Eckert, B., Chan, H., Cash, D.D., Min, B., Zhou, Z.H., Collins, K., and Feigon, J. (2013). The architecture of Tetrahymena telomerase holoenzyme. Nature *496*, 187–192.

Miller, M.C. (2002). Telomerase recognizes its template by using an adjacent RNA motif. Proceedings of the National Academy of Sciences *99*, 6585–6590.

Mitchell, M., Gillis, A., Futahashi, M., Fujiwara, H., and Skordalakes, E. (2010). Structural basis for telomerase catalytic subunit TERT binding to RNA template and telomeric DNA. Nat. Struct. Mol. Biol. *17*, 513–518.

Richards, R.J. (2006a). Structural study of elements of Tetrahymena telomerase RNA stem-loop IV domain important for function. Rna *12*, 1475–1485.

Richards, R.J. (2006b). Structure of the Tetrahymena thermophila telomerase RNA helix II template boundary element. Nucleic Acids Res. *34*, 816–825.

Rouda, S., and Skordalakes, E. (2007). Structure of the RNA-Binding Domain of Telomerase: Implications for RNA Recognition and Binding. Structure *15*, 1403–1412.

Singh, M., Wang, Z., Koo, B.-K., Patel, A., Cascio, D., Collins, K., and Feigon, J. (2012). Structural Basis for Telomerase RNA Recognition and RNP Assembly by the Holoenzyme La Family Protein p65. Mol. Cell *47*, 16–26.

Spahn, C.M., Grassucci, R.A., Penczek, P., and Frank, J. (1999). Direct three-dimensional localization and positive identification of RNA helices within the ribosome by means of genetic tagging and cyro-electron microscopy. Structure *7*, 1567–1573.

Yokoyama, T., and Suzuki, T. (2008). Ribosomal RNAs are tolerant toward genetic insertions: evolutionary origin of the expansion segments. Nucleic Acids Res. *36*, 3539–3551.

Chapter 4

Crystallization of a Minimal TERT-RBD and TER Complex

4.0 Introduction

The reverse transcriptase catalytic subunit of telomerase (TERT) contains four domains: N-terminal (TEN), RNA binding domain (RBD), reverse transcriptase domain (RT), and C-terminal extension (CTE) (Fig. 1) (Hengesbach et al., 2011). The RNA component of telomerase (TER) contains a template boundary element (TBE) that defines the region reverse transcribed by TERT (Fig. 1) (Lai, 2002; Lin et al., 2004). This is a site of strong RBD interaction that prevents TERT from transcribing past the template, defining a central feature of the mechanism of telomere synthesis.

RBD in *Tetrahymena thermophila* contains several conserved RNA binding motifs, binds to TER with high affinity, and is the main telomerase protein component that interacts with TER (Bosoy, 2002; Bryan et al., 2000; Lai et al., 2001; Miller, 2000; O'Connor, 2005). The crystal structure of a large C-terminal fragment of RBD showed the CP motif is adjacent to the T motif in an electropositive groove, suggesting that this may be the site of interaction with TBE (Rouda and Skordalakes, 2007). However, the structure does not include RNA or the third domain CP2, so it's not currently possible to define which motifs interact with which RNA domains.

Previous studies have shown several mutations in the CP2 motif can cause defects in telomerase activity and template definition (Lai, 2002; Miller, 2000). Additionally, by using Fe(II)-EDTA site-directed hydroxyl radical probing and electrophoretic mobility shift assays (EMSA), it was demonstrated the CP2 motif is a critical site of TBE interaction and is important for maintaining template definition (Akiyama et al., 2013).

Unfortunately, no high-resolution structural information exists for the CP2 motif. This is due to a required proteolysis step of RBD prior to crystallization, cleaving a short portion of the N-terminus that includes the CP2 motif. To better understand how template definition is defined, and thus construct a more complete understanding of how telomerase functions, the following work attempts to obtain a high-resolution crystal structure of a minimal CP2-containing RBD and TER complex.

4.1 Results

4.1.1 Minimal TERT-RBD Protein Construction

Attempts to express full length RBD in *E. coli* led to poor yield, with the majority of protein resulting in the insoluble portion, which is not ideal for the high quantity required for crystallization trails. Expressing a RBD construct lacking the N-terminus may have resulted in better yields and crystal formation, explaining the reason for the shorter RBD construct for crystallization. Coincidently, the cleaved RBD in the crystal structure had poor binding affinity to TER, suggesting CP2 may play a critical role in TER binding (Akiyama et al., 2013).

To find a compromise between solubility and sufficient TER binding, several RBD constructs were created. Spaced evenly from amino acid 200 up to the CP2 motif, 226, six proteins were constructed, expressed, and their solubility analyzed (Fig. 2). While the construct beginning at 211 (RBD211) appeared to have the highest solubility before purification, a higher yield of RBD217 was obtained after nickel purification and was chosen for binding and crystallization experiments (Fig. 3).

4.1.2 Minimal TER Construction

Previous studies have identified the bases at stem II as being critical for proper boundary definition (Lai, 2002; O'Connor, 2005). Other work demonstrated RBD directly contacts these bases and created several minimal TBE constructs to more specifically probe CP2-TER interactions (Akiyama et al., 2013). Following these minimal constructs, a minimal stem II construct consisting of residues 13-42 of TER was created (sII) (Fig. 4). Binding studies demonstrated that stem I may also participate in TERT binding, so a second construct consisting of Stem I and II, with a GAAA tetraloop inserted for stem IV, was also created (sI-II) (O'Connor, 2005). When constructs were analyzed by size exclusion chromatography, sI-II was sized as predicted, but sII appeared to form a higher order complex (Fig. 5). While binding affinities were lower than full length TER, both constructs were able to bind full length RBD (unpublished data). With the goal of producing a crystal structure that would show basic interactions of TER with CP2, these constructs were used for further binding and crystallization experiments. Using a minimal TER would also improve crystal formation and electron density by minimizing flexible domains.

4.1.3 Binding and Crystallization of Minimal Protein and RNA Constructs

To assess the binding affinity of RBD217 to each TER construct, the protein and RNA were incubated in a 1:1 molar ratio and analyzed using a gel filtration column. While it appears binding is occurring with both constructs, RBD217 has a slightly higher affinity for sI-II, however not all RNA appeared to be bound (Fig. 6). Increasing the protein to RNA ratio to 2:1 was able to completely shift free RNA to the bound state for sI-II, but had little affect on sII (Fig. 7). Because of this, only the sI-II complex was used for crystallization trails. The shifted RNA peak was conservatively collected and concentrated to ~3.0 mg/mL. Initial screening trays were laid using sitting drop method, placed at 4°C, and possible crystals, visible as small spiked balls, appeared within several days (Fig. 8).

4.2 Discussion

This work lays the foundation for obtaining a high-resolution structure of the key interactions between RBD and TER. Not only was a minimal CP2-containing RBD construct designed that was soluble and bound to a minimal TER

construct, but preliminary crystals were successfully obtained with initial screening.

Not surprisingly, RBD217 did not bind as well to sII, suggesting stem I contributes to RBD binding, agreeing with previous studies (O'Connor, 2005). Interestingly, a 2:1 protein to RNA ratio is required to completely shift the sI-II RNA peak. It's difficult to say whether half the protein is incompetent to bind or the complex contains two RBD217 proteins. If the former, modifying protein expression and purification may show a binding ratio closer to 1:1. Cleaving the his-tag prior to binding may also improve binding and/or crystal formation. It's also possible that we are simply under or overestimating one of the concentrations. Additionally, using a narrower range sizing column would increase the separation of the bound and unbound peaks, resulting in a more homogenous complex for crystal formation.

Only initial trays were laid and no optimization was attempted. Continuing to optimize with finer well conditions, more initial well conditions, different temperatures, and using the initial crystals for seeding new trays may prove fruitful. The fact that crystals were obtained with initial screens is very promising and warrants further exploration.

CP2 is only conserved among ciliates, but CP2 could co-evolve with the changing sequence of TBE, explaining the divergence a possible TBE binding motif (Lin et al., 2004). A region in vertebrate TERT, the vertebrate-specific

region (VSR), is in a similar position to CP2 in ciliates (Moriarty et al., 2002), and could be analogous to CP2 in that in binds TBE. Obtaining a structure of how *T. thermophila* RBD interacts with TBE would lead to a better mechanistic understanding of the unique reverse transcriptase activity of telomerase, and could be extended to telomerase in other species.

4.3 Methods

4.3.1 Minimal TERT-RBD Construction

Linear plasmids of each construct were produced by performing PCR on a plasmid containing *T. thermophila* RBD fused to an N-terminal his-tag using custom primers that would exclude the necessary portion of RBD. The linear plasmid products were treated with T4 polynucleotide kinase (NEB) and T4 DNA ligase to form circular plasmids that were transformed into 10-beta *E. coli* cells. This resulted in plasmids containing RBD200, RBD205, RBD211, RBD217, RBD222, and RBD226, with the second number denoting the first starting amino acid each construct begins in respect to full length TERT.

4.3.2 Minimal TERT-RBD Expression and Purification

Protein constructs were expressed using BL21 DE3 *E. coli* cells. After reaching mid-log growth phase at 37°C in Terrific Broth media, the cells were placed on ice for 1 hour. Expression was performed using 0.1 mM IPTG at 15°C and cells harvested after overnight growth (~16 hours). The cells were resuspended in buffer A (20 mM Tris pH 7.8, 200 mM NaCl, 1 mM $MgCl_2$, 40 mM imidazole, and 1 mM DTT). 1 mM PMSF was added just before passing the cells twice through a cell disruptor at 15,000 psi. 120 U of TURBO DNase (Ambion) was added to the cell lysate and incubated on ice for 15 minutes. The cell lysate was cleared twice by centrifugation at ~36,000 x g for 30 minutes each at 4°C and passed over Ni Sepharose 6 Fast Flow resin (GE Life Sciences) by gravity flow. The resin was thoroughly washed with buffer A that contained 400 mM NaCl until no protein was detected by Bradford assay. After washing with a column volume of buffer A , the protein was eluted with buffer A containing 200 mM NaCl and 500 mM imidazole until no protein was detected. The elution was diluted to 100 mL with buffer B (20 mM Tris pH 7.8, 200 mM NaCl, 1 mM $MgCl_2$, and 1 mM DTT), passed over a heparin ion exchange column (GE Life Sciences), and protein eluted with a 60 minute buffer B gradient to 1 M NaCl. The protein was concentrated with a spin concentrator (Amicon) and passed over a superdex 200 size-exclusion column at 4°C (GE Life Sciences) in buffer C (20 mM

Tris pH 8.0, 200 mM NaCl, 1 mM $MgCl_2$, 1 mM DTT, and 10% glycerol) 10 mg at a time maximum. The protein was aliquoted, flash frozen, and stored at -70°C.

4.3.3 Minimal TER Construction

The stem II construct consisted of residues 13-42 of TER, while the stem II+IV construct contained residues 1-42, a GAAA tetraloop for stem IV, and residues 101-107. TER constructs were transcribed using T7 RNA polymerase (NEB), DNase treated (Ambion), phosphatase treated (NEB), and purified using a superdex 200 size-exclusion column (GE Life Sciences). The RNA was end-labeled using phosphonucleotide kinase (NEB) with ^{32}P-labeled γ-ATP (Perkin-Elmer) and PAGE purified.

4.3.4 Binding Assay

Binding assays were performed by combining either a 1:1 or 2:1 protein to RNA molar ratio for a total volume of 100 μL, supplementing the mixture with NaCl to 200 mM, and letting rest at RT for 10 minutes. The complex was loaded

on a Superdex 200 size-exclusion column (GE Life Sciences) using buffer C, and analyzed at wavelengths 215, 254, and 280 nm.

4.3.5 Crystallization

Purification of the 2:1 RBD217 to sI-II molar ratio complex was performed in similar fashion as with the binding assay except for the following deviations: 20 µL of sI-II was added to 600 µL 4.52 mg/mL RBD217, left at RT for 2-3 minutes, 280 µL of sI-II added, left at RT for an additional 10 minutes, passed through the sizing column using buffer C with 10 mM DTT, and shifted RNA peak conservatively collected. The purified complex was concentrated to ~3 mg/mL using a spin concentrator (Millipore) and crystallization was performed by sitting drop vapor diffusion using a Crystal Phoenix automatic robot (Art Robbins Instruments) with 96 well plates. The trays were placed at 4°C, and at a 1:1 complex to buffer ratio with a total drop volume of 0.4 µL. Small crystals appeared in 1-2 weeks in the following conditions:

1) 0.2 M Sodium chloride

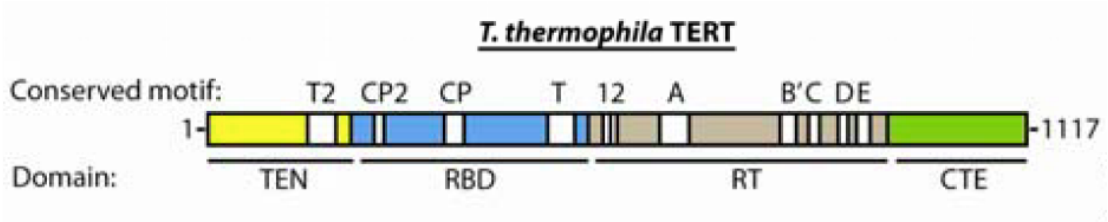
   0.1 M Sodium cacodylate pH 6.5

   2.0 M Ammonium sulfate

2) 1.0 M Sodium citrate

0.1 M Sodium cacodylate pH 6.5

A



B

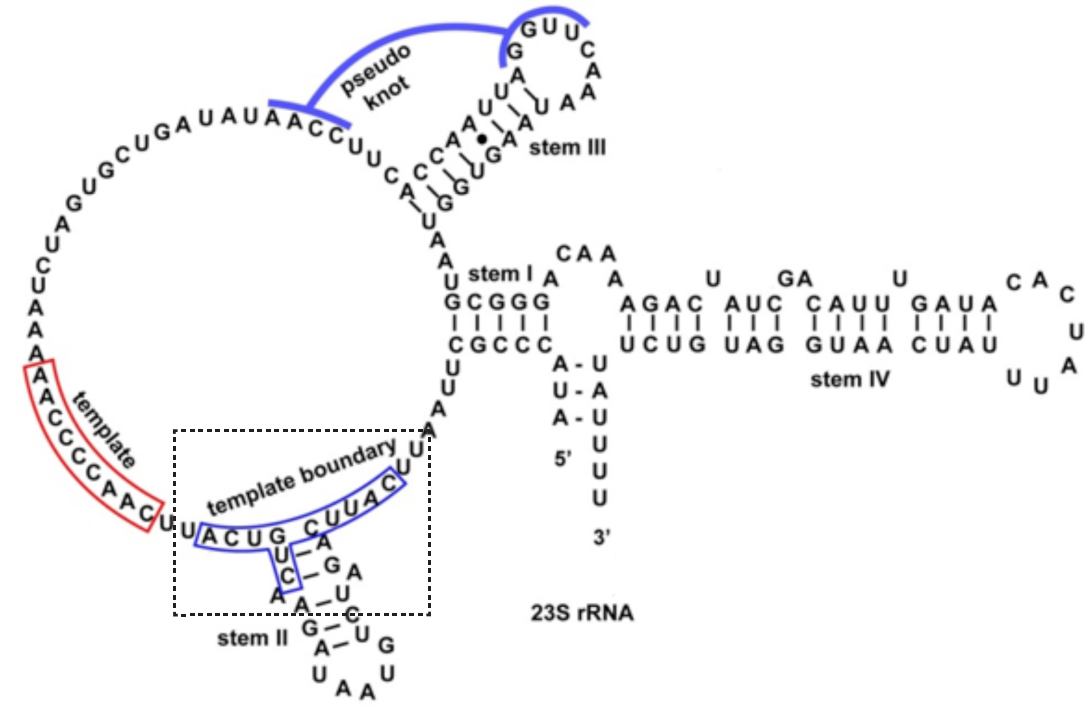


Figure 1. (A) Schematic outlining conserved motifs and domains of TERT. Note the CP2-motif at the N-terminal end of RBD. (B) Schematic of TER with the TBE outlined.

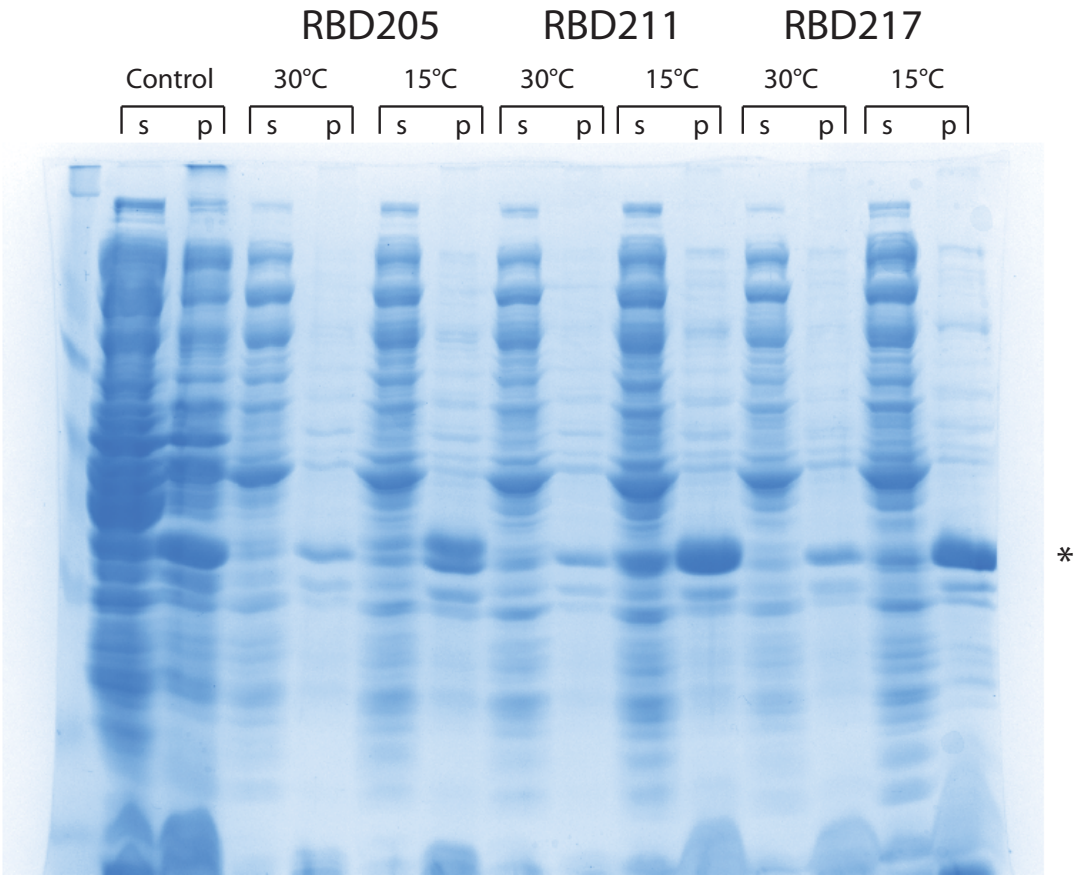


Figure 2. SDS-page gel of three RBD constructs. Proteins were expressed at either 30°C or 15°C, cells lysed, and supernatant (s) and pellet (p) analyzed by SDS-page. The asterisk marks the RBD constructs. Note that while RBD111 had the best solubility, RBD217 had the best yield after nickel purification.

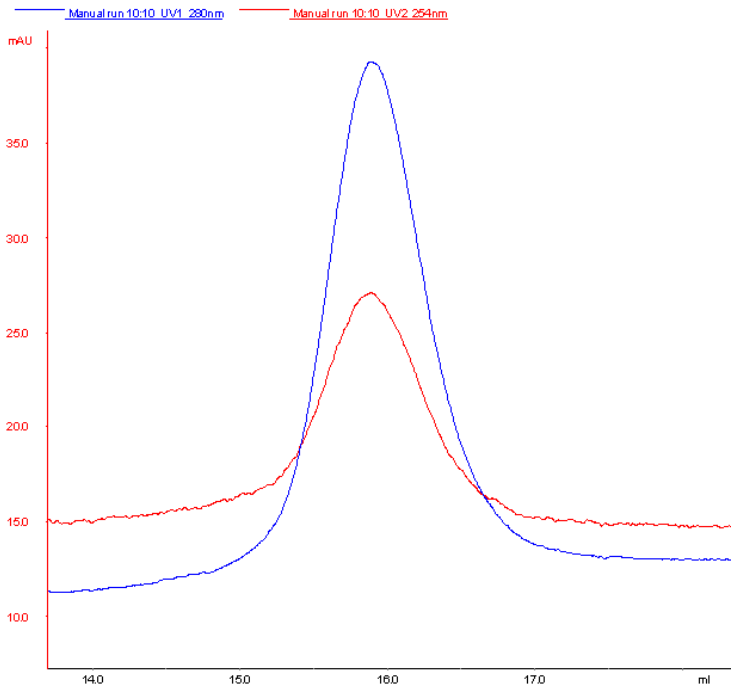
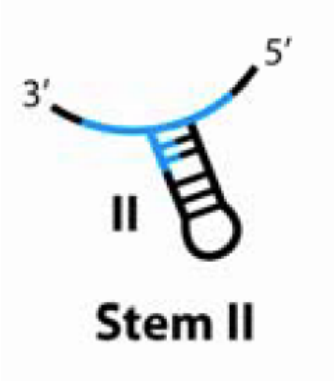


Figure 3. Trace of nickel and ion exchange purified RBD217 using a superdex 200 size exclusion column (GE Life Sciences).

A



B

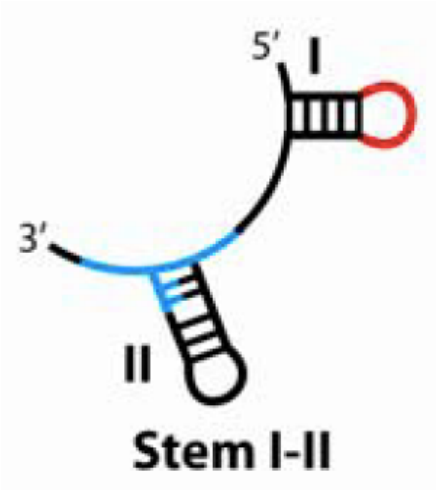
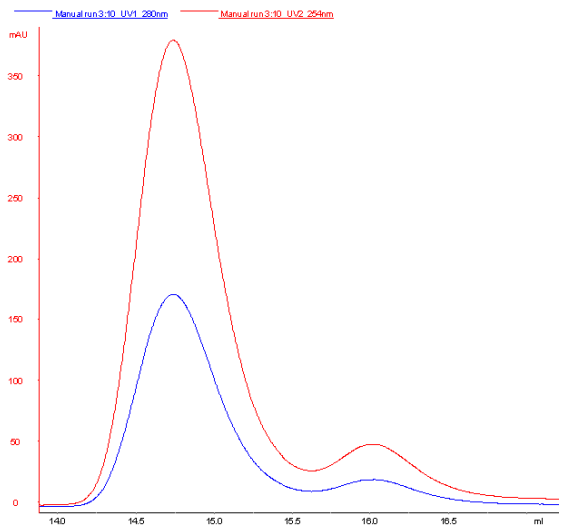


Figure 4. Minimal TER constructs. (A) Stem II construct (sII) consists of stem II and the TBE (blue). (B) Stem I-II construct (sI-II) consists of stem II, TBE, and stem I, with a GAAA tetraloop substitution for stem IV.

A



B

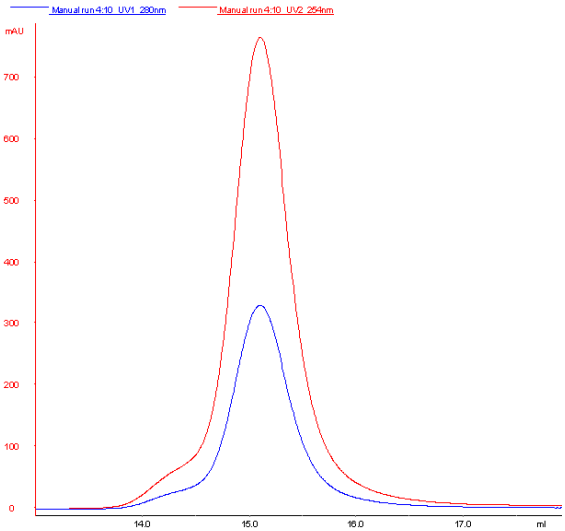
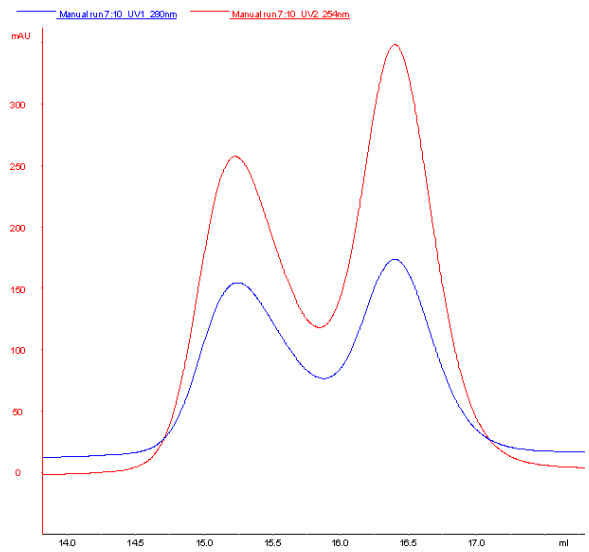


Figure 5. Size exclusion traces of TER constructs (A) sII and (B) sI-II alone. SII appeared to be forming a higher order complex.

A



B

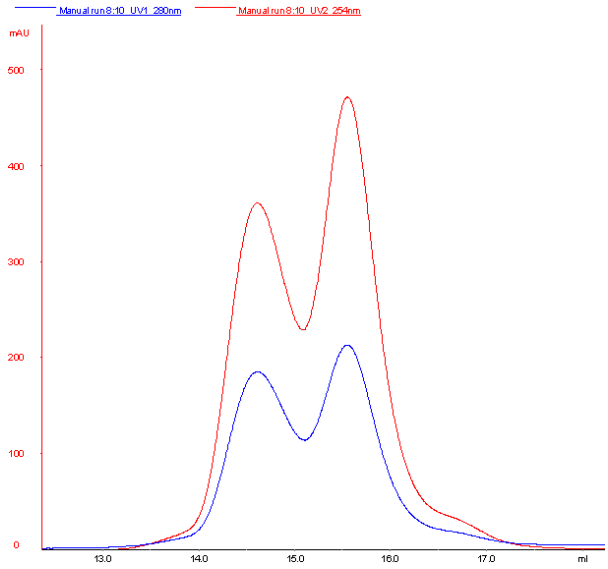
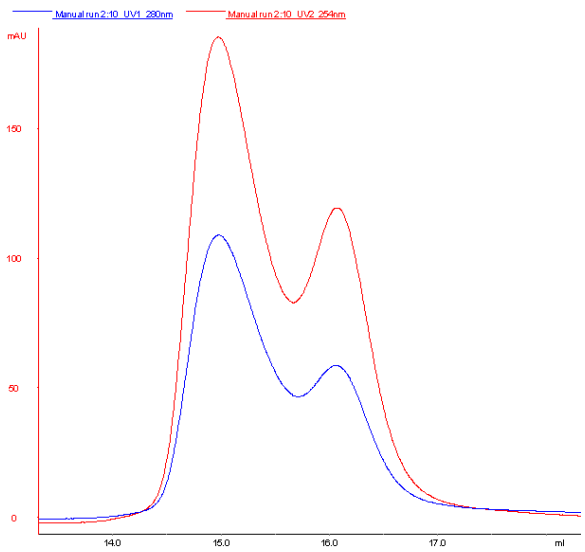


Figure 6. Size exclusion traces of 1:1 RBD217 to TER construct binding. Both (A) sII and (B) sI-II appeared to be half bound.

A



B

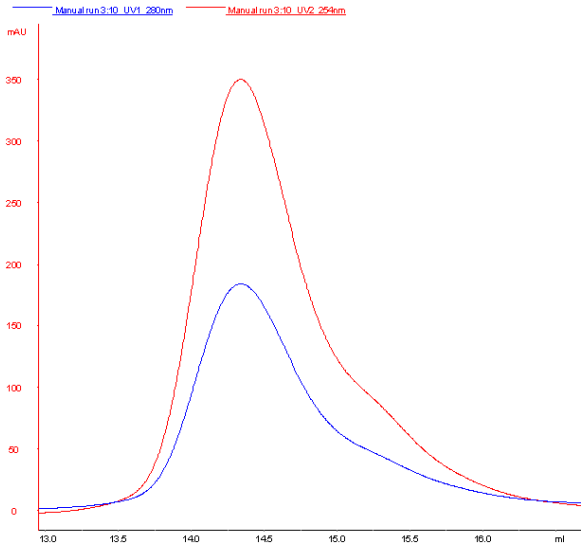
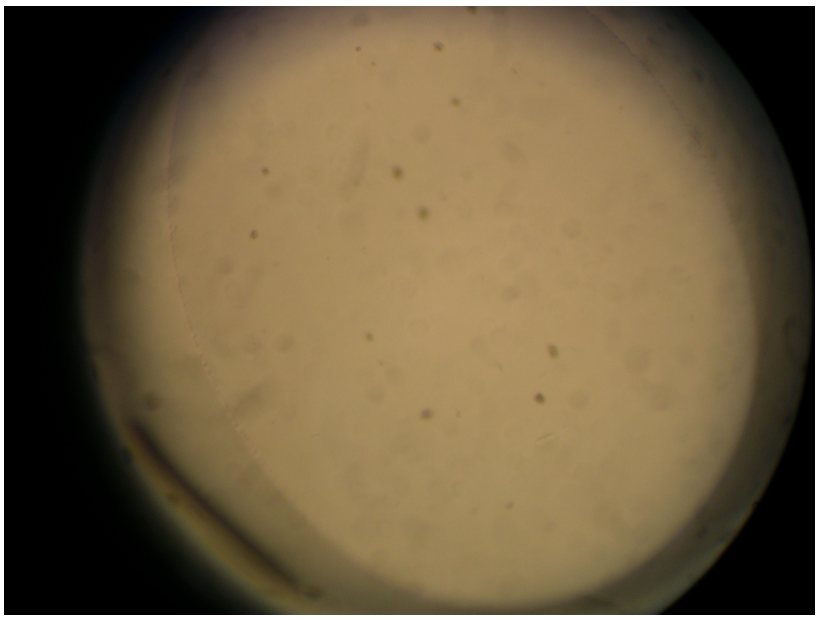


Figure 7. Size exclusion traces of 2:1 RBD217 to TER construct binding. (A) SII binding slightly improved while (B) sI-II appeared to be nearly completely bound.

A



B

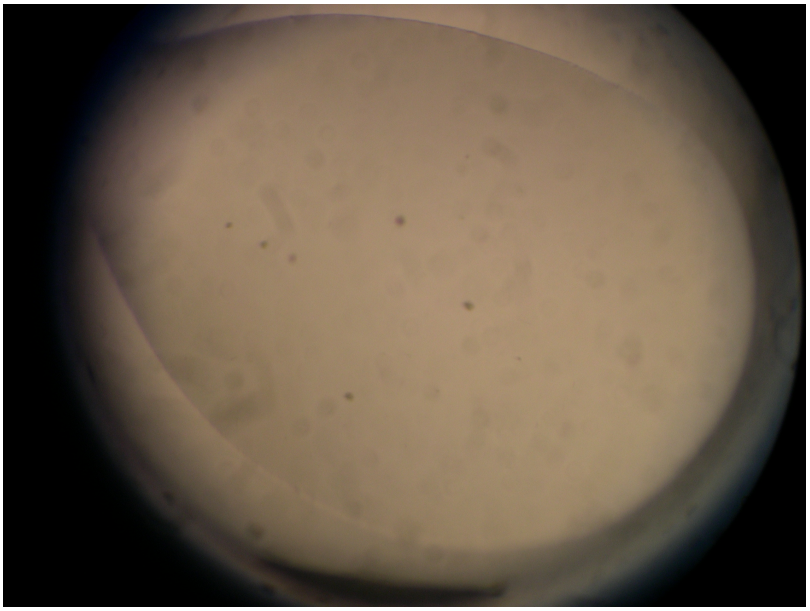


Figure 8. Crystals formed at 4°C in conditions (A) 0.2 M Sodium chloride, 0.1 M Sodium cacodylate pH 6.5, 2.0 M Ammonium sulfate and (B) 1.0 M Sodium citrate, 0.1 M Sodium cacodylate pH 6.5. They appeared within a few days as small spiked balls.

4.4 References

Akiyama, B.M., Gomez, A., and Stone, M.D. (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.

Bosoy, D. (2002). Conserved N-terminal Motifs of Telomerase Reverse Transcriptase Required for Ribonucleoprotein Assembly in Vivo. Journal of Biological Chemistry *278*, 3882–3890.

Bryan, T.M., Goodrich, K.J., and Cech, T.R. (2000). Telomerase RNA Bound by Protein Motifs Specific to Telomerase Reverse Transcriptase. Mol. Cell *6*, 493–499.

Hengesbach, M., Akiyama, B.M., and Stone, M.D. (2011). Single-molecule analysis of telomerase structure and function. Curr Opin Chem Biol *15*, 845–852.

Lai, C.K. (2002). Template boundary definition in Tetrahymena telomerase. Genes & Development *16*, 415–420.

Lai, C.K., Mitchell, J.R., and Collins, K. (2001). RNA Binding Domain of Telomerase Reverse Transcriptase. Mol. Cell. Biol. *21*, 990–1000.

Lin, J., Ly, H., Hussain, A., Abraham, M., Pearl, S., Tzfati, Y., Parslow, T.G., and Blackburn, E.H. (2004). A universal telomerase RNA core structure includes structured motifs required for binding the telomerase reverse transcriptase protein. Proc. Natl. Acad. Sci. U.S.a. *101*, 14713–14718.

Miller, M.C. (2000). Template definition by Tetrahymena telomerase reverse transcriptase. The EMBO Journal *19*, 4412–4422.

Moriarty, T.J., Huard, S., Dupuis, S., and Autexier, C. (2002). Functional Multimerization of Human Telomerase Requires an RNA Interaction Domain in the N Terminus of the Catalytic Subunit. Mol. Cell. Biol. *22*, 1253–1265.

O'Connor, C.M. (2005). Two Purified Domains of Telomerase Reverse Transcriptase Reconstitute Sequence-specific Interactions with RNA. J. Biol. Chem. *280*, 17533–17539.

Rouda, S., and Skordalakes, E. (2007). Structure of the RNA-Binding Domain of Telomerase: Implications for RNA Recognition and Binding. Structure *15*, 1403–1412.