DUPLICATION AND DIVERSIFICATION OF *Arabidopsis thaliana* TELOMERASE RNP COMPONENTS

A Dissertation

by

CATHERINE CIFUENTES ROJAS

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

December 2010

Major Subject: Genetics

Duplication and Diversification of *Arabidopsis thaliana* Telomerase RNP Components

DUPLICATION AND DIVERSIFICATION OF *Arabidopsis thaliana* TELOMERASE RNP COMPONENTS

A Dissertation

by

CATHERINE CIFUENTES ROJAS

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Approved by:

Chair of Committee,	Dorothy E. Shippen
Committee Members,	Gregory Reinhart
	Julien Leibowitz
	Xiuren Zhang
Chair of Genetics,	Craig J. Coates

December 2010

Major Subject: Genetics

ABSTRACT

Duplication and Diversification of *Arabidopsis thaliana* Telomerase RNP Components.

(December 2010)

Catherine Cifuentes Rojas, B.Sc., Universidad del Valle; M.Sc., Pontificia Universidad Javeriana

Chair of Advisory Committee: Dr. Dorothy E. Shippen

Telomerase is a highly regulated ribonucleoprotein complex that stabilizes eukaryotic genomes by replenishing telomeric repeats on chromosome ends. Defects in telomerase RNP components involving the catalytic subunit TERT or the RNA template TER lead to stem cell-related diseases such as dyskeratosis congenita and idiopathic pulmonary fibrosis, while inappropriate telomerase expression is a rate-limiting step in carcinogenesis. In this study we report the discovery of a novel negative regulatory mechanism for telomerase that stems from duplication and diversification of key components of the telomerase RNP in the flowering plant *Arabidopsis thaliana*.

We show that *Arabidopsis* encodes three distinct TERs: TER1, TER2 and a processed form of TER2 termed $TER2_S$. Although all three RNAs can serve as templates for telomerase *in vitro, in vivo* they have different expression patterns, assemble into distinct RNPs with different protein binding partners, and play opposing roles in telomere maintenance. The TER1 RNP is analogous to the

telomerase enzyme previously described in other eukaryotes, but the TER2 RNP is a negative regulator of telomerase activity and telomere maintenance *in vivo*.

Furthermore, we demonstrate that the Protection Of Telomeres (POT1) paralogs in *Arabidopsis* (POT1a, POT1b and POT1c) are novel TER binding proteins. This finding is striking because in yeast and vertebrates, POT1 is an essential component of the telomere capping complex and functions to distinguish the chromosome terminus from a double-strand break. Thus, our data argue that *Arabidopsis* POT1 proteins have migrated off of the chromosome terminus and onto the telomerase RNP, indicating that duplication and diversification of *Arabidopsis* telomerase may be the end result of the co-evolution of the TER and POT1 RNP components.

Additionally, given the dire consequences of misregulating telomerase in human cells, our discovery of a novel negative regulatory mechanism for telomerase in plants strongly suggests that additional modes of telomerase control remain to be elucidated in vertebrates.

DEDICATION

To my family, my whole source of motivation and strength. My parents William and Mariana have always backed me up in all the stages of my life and have supported all my crazy ideas, even those with which they did not agree. My brother and unconditional partner in mischief during my whole life, Andres, and my sister, Sirley, have always been there for me. And Alfredo, the love of my life, who has held my hand ever since we met, has helped me to stand up every time that I've fallen with patience and love. Without their unconditional support, I could not have accomplished this goal.

A mi familia, toda mi fuente de fuerza y motivación. Mis padres William y Mariana siempre me han respaldado durante todas las etapas de mi vida y han apoyado todas mis ideas, inclusive aquellas con las que no estaban de acuerdo. Mi hermano Andres, compañero incondicional de travesuras durante toda mi vida y mi hermana Sirley siempre han estado allí para mi. Y Alfredo, el amor de mi vida, quien a sostenido mi mano desde que nos conocimos y con paciencia y amor me ha ayudado a levantarme cada vez que he caído. Sin su apoyo incondicional yo no hubiese podido alcanzar esta meta.

ACKNOWLEDGEMENTS

First, I want to thank my mentor, Dr. Dorothy Shippen for giving me the opportunity to join her lab a few years back and for her guidance throughout my graduate life. Additionally, I am grateful for her constant motivation and support; for teaching me to be creative and proactive, to enjoy science from all perspectives and particularly to set high goals in life and to work hard to accomplish them. Because all this, I believe a better mentor could not be found.

I want to thank my committee members, Dr. Gregory Reinhart, Dr. Julien Leibowitz and Dr. Xiuren Zhang for their scientific advice, as well as for their guidance and support. I also want to thank Dr. Geoffrey Kappler for his valuable comments during our joint group meetings and for his help to prepare the presentations for my job interviews. I also appreciate the help and advice of Dr. Cheng Kao.

I am also grateful to Dr. Kalpana Kannan, Dr. Yulia Surovtseva, Dr. Xiangyu Song and Dr. Eugene Shakirov for their friendship, good advice and help when I joined and throughout the years in the Shippen Lab. I am thankful to Dr. Mark Beilstein for providing his bioinformatics expertise and figures 37-39, as well as for his advice and support. Thanks also go to my colleagues form the Shippen lab: Andrew Nelson, who is always willing to help, Kara Boltz, Katherine Leehy, Dr. Jung Ro Lee and Lin Tseng, for their friendship and collaboration during all these years.

We thank Yehuda Tzfati and Mark Beilstein for insightful comments on the manuscript and Jeff Kapler, Alfredo Hernandez, Feng Qiao and members of the Shippen lab helpful comments throughout the study. We are also grateful to Jung Ro Lee for providing the α-dyskerin antibody. This work was supported by NIH GM065383 and NSF MCB-0843399 to D.E.S.

Thanks to the Faculty of Genetics and the Department of Biochemistry and Biophysics of Texas A&M University for contributing to my scientific and personal growth. I also want to extend my gratitude to Tilie Rausch, Juanita Withem, Sherry Coronado, Melissa Kay, Julia Williams and other members of the staff for their willingness to help solve administrative issues and for their patience with the deadlines.

Finally, thanks to my family for their unconditional encouragement and support. I extend special thanks to my husband, Alfredo Hernandez, for his helpful suggestions for my experiments and manuscripts, as for being there for me every step along the road.

TABLE OF CONTENTS

CHAPTER Page

LIST OF FIGURES

Page

Page

CHAPTER I

INTRODUCTION

The history of telomeres is closely related with that of eukaryotic cellular aging, dating to the late 1800s. In 1881 the German biologist August Weismann speculated that "death takes place because a used up tissue cannot renew itself forever and because a capacity for increase by means of cell division is not everlasting but finite" (1).

In 1908, Alexis Carrel (Nobel Prize 1912) became interested in the growth of cells in culture and established a culture of chick heart fibroblast cells, which he cultured for 34 years in his lab. The result of this work challenged Weismann's model and led to the general acceptance that vertebrate cells can divide indefinitely in culture and that "the lack of continuous cell replication was due to ignorance on how best to cultivate the cells". Since individual cells were "immortal", Carrel came to the conclusion that aging is "an attribute of the multicellular body as a whole" (reviewed in (2).

In 1961, Carrel's notion of cell immortality was challenged by experiments published by Hayflick and Moorehead (3). They demonstrated that the finite replicative capacity of normal human fibroblasts corresponds to ageing at the cellular level. Fibroblast cultures derived from human skin would divide up to the

---

This dissertation follows the style of the journal *Proceedings of the National Academy of Sciences of the United States of America.*

40^{th} to 50^{th} generation, and then stop and “undergo senescence”. Further studies showed that cells from older people underwent fewer divisions than cells from younger people, suggesting that it was the total number of divisions since birth and not the total divisions in culture, that was important (4). This model is called ‘‘Hayflick limit’’ and is the most accepted, as Carrel’s experiments were never reproduced. One of several important questions raised from this work was what tells cells to stop dividing?

Work done by the American geneticists Herman Muller (Nobel Prize 1946) and Barbara McClintock (Nobel Prize 1983) in the late 1930s, in *Drosophila* and *Zea mays*, respectively, would be the start of a series of discoveries that would answer this and other related questions. Muller was studying the effects of X-rays on *Drosophila* chromosomes (5). He isolated several types of chromosome rearrangements, but he could not isolate a terminal deletion, a chromosome missing the end. Muller concluded from these results that the end of the chromosome is a distinct structure essential for chromosome stability. He named it telomere (from the Greek *telos* for end and *meros* for part) (5).

Additionally, Barbara McClintock found that broken meiotic chromosomes would frequently fuse end to end resulting in dicentric chromosomes in maize which are unstable (6). Results from this work allowed her to postulate that chromosome ends protect the terminus from being recognized as double-strand DNA breaks (7).

The End Replication Problem

James Watson (Nobel Prize 1962) proposed that an end replication problem would result from a fundamental limitation of DNA polymerases. The conventional DNA polymerases cannot carry out de novo DNA synthesis. Instead they require a template and a pre-existing 3'-hydroxyl group (3'-OH) provided by an RNA primer to initiate synthesis. This results in the direction of the synthesis being constrained to the 5'→3' direction. Removal of RNA primers generate Okazaki fragments, which then provide the 3'OH source to fill-in the empty spaces left by the RNA primers. Watson hypothesized that when the polymerase reached the end of a linear DNA, the space left by the last RNA primer could not be filled-in, and then there would be a problem to complete replication: "the end replication problem" (Fig 1).

Linear phage genomes can overcome this problem by joining multiple genomes before replication, thus reducing the number of actual ends and minimizing the damage by incomplete replication (8). Independent studies by the Russian biologist Alexei Olovnikov resulted in a similar conclusion. He predicted a problem with the beginning of the DNA replication instead of the end (9). Olovnikov also suggested a link between cell senescence and telomeres. They proposed that the progressive shortening of telomeres might explain the ''Hayflick limit" theory: that cells can only replicate for a determined number of times. He concluded that a mechanism for maintaining telomere length should exist that would act as a molecular clock. If cells cannot compensate for the end

replication problem, chromosomes would shorten with each round of cell division leading to the end of the cellular lineage due to loss of distal genes. The next question was: how are telomeres maintained to allow an organism to survive for several generations?

Telomerase and Telomeres

In 1978, Elizabeth Blackburn (Nobel Prize 2009), joined Joseph Gall's lab at Yale as a postdoctoral fellow determine the DNA sequence of the ribosomal DNA (rDNA) from the ciliated protozoan *Tetrahymena*. She found that each *Tetrahymena* rDNA molecule terminated with 200–400 bp of TTGGGG tandem repeats and later it was determined that these repetitive termini were not exclusive of rDNA molecules (10).

Later and as result from the collaborative studies with Jack Szostak in yeast, it was established that these repetitive elements are highly conserved throughout evolution. They observed that the introduction in yeast of linear DNA fragments containing the *Tetrahymena* repetitive sequences at their termini resulted in stable maintenance in yeast cells as artificial minichormosomes. Normally, linear DNA is integrated into the genome by homologous recombination or subjected to exonucleolytic degradation. Interestingly, after cloning and sequencing of these DNA fragments they found tandem arrays of what they determined was the yeast telomeric repeat were added to the *Tetrahymena* repeats (11).

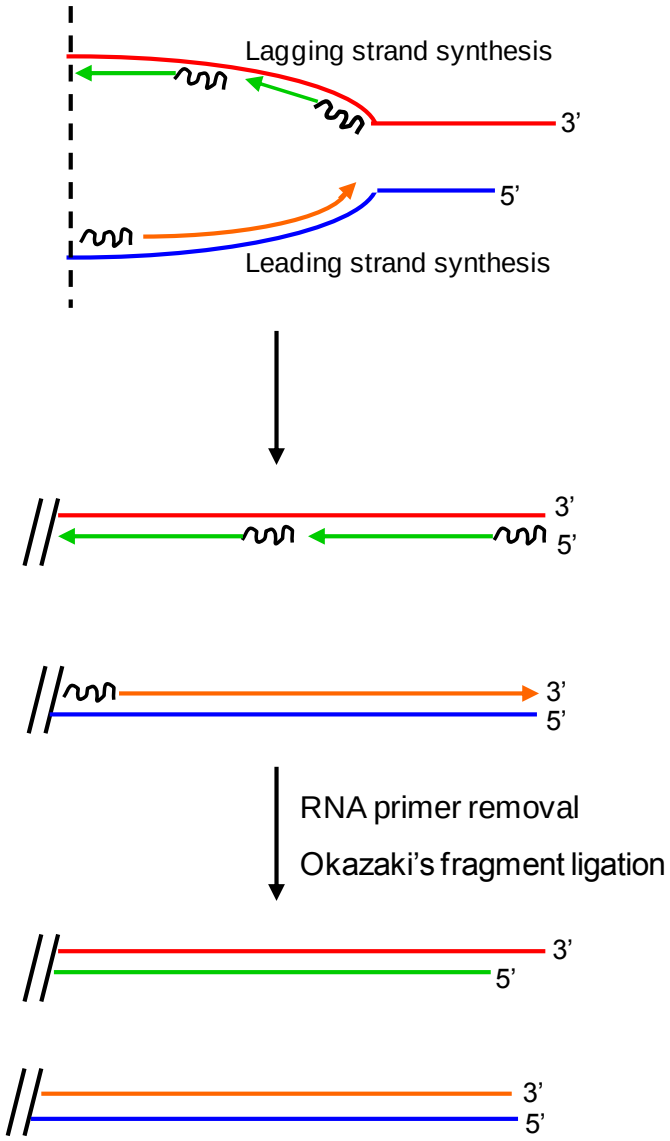


Figure 1. The end replication problem. Short RNA primers (wavy black lanes) are made by RNA primase for lagging-strand replication. These are then extended by DNA polymerase to form Okazaki fragments (Green line). The RNA primer removal leaves a single-stranded 3’-end G-overhangs in the lagging strand. After cell division a small piece of chromosome is not replicated.

It was proposed that species-specific tandem arrays of sequences confer to eukaryotic cells the capability to store their genetic material in form of linear chromosomes. But, by what mechanism were these sequences added at the chromosome end?

In 1984-85 Carol Greider, working as a graduate student in Elizabeth Blackburn's lab, found a biochemical activity in *Tetrahymena* extracts that extended telomeric sequences *in vitro* (12). Two years later she and Blackburn identified a telomere terminal transferase responsible for this activity. This activity was later renamed telomerase. Telomerase is a ribonucleoprotein complex whose protein and RNA subunits were essential for activity (13). Cloning of the RNA subunit showed that it contained CAACCCCAA, a sequence that was complementary to the telomeric repeat of *Tetrahymena* (TTGGGG), which it suggested that the RNA could be used as template for telomeric repeat synthesis (14). A year later several studies provided functional evidence for the RNA as template to direct reverse transcriptase activity of telomerase (15-16).

Further experiments resulted in the identification of yeast mutants that failed to add telomeres to a telomeric "seed" sequence introduced in a linearized and stabilized plasmid (17). The shorter telomere phenotype observed in these mutants (EST1, EST2, EST3 and EST4) was named EST for Ever-Shorter-Telomeres. These genes were subsequently shown to be essential components of the telomerase RNP or its docking site on the chromosome end (18). These data revealed a requirement for telomerase activity and pinpointed the cellular

consequences of telomerase inactivation: telomere shortening and subsequently cell senescence (17).

Altogether, these discoveries laid the ground work for an immense variety of research studies in the fields of telomere and stem cell biology as well as the application of this knowledge to human proliferative diseases like cancer among others.

Telomere Structure and Function

As discussed above, the DNA portion of telomeres is comprised of a tandem array of simple G-rich sequence repeat that vary slightly in different organisms. In vertebrates, telomeres consist of TTAGGG repeats (Moyzis 1988), while in *Arabidopsis* and most other plants the telomeric repeat is composed of TTTAGGG (19-22).

G-overhangs

Although many of the telomere repeats exist as dsDNA, the extreme 3' end of each chromosome extends beyond the 5' ends, forming a G-rich single-stranded protrusion called the G-overhang. G-overhangs were first evidenced in *Oxytricha* and *Euplotes* by sequencing of the 3' ends of macronuclear chromosomes (23). Later G-overhangs were found in budding yeast by non-denaturing in-gel Southern hybridization (24). They were also found in humans (25), chickens (26) as well and plants (27).

The G-overhang is the binding site for single-stranded telomere DNA binding proteins that are central players in telomere end protection (28-30). It is also the site where telomerase acts. The G-overhang has been proposed to play a regulatory role in yeast (31-32).

The proposed mechanism of G-overhang generation comes from studies performed in yeast. In budding yeast, G-overhangs are ~10–14 nt for the majority of the cell cycle, but their length increases transiently in late S phase when telomeres are replicated (24). C-strands are then filled in, presumably by replication or repair synthesis, yielding very short G-overhangs(24, 33). This cell cycle-dependent control of G-overhang formation is mediated by cyclin-dependent kinase 1 (CDK1) (34-35). Multiple nucleases and helicases have been identified in yeast that contribute to C-strand resection (32).

In contrast, mammalian telomeres have G-strand overhangs in the 150–350-nt range that are present throughout the cell cycle (25, 36-37). There is a length dependency of the cell type and a correlation with the length of the Okazaki fragments (25, 37). Several studies indicate that the G-stand overhangs are present on both telomeres, although their lengths may be quite different (38). The telomeres of plant chromosomes may also differ in overhang length; the longest are in the 20–30-nt range (27) and Riha personal communication).

Interestingly, transient C-overhangs have been detected during S phase in human cells and they are proposed to result from stalling of the replication fork on the leading-strand telomere (39). C-overhangs have been detected in

worms also and they appear to associate with distinct telomere binding proteins (40).

T-loops

G-overhangs fold back and insert into the double-stranded telomere DNA repeats to form higher order nucleoprotein structures that protect the ends called t-loops (41-42) (Fig 2). The first evidence of telomere folding was obtained from chromatin Immunoprecipitation (ChIP) experiments with the telomere protein Rap1 in budding yeast (43). The current models for yeast telomere structure suggest that yeast telomeres do not form t-loops, but rather a simple fold-back structure (44).

T-loops have been observed on chromosomes from a variety of organisms, including humans, mice, ciliates, trypanosomes and plants by electron microscopy (41, 45-46). Inter-strand psolaren cross-linking requires the precise positioning of T residues on opposing strands, arguing that the G-overhang is base-paired with an upstream region of the telomere (41). In addition, t-loop formation analyses using recombinant TRF2, a double stranded telomeric DNA binding protein indicated the requirement of G-overhangs for t-loop formation (42).

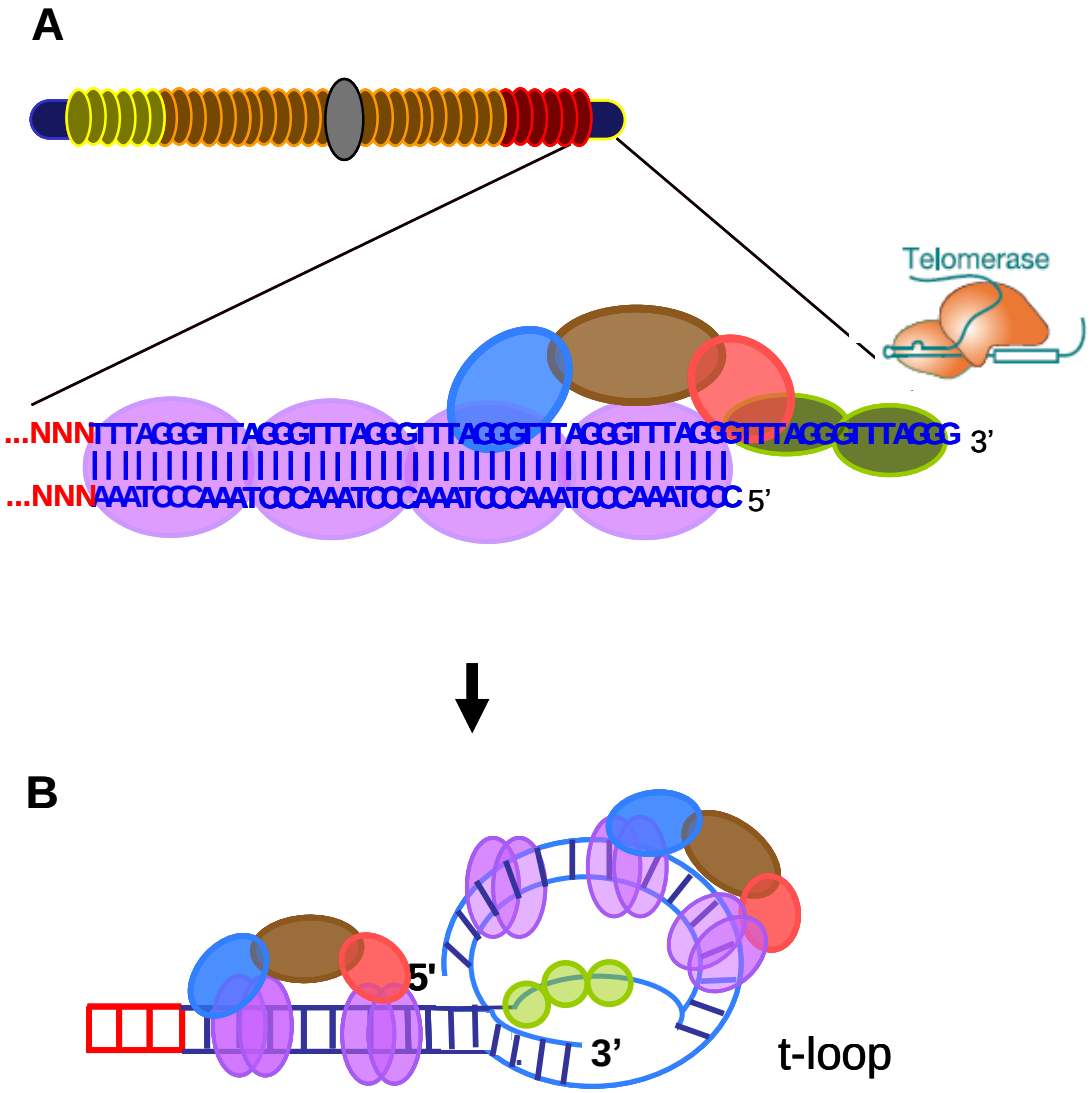


Figure 2. Telomere structure. (A) Telomeres are nucleoprotein structures at the end of eukaryotic chromosomes. The double-stranded and single-stranded DNA portions of the telomere are bound by ds-DNA and ss-DNA binding proteins. Other proteins interact with the proteins bound to the DNA forming complexes with different cellular functions. Telomerase must be accessible for telomere repeat addition. This reaction is regulated by the protein complexes at the telomere. (B) t-loop structure formed by the DNA-protein complexes at the end of the chromosomes.

Telomere Length Homeostasis

Telomere length is maintained within a strict range in each organism. In yeast telomeres are short (~300 bp), while in humans telomeres range from 10 to 15 kb in length (reviewed in(47)). Changes of telomere lengths are also observed in plants. *Arabidopsis* telomeres are 2-5 kb long (22), whereas in tobacco telomeres reach 150 kb (48). Telomere length homeostasis is achieved through a balance of forces that shorten and lengthen telomeres. Telomeres below a minimal size threshold "critically short" do not protect the chromosome end, resulting in activation of a DNA damage response and cell cycle arrest (reviewed in (49). The opposite situation, very long telomeres, also results in cell growth defects (50), which suggests that the cell takes very good care of maintaining the equilibrium of telomere shortening and lengthening *in vivo*.

Telomerase-Independent Telomere Maintenance

Telomere synthesis by telomerase is a conserved mechanism of chromosome end protection in eukaryotes (12, 51-52). However, some organisms use an alternative mechanism for this purpose. For example, the chromosomes of poxvirus are composed of highly A-T rich hairpins which contain extrahelical bases (reviewed in (53)). In *Drosophila melanogaster,* two families of non-LTR retrotransposons, HeT-A, TART, and TAHRE maintain telomere length by occasional transposition to the chromosome terminus (54-

56). This results in a complex array of repeats that form heterochromatin to protect the end of the chromosome (57).

Telomerase-independent Alternative Telomere Lenghtening, ALT was first discovered in yeast. Some yeast cells, termed “survivors” escape the cell senescence resulting from Ever-Shorter-Telomeres by maintaining telomeres through homologous recombination (58-59). Survivors can be classified as Type 1, which contain longer arrays of telomeric repeats alternated with subtelomeric sequences and which depend on Rad51 to amplify the subtelomeric sequences or Type 2, which only contain telomeric repeats and are dependent on Rad50 (60).

Although in most human cancers telomere lengthening is mediated by up-regulation of telomerase activity, in 10-15% of the cases this is achieved by ALT (reviewed in (61). Interestingly, observations made from cellular fusions between normal and ALT cells show that ALT is not observed in the resulting hybrid (62), thus ALT is not the result of a normal function and hence ALT is potentially a good target for anti-cancer therapies.

Telomere-Binding Proteins

Telomeric DNA is coated with specialized single-stranded and double-stranded DNA binding proteins that work in concert with the t-loop to protect the chromosome end and to prevent its recognition as a double-strand break. Telomere-specific proteins were identified from ciliates in the 1980s. Their high abundance and the stability of their interaction with telomeric-DNA facilitated their purification (63-64). Telomere-associated proteins contribute to the assembly of other protein complexes involved in telomere specific functions, as well as other cellular functions including DNA repair, DNA damage checkpoint and chromatin modification (65-66). There are two main classes of telomere-binding proteins: those that associate with single-stranded (ss) telomeric DNA and those that bind double-stranded (ds) telomeric DNA binding.

Within the ss telomeric DNA binding proteins group we find the telomere end binding protein (TEBP) from the ciliate *O. nova* (64, 67), Cdc13 protein from *S. cerevisiae* (18, 68), and the recently identified Protection Of Telomeres (Pot1) proteins from *S. pombe* and humans (69). Despite their highly divergent sequence, these proteins share conserved structural domains called OB-folds (oligosacharide/ oligonucleotide binding-fold).

The OB-fold domain is a compact structural motif frequently used for nucleic acid recognition, although it can also participate in protein-protein interactions. OB-fold consists of a five-stranded β barrel and is found in many single-strand nucleic acid binding proteins (70). The OB-fold containing proteins

are classified on the basis of their functional recognition: no sequence specificity (ie. hsRPA and EcSSB), sequence specificity (ie. EcRho, yeast Cdc13, OnTEBP, and *E. coli* aspartyl-tRNA synthetase) and specificity for non-helical structured nucleic acids (ie. RecG and some ribosomal proteins IF1, L2 and S12) (70).

The ds telomeric DNA binding proteins include Taz1 in fission yeast (71), Rap1 in budding yeast (72), and TRF1 and TRF2 in vertebrates (73-75). These proteins associate with DNA via a Myb-like helix-turn-helix DNA binding motif (76-80). Rap1 possess two Myb motifs. Taz1, TRF1 and TRF2 proteins encode a single Myb domain; however, these proteins homodimerize and this interaction is required for binding to telomeric DNA (71, 81-83). A specific telobox sequence within the Myb domain is conserved in yeast and vertebrates and is thought to be used for recognition of telomeric DNA (76). In addition, the Myb-like telomere binding motif shows similarity to the third repeat of human c-Myb, but displays higher specificity towards telomeric DNA than common Myb substrates (76).

Telomere protein composition in ciliates

The G-overhang binding protein from *O. nova* was the first identified (63-64). The *O. nova* telomere binding protein (TEBP) is composed of two protein subunits, α and β (84). The α subunit binds ssDNA via two N-terminal OB-folds and interacts with the OB-fold of the β subunit via a C-terminal OB-fold (85-86).

In the presence of telomeric ss DNA, a very stable α-β-ssDNA complex is formed (87). Biochemical and crystallization studies of TEBP bound to the ssDNA suggest that TEBP tightly binds the extreme 3' terminus of the *O. nova* G-overhang, thus forming a protective cap at the chromosome end (88).

Telomere protein composition in yeast

In S. cerevisiae, ds telomeric DNA is bound by the negative regulator of telomere length, Rap1, repressor activator protein 1 (89). Rap1 negatively regulates telomere length by recruitment of Rif1 and Rif2 (Rap1-interacting factors 1 and 2) proteins to telomeres (90-91). The proposed mechanism for this regulation is called "protein counting" and states that telomere size is directly proportional to the number of binding sites for ds telomeric DNA proteins (92). It is proposed the telomere length information is transduced from duplex telomeric DNA to the telomere terminus by G-overhang binding proteins (93). Ss telomeric DNA binding proteins directly regulate telomerase by affecting its recruitment and/or activation at the chromosome end (reviewed in (94).

Yeast G-overhangs are bound by Cdc13p (18, 68, 95)(Garvik et al., 1995; Lendvay et al., 1996; Nugent et al., 1996). Cdc13p plays several important roles at telomeres via dynamic interactions with distinct protein complexes (reviewed in (96). Loss of Cdc13 results in chromosome end de-protection and extension of the G-overhang (68). It has been proposed that Cdc13p positively regulates telomere length by direct recruitment of telomerase to the chromosome end.

This process is mediated trough the direct interaction of Cdc13p with Est1p, a component of yeast telomerase. Mutant alleles of Cdc13p that abolish the interaction with Est1p lead to the EST phenotype, even though telomerase is enzymatically active in those cells (81, 97-98). Thus, it has been proposed that Cdc13p recruits telomerase telomere ends in S-phase via its interaction with Est1p.

In addition to its role in telomerase recruitment, other roles have been proposed for Cdc13p. These include negative regulation of telomere length (97, 99), as well as coordination of G-strand and C-strand synthesis by interaction with the catalytic subunit of DNA polymerase α (100).

Cdc13 forms a heterotrimeric complex with the proteins Stn1 (Suppressor of cdc thirteen, 1) and Ten1 (Telomeric pathways in association with Stn1, 1). This complex is called CST and it is involved in both telomere length control and telomere capping (95, 98). Mutations in CST genes result in long G-overhangs, C-strand resection and cell cycle arrest (99, 101). Thus, CST proteins are essential. Like ciliate ss DNA binding proteins, CST proteins contain OB-fold domains (99, 101). CST components have been found in budding yeast and its closely-related species. Stn1 and Ten1 have also been identified in fission yeast (102).

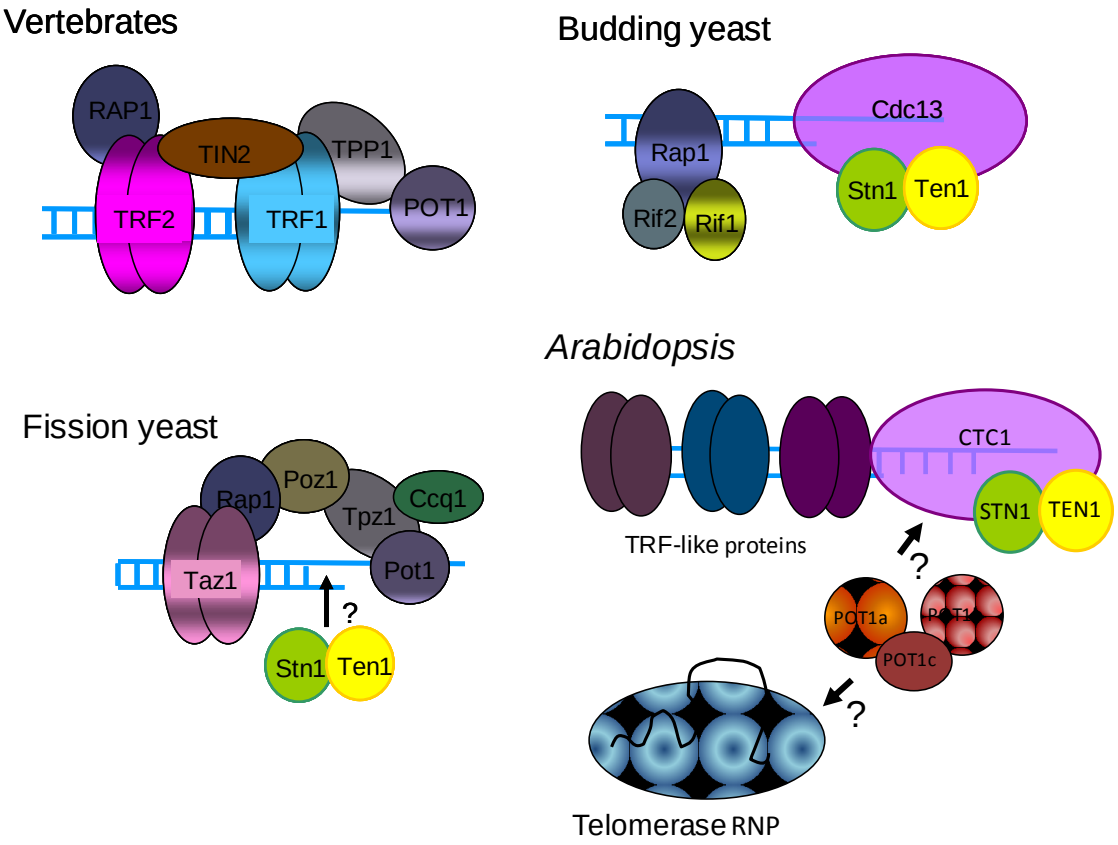


Figure 3. Telomere-associated complexes. Top left, Vertebrate shelterin complex: TRF1, TRF2, RAP1, TIN2, TPP1 and POT1. Top right, Budding yeast CST complex Cdc13/Stn1/Ten1 (CST) complex. Rap1 binds to ds telomeric DNA and associates with Rif1 and Rif2 at telomeres. Bottom left, Fission yeast telomeric complex includes Shelterin like components: Taz1 (TRF1/2 ortholog), Rap1, Poz1, Ccq1, Tpz1 (TPP1 homolog) and Pot1. It also contains CST components: Stn1 and Ten1. Bottom right, *Arabidopsis* CST complex: CTC1, STN1 and TEN1. POT1 proteins do not associate with telomeric DNA, but may be associated with telomerase RNP.

The ds DNA binding protein of *S. pombe* is Taz1 (71). Like ScRap1, SpTaz1 is a negative regulator of telomere length; extremely longer telomeres are observed in *taz1* mutants (71). Since loss of Taz1 leads to C strand degradation and increased homologous recombination at telomeres (71, 103-104), it has been suggested that Taz1 is also involved in chromosome end protection. Rap1 and Rif1 homologs have also been identified in *S. pombe* (105), but in contrast to ScRap1, SpRap1 does not bind telomeric DNA directly. Instead, both Rap1 and Rif1 are recruited to telomeres through interactions with Taz1. SpRap1 and SpRif1, share the negative regulatory role of telomere length likely via an analogous protein counting mechanism (105-106) (Fig. 3).

SpPot1 is the only single-strand telomeric DNA binding protein identified thus far. This protein is essential for chromosome end protection, since *pot1* mutants show immediate loss of telomeric and subtelomeric DNA, chromosome mis-segregation, and profound genome instability (28). In addition, SpPot1 is implicated in the negative regulation of telomere length: reduction of telomere-bound SpPot1 results in dramatic telomere lengthening (107) (Fig. 3).

Telomere-associated proteins in vertebrates

Telomeres in mammals are protected by a complex of six proteins called shelterin (108). Subunits include: TRF1 (Telomeric Repeat Factor binding), TRF2, TIN2 (TRF-interacting protein 2), POT1 (protection of telomeres 1), TPP1, and Rap1 (109-110) (Fig. 2). Shelterin performs many essential functions

including generation of t-loops, control of telomere length and protection of the chromosome terminus (110-112).

In humans, telomere length regulation is primarily accomplished by the ds telomere binding proteins, TRF1 and TRF2 (76-77). These proteins form homodimers and oligomers that coat telomeric DNA (73, 77). They play different functions in telomere biology. TRF1 functions in negative regulation of telomere length (113) while TRF2 is required for chromosome end protection. Overexpression of a TRF2 mutant lacking both its DNA-binding domain and N-terminal basic domain results in severe telomere defects, including loss of 3' G-overhangs and chromosomal fusions (114).

TIN2 serves to link TRF1 and TRF2 to the TPP1/POT1 complex at the G-overhang (115-116). Human Rap1 does not bind DNA directly, but rather associates to TRF2 and it is implicated in negative regulation of telomere (117).

POT1, Protection of telomeres 1, is highly conserved and binds the ss G-overhang. POT1 was identified as a protein similar to the α-subunit of the telomere end binding protein (TEBP) from the ciliate *Oxytricha nova* (Baumann and Cech 2001). While humans and fission yeast have only one POT1 gene, mice encode two POT1 paralogs, mPOT1a and mPOT1b. Both mPOT1a and mPOT1b are required to prevent DNA damage responses at telomeres as well as cell senescence, although functional differences do exist between the two proteins (118-119). TPP1 shows similarity to the β-subunit of TEBP and

interacts with POT1 (120-121). It has been proposed that TPP1 recruits POT1 to telomeres through contacts with TIN2 (110-111).

POT1 binds ss telomeric DNA through its two OB-folds (122). POT1 depletion results in loss of G-overhangs, chromosomal fusions and apoptosis (Hockemayer 2005 Churikov 2006). DNA binding of POT1 results in negative regulation of telomerase (123). In contrast, POT1 mutants that lack DNA binding activity have longer telomeres (93). Thus, both positive and negative regulatory roles have been proposed for POT1.

TPP1 does not directly associate with telomeric DNA. However, its interaction increases the affinity of POT1 for the 3' end of telomeric DNA (120-121). Depletion of TPP1 results in loss of POT1 localization to chromosome ends which leads to telomere lengthening and activation of DNA-damage responses (121). It is hypothesized that binding of TPP1-POT1 complexes to the telomere has an inverse relationship with telomere extension (124). In conclusion, the shelterin complex regulates the access of telomerase to the telomere which results in telomere length regulation.

Telomere-associated proteins in Arabidopsis

Functional orthologs of the yeast CST complex have been identified in *Arabidopsis* and humans ((125-126)and unpublished work). AtSTN1 is essential for chromosome end protection. This protein contains a single OB-fold, and localizes to telomeres *in vivo*. *stn1* mutants exhibit loss of telomeric and

subtelomeric DNA, longer G-overhangs, increased telomere recombination, and massive chromosome end-to-end fusions. The large subunit, Conserved Telomere maintenance Component (CTC1) exhibits many properties similar to budding yeast Cdc13. Several OB-fold domains are predicted for CTC1 and *ctc1* mutants display telomeric phenotypes similar to the *stn1* mutants; this indicates that CTC1 also function in chromosome end protection (126). Finally, a TEN1 ortholog was recently found in *Arabidopsis*. It binds AtSTN1 and seems to be important for telomere length maintenance (Leehy K, Song X and Shippen D, in preparation) (Fig 3).

The shelterin components Rap1, TIN2 and TPP1 do not have obvious sequence homologs in *Arabidopsis*. However, there are six TRF-like proteins which contain a myb-like domain in their C-terminus and bind ds telomeric DNA *in vitro* (127). Analysis of the function of these genes is difficult due to genetic redundancy. However, recent data indicate that AtTBP1 functions in telomere length regulation in a manner similar to vertebrate TRF1 (128).

Arabidopsis contains three divergent POT1-like proteins: AtPOT1a, AtPOT1b and AtPOT1c. AtPOT1a and AtPOT1b contain two OB-folds (129), while AtPOT1c only bears one (A. Nelson, Y. Surovtseva, and D. Shippen, manuscript in preparation). AtPOT1a is required for telomere length maintenance; it physically associates with the telomerase RNP and acts in the same genetic pathway as telomerase, in fact, *pot1a* mutants show progressive telomere loss, like *tert* mutants (130). In contrast, a dominant negative allele of

AtPOT1b leads to chromosome fusions,suggests that could play a role in chromosome end protection (129). AtPOT1c corresponds to a partial duplication of AtPOT1a. AtPOT1c over-expression results in increased G-overhang length and chromosome end protection (A. Nelson, Y. Surovtseva, and D. Shippen, manuscript in preparation).

Notably, a ss telomeric DNA binding activity has not been found yet for *Arabidopsis* POT1 proteins (131-132). Instead, I present evidence in Chapters I-IV that *Arabidopsis* POT1 proteins have evolved to bind the RNA subunit of telomerase and to play distinct roles in telomerase regulation.

Telomerase Ribonucleoprotein Complex

The telomerase enzyme is a ribonucleoprotein (RNP) composed of two essential core elements: a catalytic reverse transcriptase component (TERT), and an RNA subunit (TER). Telomerase activity can be reconstituted *in vitro* from only these two components (133-134). Typically, in these reactions, TERT is transcribed and translated in rabbit reticulocite lysate (RRL). The TER is either co-expressed with TERT or *in vitro* transcribed and added to the RRL to carry out the *in vitro* reconstitution reaction. Additional components present in the RRL (e.g. chaperones) may contribute to forming an active telomerase complex (16). To evaluate telomerase activity *in vitro*, the telomeric repeat amplification protocol (TRAP) assay is widely used (Fig 4). In this assay, the products of the telomerase reaction are amplified by PCR (135). First, a G-rich ss primer

corresponding to the ss overhang of the chromosome end is elongated by telomerase and then a reverse primer complementary to the telomeric repeat is added along with DNA polymerase to amplify the telomerase elongation products.

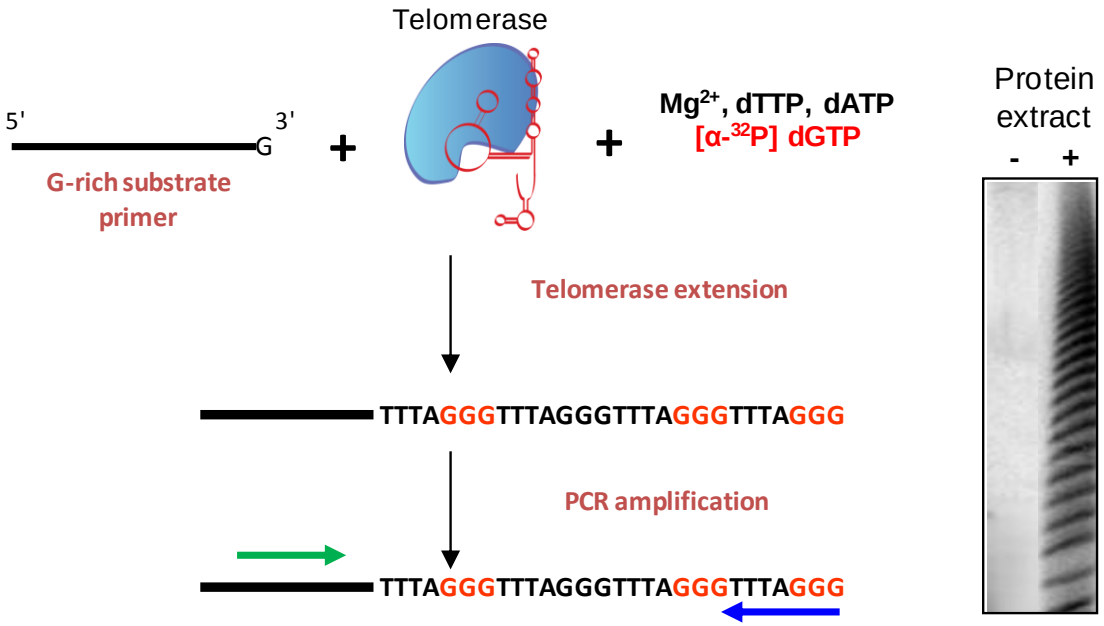


Figure 4. Telomere Repeat Amplification Protocol (TRAP). A G-rich primer is extended by telomerase in the presence of Mg2+, dTTP, dATP and [α-32P]-dGTP. After telomerase extension, dCTP is added and PCR is performed using a reverse primer complementary to the *Arabidopsis* telomeric repeat. Products are resolved by denaturing PAGE and visualized by phosphorimaging (Right panel).

Telomerase reverse transcriptase

The catalytic TERT component was initially identified in the ciliate *Euplotes aediculatus* (136). The TERT sequence is well conserved and several homologs of TERT have been identified in other ciliates (137), yeast (138), vertebrates (including humans) (51, 139-142) and plants (52, 143). All TERTs share sequence homology with reverse transcriptases (RT) from retroelements and retroviruses. Phylogenetic analyses suggest that TERT is most closely related to RTs encoded by non-long terminal repeat (LTR) retrotransposons (144). Therefore, the TERT 3D structure is predicted to be similar to the HIV and MLV RTs, which resemble a right hand with fingers, palm, containing the active site, and thumb domains (145). Each domain of telomerase has distinct functional roles that promote catalysis, nucleolar localization, RNA binding, dimerization and recruitment of telomerase to the telomere (146-147). Telomerase adds one nucleotide at a time to the 3' end of telomeric DNA using its RNA subunit as template (Fig 5). The telomerase-dependent elongation process entails substrate recognition, elongation and translocation (148-149). The mechanisms for ssDNA substrate recognition are variable between telomerases from different organisms and have not been yet well characterized (150). The recognition can involve hybridization of the RNA subunit to the substrate, although very little complementarity to the 3' end is needed (136).

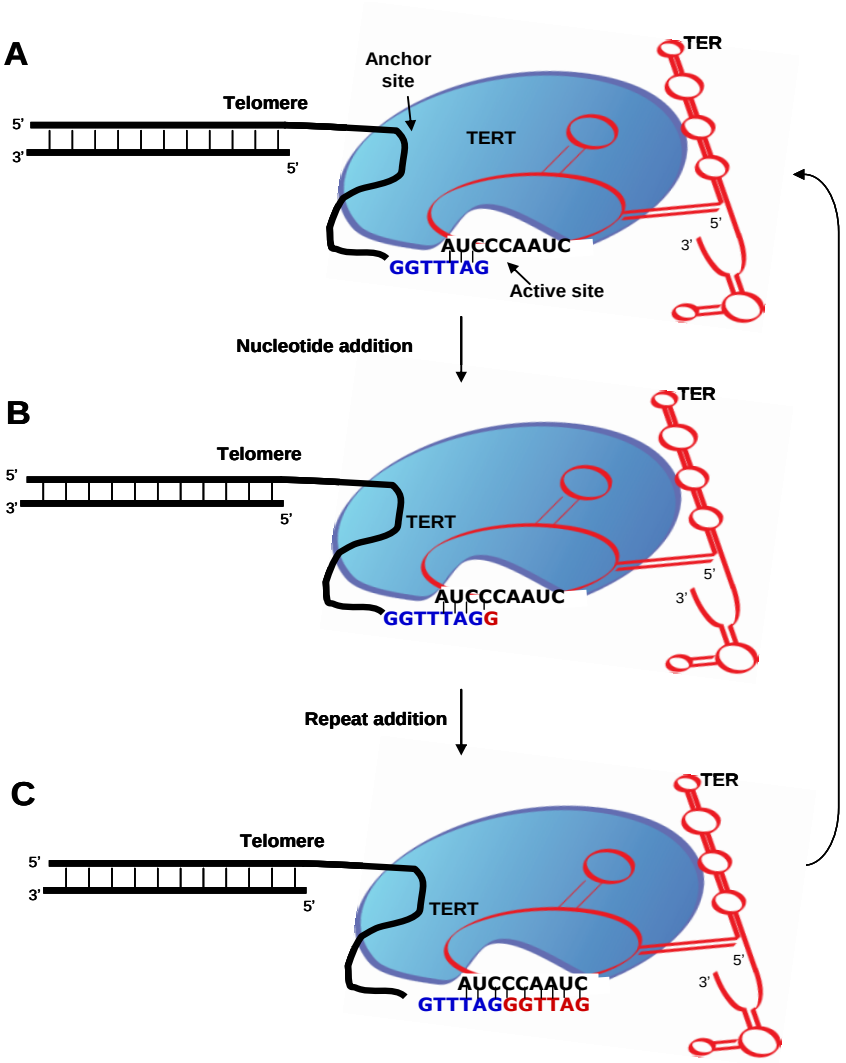


Figure 5. Mechanism of telomere elongation by human telomerase. The G-overhang at the telomere binds to the anchor site in TERT. The 3' end is positioned in the active site and base-paired to the RNA subunit template region. (B) Telomerase adds one nucleotide at the time to the telomeric 3' end. (C) Once the 5' end of the RNA template is copied into DNA, the telomerase RNA template re-aligns it in the catalytic site and two different translocation processes are thought to be carried out at this time. One involves the movement of the DNA-RNA hybrid and the TERT catalytic site and the other translocation process involves repositioning the 3'end of the primer after the entire RNA template has been copied. (Adapted from 22).

Nucleotides will be added depending on the alignment position of the primer on the template. Telomerase shows highest affinity for dGTP, but other dNTPs and rGTP can be also used by the enzyme (137). During substrate elongation, not all nucleotides in the template region are base paired to the telomeric DNA and the RNA-DNA bonds are broken at the distal end of the template as new bonds are formed at the proximal end (151). Telomerase processivity allows the synthesis of several telomere repeats without dissociating. Human and ciliate telomerases catalyze more than one round of telomere repeat synthesis while they are still bound to the same telomeric DNA substrate (152).

On the contrary, telomerases from mouse and yeast are non-processive and generate only short products in standard reactions (138, 153). After elongation, the telomerase RNA template re-aligns in the catalytic site. The telomeric DNA interacts with TERT in a site called the anchor site (154), which allows the primer to remain associated with TERT when the primer translocation occurs (Fig 5 A, B and C).

Both TERT and TER form dimeric complexes. Moreover, dimerization has been shown for endogenous and recombinant RNPs and for telomerase core components separately. Although the yeast and human telomerases appear to function as dimers, telomerases from *Euplotes aediculatus* and *Tetrahymena* act as monomers (146, 155-156) (Fig 6).

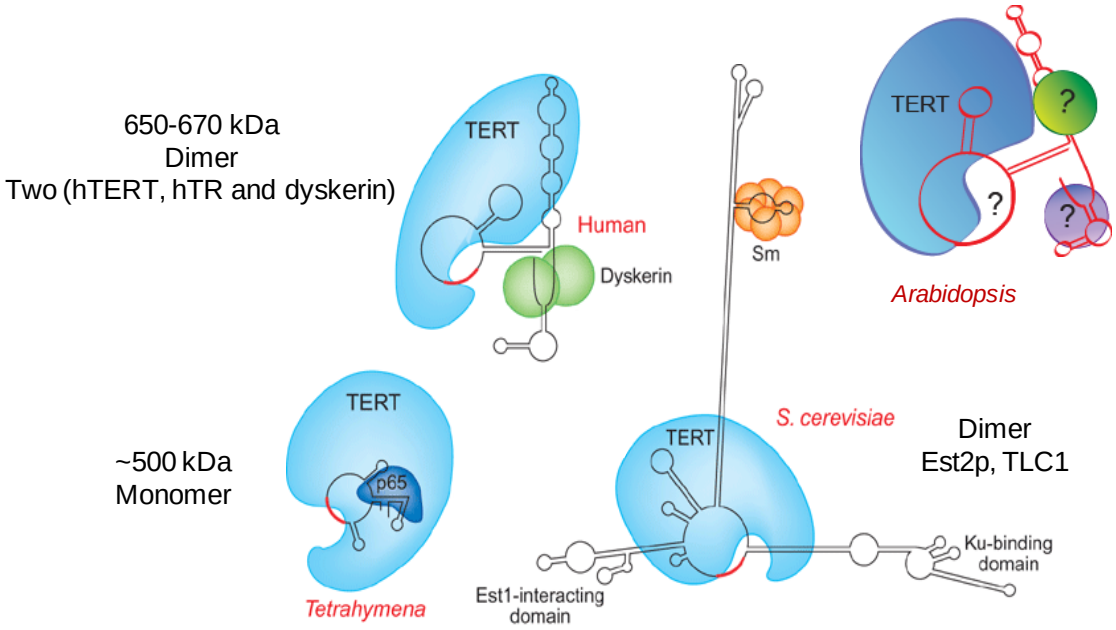


Figure 6. Telomerase RNP complexes of humans, *Tetrahymena*, *S. cerevisiae* and *Arabidopsis*. Active human telomerase is a dimer of ~650kDa (233) that includes two TERT, two TR and two dyskerin molecules. *Tetrahymena* telomerase is a monomer of ~500kDa composed by TERT, TER and p65 (155). *S. cerevisiae* is proposed to be a dimer containing Est2p and Tlc1 (159), as well as the telomerase-associated proteins Est1p, Est3p and Ku. The characterization of the composition of *Arabidopsis* telomerase is the subject of this dissertation (Adapted from (157).

Functional studies indicate that yeast telomerase contains two active sites (158-159). Several hypotheses have been proposed for the role of telomerase dimerization. One postulates that dimerization allows telomerase to extend two independent substrates such as a pair of sister chromatids (123, 159). Another hypothesis involves template switching, the transfer of the 3' end of the DNA from one template to the alternate one after one round of elongation (123). A third postulate is that the template would be supplied by one TERT-TER pair and the anchor site would be supplied by the other pair (123, 160).

In addition to the core TERT and TER components, other proteins that associate with telomerase have been identified. Many of these have roles in RNP biogenesis, assembly and regulation of telomerase at the end of the chromosome. Some of these factors are stably associated with telomerase and other may only interact during telomerase biogenesis (161). Like the RNA subunit, these accessory proteins vary from species to species (Fig. 6).

In ciliates biochemical purification led to the identification of the telomerase-associated protein p43 in *Euplotes aediculatus* (136) and its homolog p65 in *Tetrahymena* (162). These proteins may have roles in telomerase maturation and activity. In yeast the EST proteins are essential for telomerase activity *in vivo* (17-18). Est1p binds the RNA subunit and recruits telomerase to the chromosome ends through interaction with the ssDNA binding protein Cdc13p (Est4) (81, 163). Est3p is also required for telomerase activity *in*

vitro, but its function is not yet understood (164). Est3 requires Est2p to interact with the RNA subunit (165).

Cdc13 and Ku70/80, two protein complexes that bind chromosome ends, have also been shown to interact with telomerase, although they may not be stable constituents of the RNP in cell extracts (166). In humans, the list of proteins associated with telomerase is extensive. For example, chaperones like p23 and HSP90 stably associate with telomerase RNP (167). Other proteins like Ku70/80, hEST1, TEP1, KIP and PINX1, are suggested to regulate TERT activity through direct binding (168).

In conclusion, although the composition of the telomerase holoenzyme greatly differs among diverse species, telomere maintenance is a highly conserved and essential function.

Telomerase RNA

TER subunits diverge greatly between eukaryotes. They show little sequence conservation and their sizes range from 150nt in *Tetrahymena* (14), to 382-451nt in vertebrates (169), and ~1200nt in budding and fission yeast (170-171). Despite the divergence in sequence and size, secondary structure models for several organisms have been constructed independently using a combination of phylogenetic comparative analyses and mutagenesis studies (172).

Based on these secondary structure models some common characteristics are conserved between all TERs (Fig 7). The template region is

single stranded and typically corresponds to one and a half to two telomeric repeats. It is flanked by a 5' boundary element and a 3' pseudoknot domain (172). The 5' boundary element blocks telomerase from reverse transcription beyond the end of the template sequence in TER. In *Tetrahymena,* the 5' boundary element corresponds to a small TERT interacting region (173). In yeast there is a pairing element right upstream of the template (174-175). The pseudoknot domain helps in positioning the template in the catalytic pocket of human and yeast TERT through protein-RNA interactions. By contrast, this domain does not appear to be important for TERT binding in ciliates, but it seems to be required for telomerase assembly *in vivo* (176). NMR structural studies of human TER suggest that the pseudoknot is not a stable structure (177). TER may exist in equilibrium between a pseudoknot state and a single stem loop and might this equilibrium be shifted one way of the other in the prescense of TERT.

In addition to this conserved portion of the RNA, species-specific domains exist (Fig. 7). In ciliate TERs, there is helix IV, a potential TERT binding site that may be involved in nucleotide addition and processivity (178) and helix I, a probable binding site for the protein p43 (179-180). Due to its large size, the yeast RNA works as a scaffold for telomerase-associated proteins. Besides the pseudoknot, where TERT (Est2) is bound, there are three vital stem loop structures. Two are binding sites for the telomerase-associated proteins Est1 and Ku70/Ku80 heterodimer (181-182). The third site at the TER 3' end serves as a binding site for Sm proteins, a key factor for TER biogenesis (181).

In vertebrates, the CR4-CR5 and scaRNA domains have been identified. Like the pseudoknot domain, the CR4-CR5 domain may be important for telomerase activity *in vitro* (183-184). The scaRNA domain contains the Box H and ACA motifs, which are essential for RNA stability *in vivo* and correspond to the binding sites for the dyskerin complex (184). An additional element in the scaRNA domain, CR7, may be important for telomerase RNP localization to the Cajal body (185).

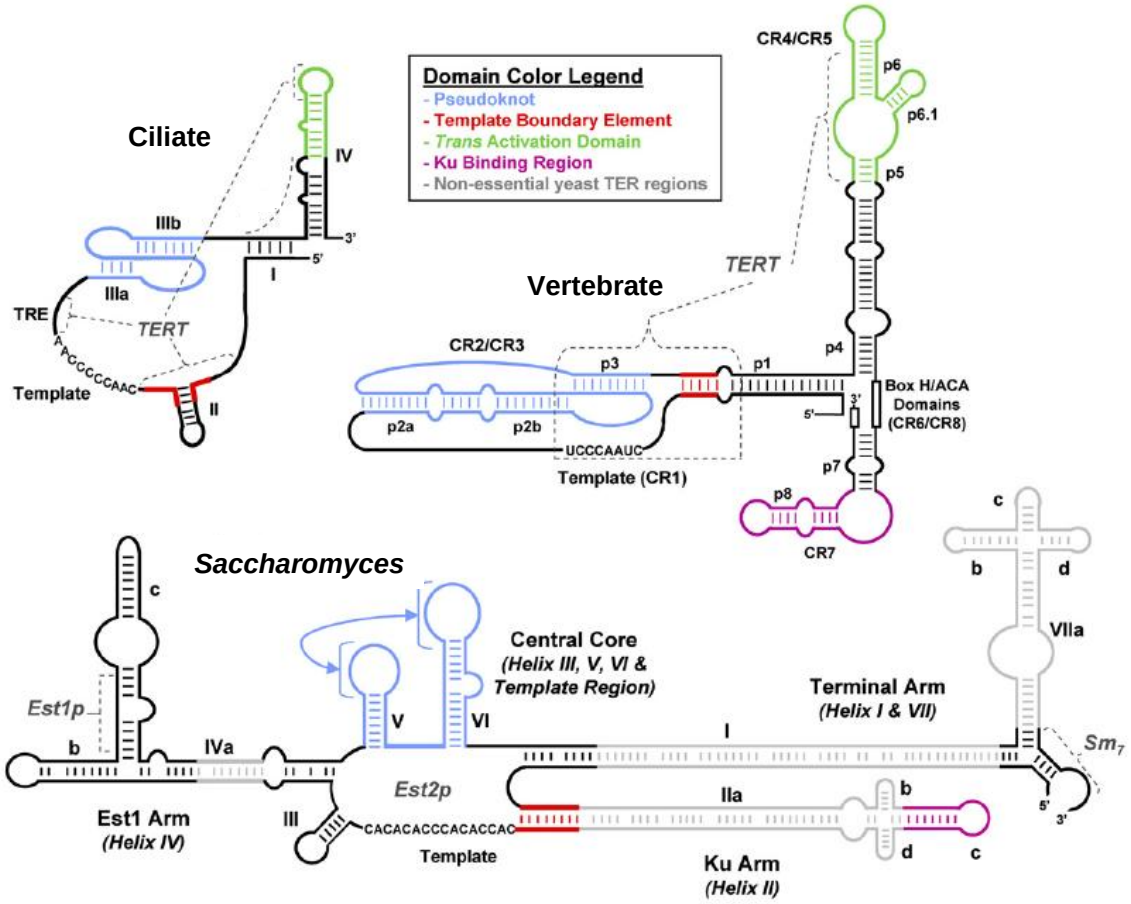


Figure 7. Proposed secondary structures of ciliate, vertebrate and *Saccharomyces* telomerase RNAs. The domain color legend indicates the color code for elements common between RNAs. Brackets illustrate regions defined to interact with a specific protein. Taken from (186).

Studies carried out in *Tetrahymena* (178) and vertebrates (183, 187) show that telomerase activity is reconstituted *in vitro* when two separate TER fragments, one including the template, the template boundary and the pseudoknot and a second RNA containing the activator domain, which interacts with TERT (Stem loop IV for *Tetrahymena* and stem loop P6 for humans), are added independently to the reaction. These experiments strongly suggest that the RNA domains required for activity do not need to be part of only one molecule. Instead the independent domains can be provided in *trans*. Thus, two RNA molecules can interact in the same complex to promote catalysis.

Telomerase Regulation

The tight association of telomerase activity in cancer processes has captured the attention of the scientific community. The detrimental outcomes of telomerase activity imbalance imply that the telomerase RNP must be tightly regulated at different levels to guarantee proper cell proliferation. In non-reproductive proliferative cells, progressive telomere shortening occurs and when telomeres reach a determined size, the cells enter in replicative senescence (188).

In this sense, telomere shortening could be a potent tumor suppressor pathway since it is the way to measure the “permited” number of cell divisions. This mechanism would protect the human cells from becoming carcinogenic. This hypothesis is supported by the observation that expression of hTERT in

normal telomerase-silent cells restored telomerase activity (189) and allows telomeres to be elongated. Moreover, these cells showed an indefinite life span (135, 190). Human tumor cells derived from carcinomas bypass cellular senescence and DNA damage-signalling pathways. The fact that telomerase is frequently activated and highly expressed in advanced human cancers (190), coupled with the requirement of telomerase to sustain the replicative potential of primary cells and tumor cell lines (191) provided the basis for the hypothesis that telomerase is a key player in cellular immortalization (192).

Telomerase action is controlled at several levels. Transcriptional and post-transcriptional regulation of hTERT are major determinant of enzyme activity (193-194) and in some cell types, hTER expression is controlled (194). The stability and processivity of the telomerase enzyme are also modulated by RNP accessory factors (120, 195-199). In addition, telomerase access to the chromosome terminus is controlled in *cis* by telomere binding proteins.

Transcriptional regulation of hTERT and hTR

In vitro telomerase activity can be reconstituted with TERT and TER (134), suggesting that these components are the major targets of telomerase regulation. The majority of human cells express hTR during all stages of development, while hTERT is expressed during early development, but it is undetectable in most normal somatic cells except for proliferative cells or

renewal tissues (200-201). Recent studies indicate that like hTERT, hTR may also be subject to transcriptional regulation in some cell types (194).

Transient transfection of the hTERT promoter-luciferase reporters reveals an expression pattern that mirrors that of telomerase activity: luciferase is inactive in normal cells and active in immortal cells (202). These observations suggest that the control at transcriptional level is a major mechanism to regulate the amount of hTERT mRNA and protein, and subsequently telomerase activity. Computational predictions showed that the hTERT promoter contains binding sites for many transcription factors and hence may be subjected to several levels of control in different cell types (202). Two examples of elements that up-regulate transcription of hTERT mRNA include c-Myc and Sp1 transcription factors, as well as steroid hormones, among others.

Near the hTERT transcriptional start site, two E-box elements are potential binding sites of basic helix-loop-helix zipper (bHLHZ) transcription factors that are encoded by the Myc family of oncogenes (203). Over-expression of c-Myc results in increased activity of the hTERT promoter, leading to increased hTERT mRNA and increased telomerase activity. This is dependent on E-box elements (193, 204-205) induced expression of hTERT mRNA and telomerase activity in normal cells (205).

Another element that influences hTERT mRNA transcription is Sp1 (193), a general transcription factor that interacts with TATA binding protein to promote transcription. Interestingly, TATA boxes are not found in the hTERT promoter;

however, mutations in Sp1 binding sites completely abolish hTERT promoter activity (193, 202).

Negative regulators of hTERT transcription have also been reported. The first identified was WT1, Wilm's tumor suppressor 1. WT1 represses hTERT transcription by direct interaction with the hTERT promoter (206-207) and its expression is restricted to kidney, gonad and spleen. Since WT1 is an important factor in growth of kidney and gonad cells, it has been proposed that WT1 inactivation in target tissues contributes to activation of telomerase during tumorigenesis (208).

Interestingly, Mad1, an E-box binding factor, has also been reported to suppress hTERT expression. Mad1 competes with c-Myc (see above) for a common interaction partner Max and for the E-box binding site (209). The switch from c-Myc/Max to Mad1/Max at the E-boxes represses hTERT expression and subsequently down-regulates telomerase activity (210).

Other factors that repress hTERT transcription by binding to the promoter include MZF-2, a zinc finger transcription factor (211), and a member of the E2F family transcription factors involved in cell-cycle progression E2F-1 (212). Repression of hTERT expression may also be accomplished by inhibiting potential activators like c-Myc, via the TGFβ/Smad signaling pathway (213) or BRCA1, a tumor suppressor for hereditary breast and ovarian cancers which is involved in DNA repair (214). Finally, it has been suggested that p53 may be

involved in the negative regulation of hTERT expression, because most cancers are deficient in p53 (215).

Post-transcriptional regulation of TERT

Telomerase is regulated at the post-transcriptional level. Although hTERT mRNA and hTER are expressed in normal ovarian cells, telomerase activity is not detected (Ulaner 2000). In addition the expression level of hTERT mRNA is similar in human peripheral blood T and B cells, and in tonsils, but the levels of telomerase activity differ in these settings (216).

Observations made in blood cells reveal that telomerase activity can be regulated by direct phosphorylation. Telomerase activity is up regulated by an activator of protein kinase C (PKC) and inhibited by PKC inhibitors (Bodnar 1996). PKC belongs to a family of phospholipid-dependent kinases that function in cell growth, differentiation, and carcinogenesis (216). Another kinase involved in telomerase regulation is the Akt protein kinase, an important component of the phosphatidylinositol 3-kinase signaling pathway that promotes cell proliferation (217). Two phosphorylation sites for this kinase have been found in hTERT and Akt-mediated phosphorylation of peptides containing these sites occurs *in vitro* (218). It is proposed that hTERT phospholylation may mediate its translocation from cytoplasm to nucleus (219-220).

Ubiquitination has also been proposed to play a role in telomerase regulation. An ubiquitin ligase, MKRN1 interacts with hTERT through yeast-two

hybrid assays (221). Over-expression of MKRN1 causes degradation of TERT and results in decreased telomere activity and shortened telomere length. This finding suggests that MKRN1 functions in telomere maintenance by modulating TERT protein stability (221).

Localization of TERT

As mentioned above, telomerase activity can be also regulated by localization of its components. The hTERT N-terminus contains amino acids responsible for proper localization to the nucleus during specific stages of the cell cycle. hTERT is found in subnuclear foci during most of the cell cycle. However, during S-phase, hTERT, together with hTR are present in Cajal bodies, where telomerase RNP is assembly occurs (222). In addition, it has been proposed that hTERT transport to and from the nucleolus may also modulate telomerase activity at the telomere (223). Ectopic expression of GFP-hTERT fusion protein results in telomere length maintenance and cell immortalization. In normal fibroblasts this fusion protein localizes in nucleoli for most of the cell cycle, but in late S/G2, when human telomere synthesis occurs, the fusion protein moved to the nucleoplasm. Since telomerase activity levels did not change during these phases of the cell cycle, the change in localization is not thought to prevent the enzyme from reaching telomeres, leading to cell cycle dependent regulation of telomere extension (223). In immortalized cells

telomerase re-localization is not observed (224), suggesting that the release of hTERT from nucleoli allows telomerase uncontrolled access to telomeres.

Telomerase RNP structure and biogenesis

Ciliates

Ciliate TERs are transcribed by RNA polymerase III (14). Affinity purification of ciliate TERT recovered La motif containing proteins, which function in maturation in RNA polymerase III transcripts. p65, one such protein, associates with stem IV of *Tetrahymena* TER and helps TERT bind TER with higher efficiency (225). p65 depletion results in inappropiate telomerase assembly (162), suggesting that TER biogenesis may be similar to other RNA polymerase III transcripts.

Other proteins associated with *Tetrahymena* telomerase include p20, p45 and p75. Depletion of any these proteins results in telomere shortening rather than defects in RNP assembly (162). Thus, these proteins are implicated in telomerase regulation and recruitment. Finally, p80 and p95 associate with *Tetrahymena* telomerase, but are not telomerase specific (226). Interestingly, *Euplotes crassus* contains three different TERT genes, with distinct expression patterns during macronuclear development. EcTERT-2 expression is limited to the time of de novo telomere formation, whereas EcTERT-1 and -3 are mainly expressed when telomere maintenance is required. Furthermore, E. crassus

telomerase oligomerize *in vitro* and potentially assumes different conformations *in vivo* (227).

Yeast

The telomerase RNA subunit of *S. cerevisiae* is called TLC1 and it is the product of RNA polymerase II transcription. A poly(A) transcript is detected in the cells, however the most abundant TLC form lacks a poly(A) tail. This poly(A) minus form associates with TERT (228). TLC1 shares several characteristics with small nuclear RNAs, inlcuding the presence of a 5'-2, 2, 7-trimethylguanosine or TMG cap (229)and a binding site for Sm proteins (Khusial, P 2005). Deletion of the Sm-motif or loss of the members of the Sm complex leads to reduced levels of TLC1 (229). Thus, the biogenesis pathway of TLC1 seems to parallel that of snRNAs.

Fission yeast TER1 also contains a Sm binding site at its 3' terminus. Mutations in this site lead to telomere shortening and loss of telomerase activity (171). Interestingly, the 3' end of SpTER1 is processed by an unconventional splicing reaction in which only the first transesterification step is perform followed by 5' exon (mature TER1) release (230).

Vertebrates

Human TER, hTR, like yeast TER is an RNA polymerase II transcript and posseses a 5'-TMG cap (231). In addition it contains two regions at its 3' end that are required for proper RNA localization and stability: the H/ACA box and the CR7 domain.

The H/ACA motif is present in small nucleolar RNAs (snoRNAs) and it is bound by the protein complex dyskerin/Nhp2/Nop10/Gar1. snoRNPs function in the pseudouridylation of ribosomal RNA and snRNAs (232). No pseudouridylation target has been identified in hTR thus far. Instead, this complex seems to function in stabilization of hTR. In fact core active human telomerase is composed of TERT, TR and dyskerin (233). Dyskerin is an essential protein and mutations in dyskerin lead to the disease dyskeratosis congenita (He 2002). The ATPases pontin and reptin also function in assembly of telomerase RNP. Pontin and reptin are well conserved AAA+ ATPases (ATPases associated with various cellular activities) that were found in chromatin remodeling complexes, as co-factors for the transcriptional regulators c-Myc and β-catenin and as proteins that interact with small nucleolar RNA (snoRNA) complexes (234). Interestingly, these ATPases interact with the telomerase RNP through direct interactions with dyskerin and TERT and the ATPase activity of pontin is required for telomerase assembly (235), thus pontin and reptin are proposed targets for anti-cancer therapies.

The CR7 loop of hTR is also important for telomerase assembly and it contains a consensus localization signal for cajal bodies (CAB) (185) the nuclear compartments required for modification and assembly of RNPs. hTR accumulates in cajal bodies throughout most of the cell cycle (185, 236). At S phase, hTR co-localizes with hTERT at foci near telomeres, enabling telomerase to act at telomeres (222). Nucleotide substitutions of the CAB box do not affect telomerase activity, but do diminish telomerase recruitment and telomere length (124). Interestingly, Sm binding proteins are also found associated with active human telomerase and it has been proposed that this association is mediated through the CAB motif, since of hTR does not contain a Sm binding sites (237).

Hsp90 and p23 chaperones associate with hTERT. These proteins are required for telomerase activity and are involved in the proper assembly of the telomerase RNP (238). In addition, 14-3-3 protein binds hTERT and it is important for hTERT localization (239).

Prior to this study, there was little biochemical information about the telomerase RNP complex from plants. Partial purification of telomerase from cauliflower suggested a molecular mass of ~670kDa (240). The TERT subunit from *Arabidopsis* was identified several years ago (143) and shown to encode a 131kDa protein containing reverse transcriptase motifs as well as a TERT-specific motif T-motif. Disruption of this gene by a T-DNA insertion leads to progressive telomere shortening, demonstrating its essential role in telomere maintenance (143, 241).

Dyskerin, a core telomerase subunit required for human RNP biogenesis and enzyme function *in vivo*, was also shown to be a component of the *Arabidopsis* telomerase RNP complex which localizes to the nucleolus and is essential for telomere maintenance *in vivo* (242). Although a null mutation the *Arabidopsis* ortholog of dyskerin, AtNAP57 is lethal, the introduction of an AtNAP57 allele containing a T66A mutation result in decreased telomerase activity *in vitro* and disrupted telomere length regulation on individual chromosome ends *in vivo*, suggesting that, like in humans, dyskerin is essential for *Arabidopsis* telomerase function (242).

I show in this study that *Arabidopsis* telomerase associates with three different TER transcripts: TER1, TER2 and the processed form of TER2, $TER2_S$. These transcripts assemble into different telomerase RNPs with diverse protein composition and distinct functions in telomere metabolism. I also show that AtPOT1 family members associate with telomerase RNPs and play different roles in telomerase regulation.

Regulation of telomerase access to the chromosome end by telomere-associated proteins

A crucial task for telomerase is maintenance of telomere length. As described above, several telomere-associated proteins are negative regulators of telomere length. This negative regulation is most cases does not reflect on

telomerase activity, but rather an interference with telomerase access to the telomere.

T-loops are proposed to play a key role in telomere length maintenance. Elizabeth Blackburn has proposed a two-state model to describe telomere length homeostasis (66). In this model, telomeres alternate between “open or uncapped” and “closed or capped” conformations in which the chromosome end can be either accessible or inaccessible to telomerase or other enzymes (66). The model implies that the overall telomere length is a primary determinant for making telomeric DNA accessible to telomerase action. Thus, in agreement with the “counting model”: a shortened telomere, which would be associated with fewer proteins, would be more accessible to telomerase action while a longer telomere, associated to more protein molecules, would be inaccessible to telomerase action (243). This mechanism would maintain telomere length within a species- and cell type-specific narrow range and would prevent excessive telomere shortening or over-extension. The species-specific telomere protein complexes that regulate telomerase access to the telomere were described in previous sections.

Regulation of enzyme activity and recruitment of telomerase to the telomere by holoenzyme components

In addition to telomere-associated proteins, telomerase-associated proteins also contribute to the recruitment of the RNP to the telomere. In budding yeast the Est1 holoenzyme component is proposed to play this role. *est1* mutants display an Ever-Shorter-Telomere phenotype, but no effect in telomerase activity (244-245). Est1 binds a bulged stem of TLC1 and this association is not dependent of Est2, the yeast TERT (181, 244, 246) .Interestingly, fusion of Est1 with the DNA binding domain of Cdc13 results in telomere lengthening and fusion of Est1 with Cdc13 overcomes the EST phenotype observed in these mutants (247). These data indicates that Est1 functions as a telomerase recruitment factor. Additional experiments suggest that telomerase recruitment by Est1 is specific for S phase.

The Ku heterodimer is also implicated in telomerase recruitment through a direct interaction with a stem loop in TLC1 (248). *ku* mutants have shorter telomeres with long G-overhangs (182, 249). Defects in TLC1 nuclear localization are observed also observed in *ku* mutants (250-251). It has been proposed that Ku functions in recruiting telomerase to the chromosome end during G1 phase of the cell cycle (252).

SpEst1 is associated with telomerase in *S. pombe*. Interacts with TER1 and may function in telomerase recruitment (253). Likewise, human Est1 associates with telomerase *in vivo*, although its role in recruitment is still under

investigation (254). Interestingly, in *Arabidopsis*, EST sequence orthologs do not seem to play a role in telomere metabolism, but instead Est1b is important for meiosis. However, POT1a from *Arabidopsis* has several similarities to Est1 protein from yeast (discussed in Chapter IV).

Other telomerase associated proteins function in regulating enzyme activity. In humans, PinX1, a TRF1 and hTERT interacting protein functions as a negative regulator of telomerase (255). Likewise, yeast PinX1 binds to Est2 and interferes with TLC1 binding (256), thus acting as a negative regulator of telomerase function. The Pif1 helicase has also been proposed to negatively regulate telomerase in budding yeast by unwinding the telomeric DNA:template RNA duplexes and thus removing telomerase from the telomeres (257). Thus, a highly conserved and complex network of protein interactions regulates telomerase at several stages of the RNP biogenesis, recruitment and accessibility to the chromosome end, although the particular proteins involved differ with organisms.

Telomeric Repeat Containing RNA (TERRA)

One of the striking features of telomeres is that they are heterochromatic. Thus, it was surprising when recent studies showed that they are transcribed to generate TERRA. Telomeric Repeat Containing RNA, TERRA, is a long non-coding RNA transcribed by RNA polymerase II that contains subtelomeric and telomereic sequences and that is part of telomeric heterochromatin in animals and fungi (258-260). TERRA from humans and yeast contain a 7-methylguanosine (m7G) cap structure (reviewed in (261)) and poly(A) tails at the 3' end, although in humans only 7% of the TERRA molecules appear to contain poly(A) (259-260). This finding indicates that the ends of TERRA transcripts are not post-transcriptionally processed. TERRA sizes vary widely with 100nt-9000nt in mammals and about 380nt in yeast (258).

In humans, TERRA transcription is thought to be driven from promoter-active CpG-islands present in the subtelomeric region, since their methylation state inversely correlates with TERRA abundance (262). TERRA transcription is also affected by other chromatin imprints like H3K9 and H4K20 trimethylation, H3K4 di/trimethylation, H3K79 dimethylation, and the acetylation state of the H3 and H4 tails (reviewed in (261).

Interestingly, ScRap1, a telomere-associated protein, regulates TERRA synthesis and decay (reviewed in (261). These outcomes depend on the presence or absence of Y' elements, which are conserved sequences in the subtelomeric region (263). Interestingly, Y' telomeres do not contain

characteristic marks of heterochromatin like histone acetylation and high Sir3 occupancy (264).

In humans, TERRA is restricted to the nucleus (258, 260) and interestingly, only TERRA lacking poly(A) is detected associated with chromatin (reviewed in (261), suggesting that TERRA transcripts may have different functions. Finally, the non-sense-mediated RNA decay (NMD) machinery may play a role in TERRA removal from the telomere (258).

Arabidopsis thaliana also contains TERRA as well as TERRA antisense transcripts, ARRET (265). In this plant, TERRA and ARRET are not exclusive of telomeres, but also are transcribed from centromeres containing telomeric DNA and siRNAs derived from these transcripts depend on the RNA-dependent DNA methylation (RdDM) pathway for their generation (265).

Finally, TERRA has been proposed to regulate telomerase (260, 266). A fraction of human TERRA is associated with telomerase. *In vitro* experiments suggest that the 3' telomeric sequences of TERRA base pair with the complementary template region of TR. TERRA may also interact with TERT independently of TR (266). TERRA is not thought to act as a substrate of telomerase, but rather a competitive inhibitor for telomeric DNA. Further evidence for telomerase regulation by TERRA has been obtained in yeast *rat-1* mutants, in which TERRA levels increase and telomeres shorten (259).

The Flowering Plant *Arabidopsis* as a Model Organism for Telomere Biology

Arabidopsis has emerged as the premiere model system for plant biology. It possesses a small sequenced genome (~125Mb), has a growth period of less than 6 weeks and it is genetically tractable, which allows for both crosses and transformations. *Arabidopsis* is also an excellent model for telomere biology due to its extraordinary tolerance to telomere dysfunction.

Similar to vertebrates, telomerase expression is tightly regulated and correlates with cell proliferation in *Arabidopsis*. Telomerase is active in highly dividing tissues such as flowers, seedlings or cell culture, and is not detected in leaves and stems (143). *tert* mutants show progressive telomere loss (143, 241), with the appearance of anaphase bridges in the sixth generation when telomeres reach ~1kb in length. Plants survive up to 5 more generations. In the ninth or tenth generation the plants arrest in a terminal vegetative state demonstrating *Arabidopsis* high tolerance to chromosome instability (241).

Arabidopsis telomeres were the first telomeric sequence cloned and characterized from a multicellular organism (22). They are are short relative to many eukaryotic models, making the measurement of telomere length simple by Terminal Restriction Fragment (TRF) analysis (267). In addition, unique subtelomeric sequences are present on 8 out of 10 chromosome arms in *Arabidopsis*, allowing to follow the dynamics of individual chromosome arms and to characterize the nature of chromosome end fusions (268).

Dissertation Overview

In this dissertation I present the biochemical characterization the *Arabidopsis thaliana* telomerase RNP. Chapters I and II describe the identification of three telomerase RNAs in *Arabidopsis thaliana*: TER1, TER2 and the processed form of TER2, termed $TER2_S$. Like TER1 and TER2, $TER2_S$ levels peak in highly proliferative tissues, however the $TER2_S$:TER2 ratio is higher in flowers than in cell culture, indicating that TER2 processing is regulated in different settings.

All three RNAs serve as templates for telomere repeat incorporation *in vitro*. However, TER1 is the primary template for telomere maintenance *in vivo*. Depletion of TER1, but not TER2, leads to progressive telomere shortening. Moreover, mutation of the templating domain in TER1 results in the incorporation of mutant telomere repeats on chromosome ends at a higher frequency than TER2 mutant repeats. TER2, by contrast, is a negative regulator of telomerase. Telomerase activity increases in *ter2* mutants, while TER2 over-expression reduces the TER1 templating function and results in shortened telomeres *in vivo*. TER1 TER2 and $TER2_S$ assemble into different RNP complexes with distinct protein composition.

Chapters II, III and IV show a novel RNA binding activity for AtPOT1 family proteins. In other organisms POT1 proteins bind ss telomeric DNA, whereas in *Arabidopsis* they have evolved to bind TER. POT1a, a positive regulator of telomerase function, associates with TER1 to form an enzymatically

active RNP. The Kd of the AtPOT1a-TER1 association is 2 X 10-7 M, which similar to the Kd of human POT1a for single-stranded DNA. In contrast, POT1b and KU, negative regulators of telomerase assemble with TER2.

Chapter IV describes the structural and sequence requirements of the POT1a-TER1 association. I define a minimal POT1a binding region within the highly divergent TER1 5'-end. Structurally, this region is characterized by a 6nt U-C rich internal loop flanked by a two base paired stems. Structures with similar characteristics were found in putative TER subunits from *Arabidopsis* close relatives. Also, I show that AtPO1 OB-folds act as nucleic acid recognition domains. TER1 binding is supported by the first OB-fold, although a more robust interaction is observed with the combination of both OB-folds.

Chapter V shows the initial characterization of TER2 processing into $TER2_S$. I provide evidence that $TER2_S$ is a bona fide product of TER2 self-splicing. TER2 processing results in lower levels of telomerase activity in *in vitro* reconstituted RNP reactions. Notably, TER2s production does not require guanosine or Mg2+ cofactors, and thus the processing of TER2 is not mediated by a self-splice group I or II intron mechanism.

Finally, in Chapter VI, I discuss the contribution of this study to understanding the mechanisms of telomerase regulation in other eukaryotes and propose future experiments that will provide new knowledge into the telomerase RNP regulation.

CHAPTER II

DISSECTING THE *ARABIDOPSIS* TELOMERASE RNP: DUPLICATION OF TER AND A NEW RNA BINDING COMPONENT, POT1a

Summary

Telomerase is a ribonucleoprotein reverse transcriptase whose essential RNA subunit (TER) functions as a template for telomere repeat synthesis. Here we report the identification of two divergent TERs in the flowering plant *Arabidopsis thaliana*. Although both TER1 and TER2 serve as templates for telomerase *in vitro*, depletion of TER1, but not TER2, leads to decreased telomerase activity and progressive telomere shortening *in vivo*. Moreover, mutation of the templating domain in TER1 results in the incorporation of mutant telomere repeats on chromosome ends. Thus, TER1 provides the major template for telomerase *in vivo*. Unexpectedly, we found that POT1a binds TER1 with a Kd of 2 X 10-7 M and the two components assemble into an enzymatically active RNP *in vivo*. In contrast, TER1-POT1b and TER2-POT1a associations were not observed. In other organisms POT1 proteins bind telomeric DNA and provide chromosome end protection. Thus, we propose that duplication of TER and POT1 in *Arabidopsis* fueled the evolution of novel protein-nucleic acid interactions and the migration of POT1 from the telomere to the telomerase RNP.

Introduction

The telomerase reverse transcriptase is responsible for synthesizing and maintaining tracts of telomeric DNA on chromosome ends. A ribonucleoprotein (RNP), telomerase contains two essential components: a catalytic reverse transcriptase subunit (TERT), and an RNA subunit (TER) that serves as a template for telomeric DNA addition. Telomerase activity is primarily confined to cells with extended proliferation potential (e.g. the germline and stem cell populations)(269), and in settings defined by restricted proliferation programs, the enzyme is inactive. The absence of telomerase leads to progressive telomere loss, ultimately eliciting a DNA damage response that culminates in cell cycle arrest, genome instability and in humans, replicative cell senescence and apoptosis (270).

Telomerase access to the chromosome terminus is controlled in cis by telomere binding proteins. One key telomere protein is Protection Of Telomeres (POT1). POT1 binds single-strand G-rich telomeric DNA via oligonucleotide/oligosaccharide binding folds (OB-fold) (29, 122, 271) that evolved to specifically exclude RNA (272). POT1 is implicated in many facets of telomere biology, including control of telomere replication (both positive and negative regulation of telomerase), suppression of the DNA damage response, and protection of chromosome ends from inappropriate recombination, nuclease attack and end-to-end fusion (108, 273). Notably, *Arabidopsis thaliana* encodes two highly divergent POT1-like paralogs, AtPOT1a and AtPOT1b (129), neither

of which binds to telomeric DNA *in vitro* (130, 274). AtPOT1a is not required for chromosome end protection, but rather serves as a positive regulator of telomerase activity that physically associates with the RNP *in vivo* (130).

While TERT can readily be identified by its conserved reverse transcriptase motifs, TER has diverged dramatically and exhibits little sequence similarity and vastly different sizes, ranging from 150nt in *Tetrahymena* (14) to >1200nt in yeast (170, 253). Phylogenetic and mutational analyses reveal functionally conserved elements within TER including a single-stranded templating domain typically corresponding to one and a half telomeric repeats flanked by a 5' boundary element and a 3' pseudoknot domain (133, 174, 275-276). Vertebrate TERs harbor a box H/ACA snoRNA motif, which binds dyskerin and is required for RNP maturation and nuclear localization (277).

Telomerase activity can be reconstituted *in vitro* (278-280) with TERT and TER, although additional proteins assemble with the core telomerase RNP *in vivo*. For example, Est1 is a non-catalytic component of the yeast telomerase that facilitates enzyme recruitment/activation at chromosome ends (98, 281). Relocation of the Est1 binding site to a different position in TER does not diminish telomerase activity *in vivo* (282), indicating that TER is a modular, highly flexible scaffold for telomerase proteins. The Ku70/80 heterodimer also binds TER in budding yeast and vertebrates (283-284), but not in fission yeast (285). Ku promotes telomerase recruitment to non-telomeric DNA for de novo

telomere formation in *S. cerevisiae* (286); its role in vertebrate telomere maintenance is unknown.

We report the discovery of two distinct TERs in A. *thaliana*, the first TERs identified in the plant kingdom. Both TER1 and TER2 function as templates for TERT reverse transcription *in vitro*, but only TER1 plays a significant role in telomere maintenance *in vivo*. Moreover, we show that TER1, but not TER2, specifically binds POT1a *in vitro* and the two components assemble into an enzymatically active RNP *in vivo*. These findings underscore the evolution of telomerase and telomere proteins in *Arabidopsis* and argue that the process is driven by gene duplication.

Results

Identification of two TERs in Arabidopsis

We took a biochemical approach to identify the *Arabidopsis* TER. Telomerase was partially purified from cell culture nuclear extracts by three sequential chromatographic steps (Fig. 8). Enzyme activity was monitored after each step by the telomere repeat amplification protocol (TRAP). Size exclusion chromatography revealed that the peak of telomerase activity correlated with a complex of ~670kDa, similar to the human telomerase RNP (287). Total RNA was extracted from heparin column peak activity fractions, 3' end radiolabeled and separated on a 6% acrylamide gel. Numerous RNAs were detected (Fig. 9). The gel was divided into slices, RNA was extracted and cDNA was generated by

reverse transcription using random pentadecamers. Sequencing reactions were performed with primers corresponding to the seven permutations of the *Arabidopsis* telomere repeat sequence (T3AG3)n. BLAST searches of resulting sequences retrieved a hit termed TER1, which maps to chromosome 1 in a region overlapping the 5' UTR and the first 2 exons and introns of an unknown protein coding gene, AT1G71310. TER1 encodes a putative template sequence (5'-CTAAACCCTA-3') corresponding to 1.5 copies of *Arabidopsis* telomere repeat (Fig. 10A).

Unexpectedly, a second BLAST hit, termed TER2, was also uncovered (Fig. 10A). TER2 maps to chromosome 5, in the opposite direction and partially overlaps the 5' UTR of another unknown protein coding gene, AT5G24670. TER2 also contains 1.5 copies of the *Arabidopsis* telomere repeat. RT-PCR confirmed that both TER1 and TER2 RNAs were enriched in purified telomerase fractions, in contrast to a U6 snRNA control (Fig. 9C and 9D). 5' and 3' mapping experiments revealed that TER1 encodes a 748nt RNA and TER2, a 784nt RNA (Fig. 10A, 10B and Fig. 11). Unlike human (288) and yeast TER (228), we found no evidence of a poly-A+ tail for either *Arabidopsis* RNA.

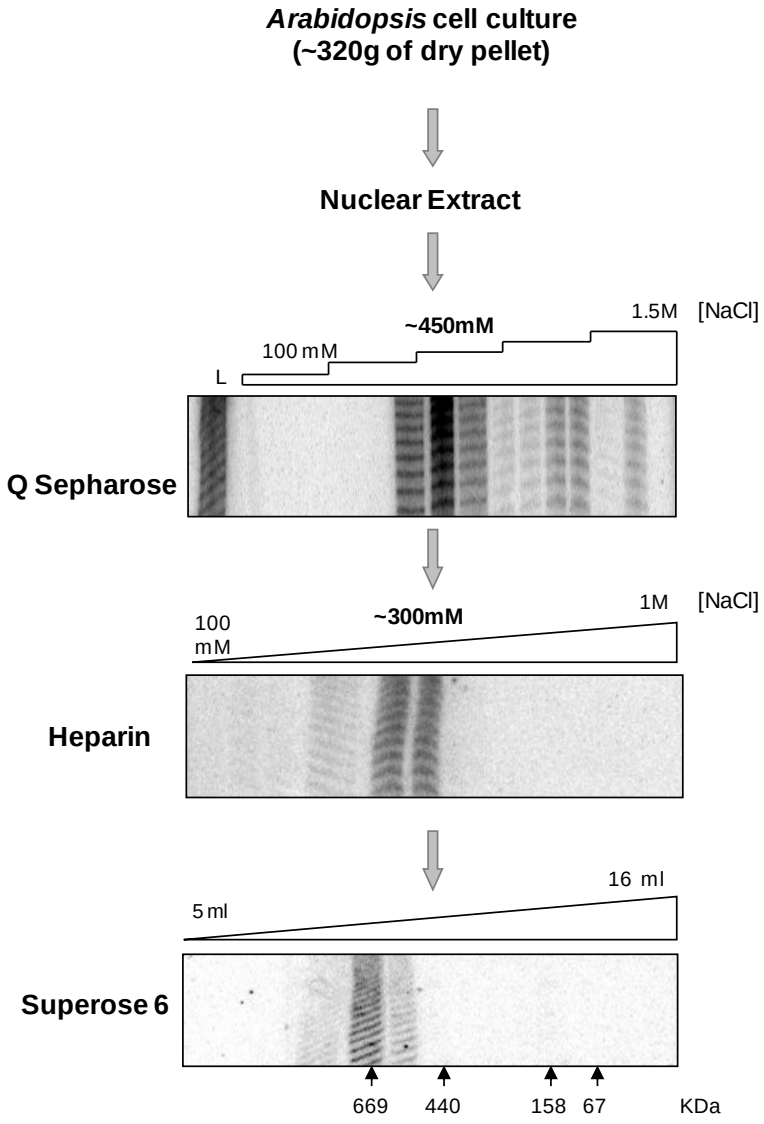


Figure 8. Purification of *Arabidopsis* telomerase. The purification scheme is shown. Telomerase activity was followed by TRAP after each purification step. The Q-Sepharose column was eluted using a 100mM-1.5M NaCl pH 8.4 step gradient. The Heparin column was eluted using a 100mM-1M NaCl pH8,0 continuous gradient. The Superose 6 size exclusion column was eluted with 20 ml of 100mM NaCl pH 8.0 and calibrated with size markers run under the same conditions. BSA (67kDa), Aldolase (158kDa), Ferritin (440kDa) and Thyroglobulin (669kDa) were used as markers. Telomerase activity was present in fractions corresponding to ~670 kDa.

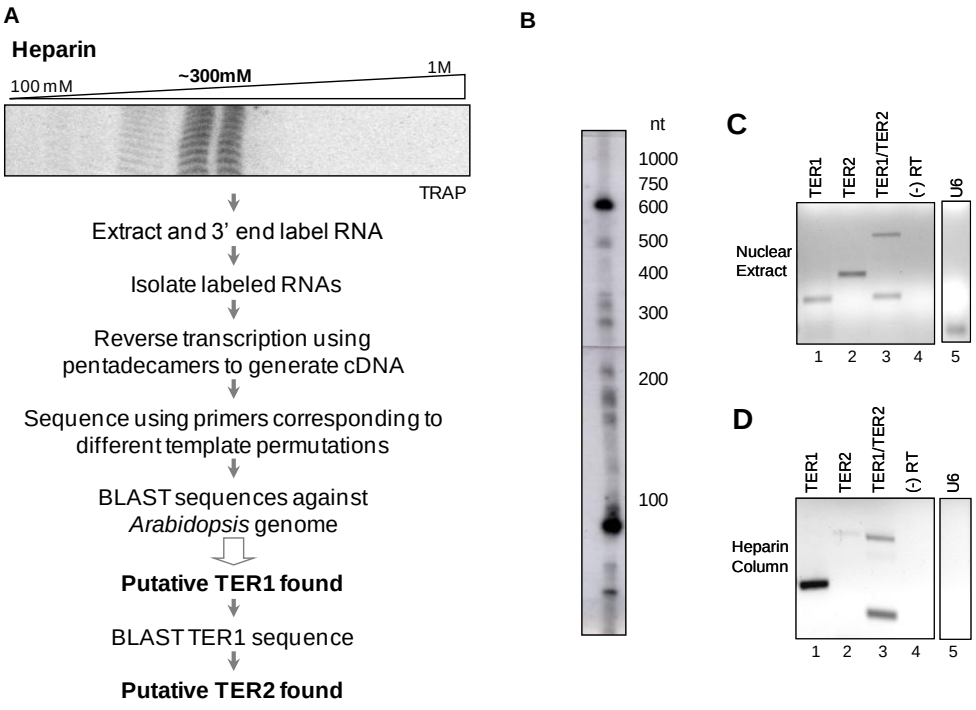


Figure 9. Purification of putative TERs from partially purified telomerase complexes. (**A**) Strategy for identification of TERs. cDNA from the sub-pool of RNAs isolated from each gel slice was sequenced using seven different permutations of the predicted *Arabidopsis* RNA template and the resulting sequences were used in a BLASTN search against the nucleotide collection database (NCBI) and the *Arabidopsis* transcript database (TAIR8 – introns +UTRs). The criterion to accept the BLAST hits as TER *Candida*tes was the presence of the *Arabidopsis* telomeric sequence in the sequenced region. Most hits corresponded to rRNA. Only two RNAs were found to include the *Arabidopsis* telomeric sequence. (**B**) Denaturing 6% PAGE gel showing PCP labeled RNA from the peak of activity fractions of the heparin column. (**C**) RT-PCR results from cell culture nuclear extracts. cDNA was generated with random pentadecamers. Reactions with primers specific for TER1 or TER2 or both are indicated. Lane 3, top band is derived from TER2 and bottom band from TER1. Lane 4, U6 snRNA amplified as a control. Lane 5, primers targeting the TER conserved region were used for the (-) RT control. (**D**) RT-PCR results from telomerase-positive fractions obtained from two steps of ion-exchange chromatography (see Materials and Methods). Lanes 3-5, same as in C, except a different set of TER1 and TER2 specific primers was used.

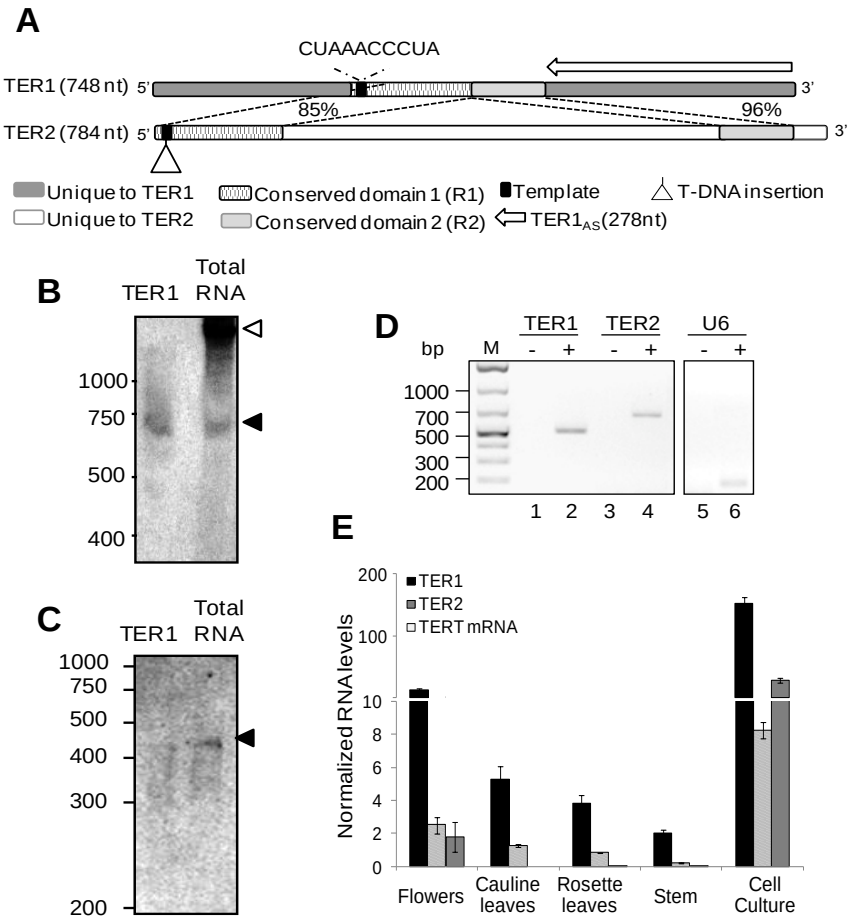


Figure 10. Identification of two TERs in *Arabidopsis*. (**A**) Diagram of TER1 and TER2. Templates, unique and conserved domains are shown. The region in TER1 targeted for antisense and the T-DNA insertion in *ter2-1* are indicated. Northern blot (**B**) and primer extension analysis (**C**) of total RNA from *A. thaliana* cell culture are shown. The probes for northern blotting and primer extension were directed at unique regions in the 5' and 3' ends of TER1, respectively. Closed triangles, TER1 transcript; open triangle, transcript that appears to be derived from AT1G71310, the protein encoding gene into which TER1 is embedded. *In vitro* transcribed TER1 served as a control for B and C. (**D**) RT-PCR analysis of total RNA from cell culture. cDNA was generated using random pentadecamers. Odd lanes show minus reverse transcriptase controls U6 snRNA was amplified as a control (lanes 5,6). (**E**) Quantitative RT-PCR results for TER1, TER2, and TERT mRNA. RNA levels were normalized to U6 snRNA and to the efficiency of each primer pair to allow comparison.

TER1 and TER2 contain a 220nt stretch of ~90% identity (Fig. 10A), but in TER2 this segment is divided into two regions (R1 and R2) separated by an unrelated segment of 529nt. The 5' region (R1) in TER2, which corresponds to 114nt and includes the telomere sequence, shares 85% identity with TER1. The downstream 75nt block (R2) exhibits 96% identity with TER1. The telomere sequence is located in different positions in the two RNAs. The telomere sequence is located in different positions in the two RNAs. In TER1 it is embedded in the RNA (241nt from the 5' end) similar to the human TER, while in TER2 it lies at the extreme 5' end of the RNA (8nt from the terminus) as in mouse TER (172) (Fig. 10A). Putative H/ACA boxes are present in the 3' ends of both TER1 and TER2, consistent with our previous finding that dyskerin is a component of *Arabidopsis* telomerase (242).

Northern blotting of total cell culture RNA revealed ~750 nt transcript corresponding to TER1 (Fig. 10B). Primer extension confirmed the presence of a TER1 transcript (Fig. 10C). Attempts to detect TER2 by Northern blotting or primer extension using probes that target the conserved region in the two RNAs or the unique region in TER2 were unsuccessful, likely due to the low abundance of this RNA. As an alternative approach, end-point RT-PCR was employed. Both TER1 and TER2 were detected by this method (Fig. 10D). Quantification by real time RT-PCR showed that like TERT (34), the abundance of TER1 and TER2 peak in cells with high telomerase activity (e.g. flowers and cell culture) (Fig. 10E). However, TER1 levels were consistently higher than

TER2. In flowers, TER1 abundance was 10-fold higher than TER2 and in cell culture the difference was 20-fold (Fig. 10E).

Both TER1 and TER2 serve as templates for telomerase in vitro

To test whether TER1 or TER2 form a functional telomerase RNP with TERT *in vitro*, T7-tagged TERT was expressed in rabbit reticulocyte lysate (RRL) and incubated with *in vitro* transcribed TER1 or TER2. RNP complexes were immunoprecipitated (IP) and telomerase activity was monitored by TRAP (Fig. 12A). Telomerase activity was dependent on the addition of TER1 or TER2 (Fig. 12A and 11B). Activity was not detected with an antisense version of TER1 or an unrelated RNA (At-oRNA-415) containing a 10nt telomeric sequence (Fig. 12B). In addition, telomerase activity was not observed in the absence of TERT or a boiled sample control (Fig. 12B). We conclude that both TER1 and TER2 assemble into enzymatically active telomerase RNPs *in vitro*.

Deletional mutagenesis was used to establish the minimal TER1 and TER2 sequences required for telomerase activity. For TER1, telomerase activity was detected from *in vitro* assembled RNPs containing a fragment of 534nt (TER1-A, Fig. 12C and E) or 478nt (TER1-B, Fig. 12C and E). Notably, the 220nt conserved region of TER1 was sufficient to reconstitute telomerase activity (TER1-C, Fig. 12C and E). A minimal TER1 of 150nt, roughly corresponding to R1 in TER2 produced activity comparable to full length TER1

(TER1-D, Fig. 12C and E). As expected, no activity was recovered when a 5nt deletion in the TER1 template region was introduced (TER1-E, Fig. 12C and E). Fusion of the two conserved segments in TER2 (R1 and R2) to generate a contiguous 220nt segment (TER2-B), reconstituted telomerase activity to the level of full length TER2 (Fig. 12D and F). As with TER1, the R1 segment in TER2 was sufficient to fully reconstitute telomerase activity (TER2-D, Fig. 12D and F). However, telomerase activity was not observed when TER2 R1 plus 129nt of downstream sequence was expressed (TER2-C, Fig. 12D and F). This additional RNA segment may interfere with the folding of functional domains within TER2 R1. Thus, the core elements required for TER1 and TER2 function reside within a 150nt region conserved in both RNAs.

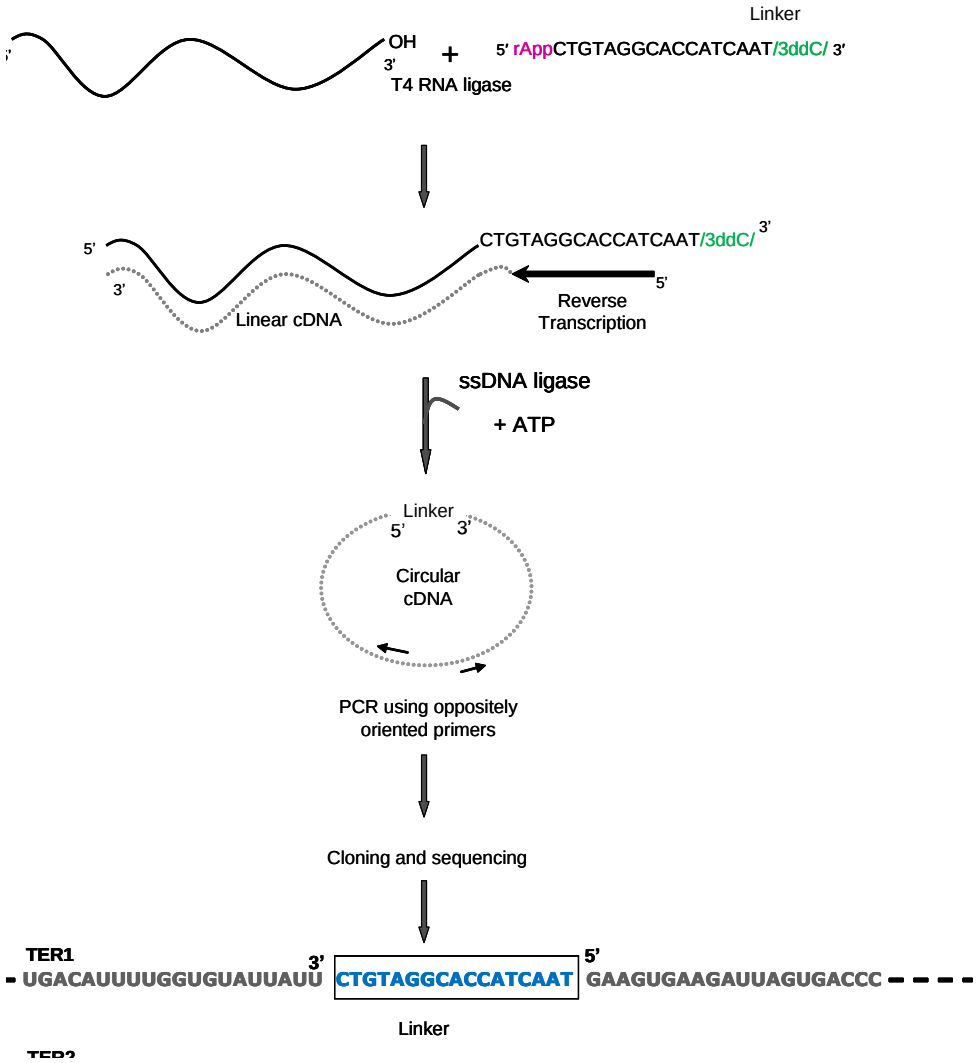


Figure 11. Strategy for 5' and 3' end mapping of TER1 and TER2. The approach is described in detail in the Supplemental Materials and Methods section. Briefly, a 3' unique DNA linker containing a 5'-5' linked adenylate (rApp) and a 3' dideoxynucleotide (ddCTP) was ligated by T4 RNA ligase to RNA extracted from the peak activity fractions from the heparin column. cDNA was generated and circularized using ssDNA ligase in the presence of ATP. PCR was carried out using oppositely oriented primers. The product was cloned and sequenced. The resulting sequenced region is shown and includes the sequence of the linker.

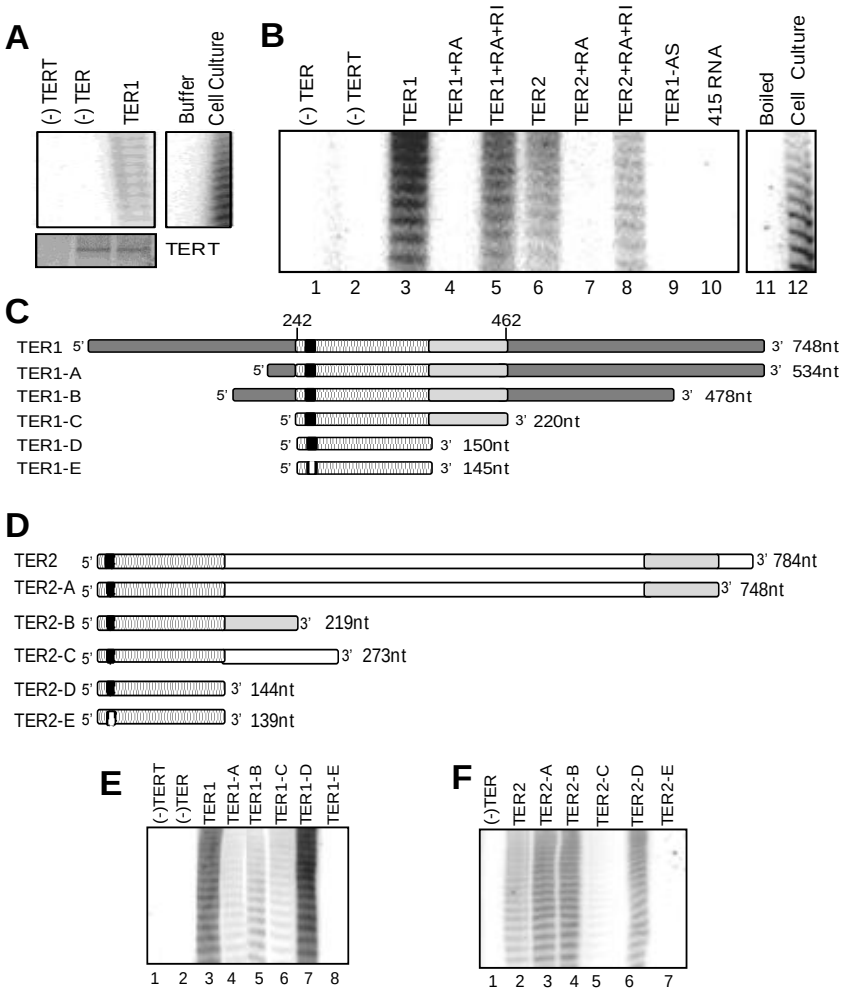


Figure 12. TER1 and TER2 function as templates for TERT *in vitro*. (**A**) Telomerase reconstitution with recombinant TERT and TER1. Top, TRAP results. *A. thaliana* cell culture, positive control; buffer, negative control. Bottom, western blot analysis to monitor TERT expression. (**B**) TRAP results with recombinant TERT, TER1 or TER2. Controls include treatment with RNase A (RA) or RNase A plus RNase Inhibitors (RI); antisense TER1 ($TER1_{AS}$); unrelated *Arabidopsis* transcript containing an 10nt telomere sequence (415RNA); boiled reconstituted RNPs and total protein extract from cell culture. (**C**, **D**) Diagram of TER1 and TER2 deletion mutants. White block in the template indicates a 5nt deletion. Nucleotide positions for TER1 mapping experiments (Fig. 14D) are indicated. (**E**, **F**) TRAP results for reconstitution reactions with truncated TER1 and TER2 constructs.

TER1 is the template for telomere maintenance in vivo

A T-DNA insertion line (ter1-1) is available for TER1, but it is not a null allele as transcripts continue to be produced from this locus. As an alternative strategy to study TER1 function *in vivo*, we expressed a TER1 antisense RNA from the constitutive KU70 promoter (35) in wild type plants, targeting a unique 278nt region at the 3' end of TER1 ($TER1_{AS}$) (Fig. 10A). Relative to wild type and untransformed siblings, TER1 was reduced 10-fold in first generation (T1) KU::TER1AS transformants and ~16-fold in the second generation (T2) (Fig. 3A). TER2 levels were unaffected (Fig. 13A).

Quantitative TRAP (Q-TRAP) showed that telomerase activity decreased 8-fold in T1 and 18-fold in T2 plants (Fig. 14B), paralleling the reduction in TER1. Terminal restriction fragment (TRF) analysis was performed to examine telomere length. As expected, plants transformed with the KU promoter alone or a 35S::GFP construct showed no defects in telomere maintenance (Fig. 13B). In contrast, telomere tracts were more heterogeneous and shorter in T1 KU::$TER1_{AS}$ mutants versus wild type siblings (Fig. 14C and 13C). T2 KU::$TER1_{AS}$ plants showed preferential loss of longer telomere tracts compared to T1 transformants, and increased length homogeneity (Fig. 14C and 13C). This result reflects reduced telomerase activity, since telomerase is known to preferentially extend the shortest telomeres in the population (289). Telomeres below 1kb are subject to end-to-end fusion in *Arabidopsis* (290). None of the telomeres in T2 TER1AS mutants fell below this critical threshold, and therefore

it was not surprising that no defects in plant morphology, fertility or genome instability were observed.

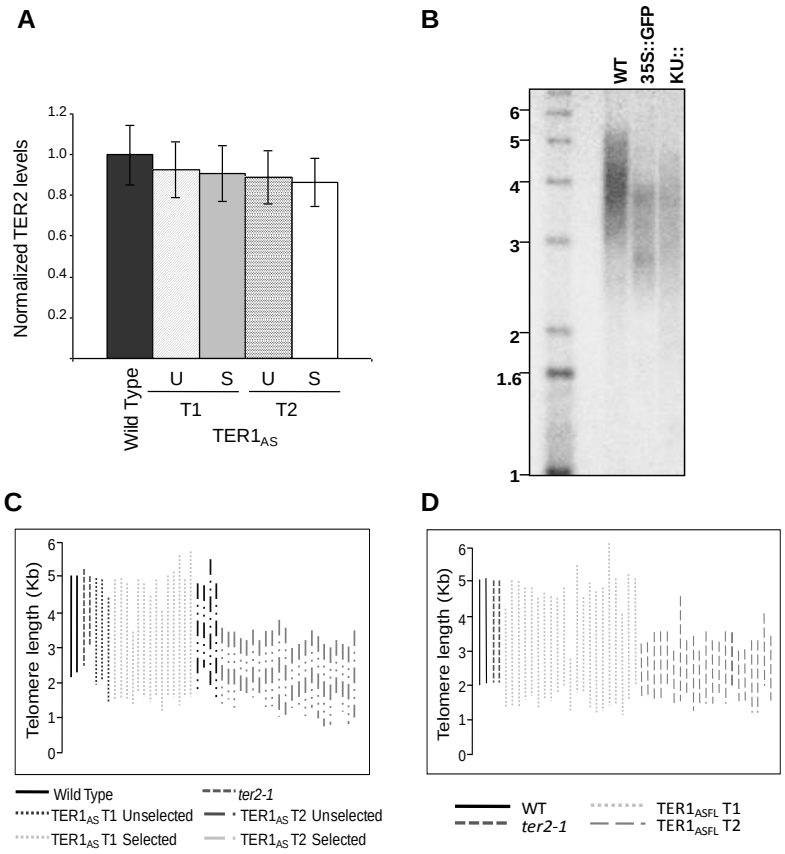


Figure 13. TER1 knock-down lines. (**A**) TER2 levels in $TER1_{AS}$ plants**.** TER2 measured by quantitative RT-PCR was normalized to β-actin and to TER2 in wild type plants. The wild type group included 15 plants. The wild type group included 15 plants. There were 5 plants in each $TER1_{AS}$ unselected (U) T1 and T2 groups and 10 plants in each $TER1_{AS}$ herbicide selected (S) T1 and T2 groups. (**B**) Telomere length in antisense controls plants. TRF analysis of wild type (WT) plants transformed with GFP under the control of the 35S promoter (35S::GFP) and plants transformed with the KU promoter alone (KU::). Telomere length falls within the wild type range. (**C-D**) Graphic depiction of telomere length in $TER1_{AS}$ (**C**) and $TER1_{ASFL}$ (**D**) T1 and T2 populations. Bars indicate range of telomere lengths from TRF analysis.

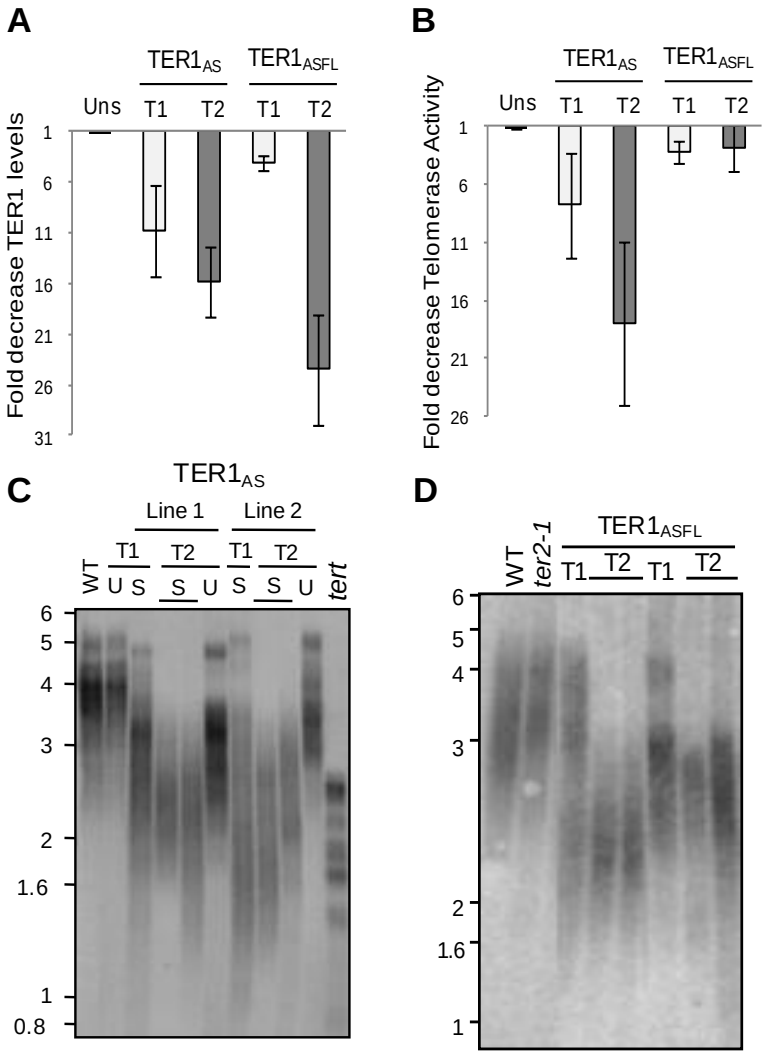


Figure 14. Telomere maintenance in *Arabidopsis* is dependent on TER1, but not TER2. (**A**) Real time RT-PCR results for first (T1) and second (T2) generation $TER1_{AS}$ and $TER1_{ASFL}$ transformants. Results for control plants not selected on kanamycin (Uns) are included. Signals were normalized to β-actin and to wild type (WT) TER1 levels. (**B**) Q-TRAP results for $TER1_{ASFL}$ and $TER1_{AS}$. Values are normalized to telomerase activity from WT plants. (**C, D**) TRF analysis. In C, the T1 $TER1_{AS}$ plant selected for kanamycin resistance (S) is the parent of the T2 progeny. Molecular weight markers in kbp are shown.

Unlike TER1, there are several T-DNA insertions available for TER2. In the ter2-1 line (SAIL_556_A04), the T-DNA insertion lies directly within the templating domain of TER2 (Fig. 10A). No TER2 transcript is detected in plants homozygous for *ter2-1*. In striking contrast to $TER1_{AS}$ mutants, we found no defects in telomere maintenance in *ter2-1* (Fig. 14D). Simultaneous reduction of TER1 and TER2 was achieved by expressing an antisense RNA targeting full-length TER1 (35S::$TER1_{ASFL}$) in a *ter2-1* background. In this setting, TER2 was undetected and TER1 was reduced 3-fold in T1 plants and 24-fold in T2 (Fig. 14A). Q-TRAP showed that telomerase activity decreased ~3-fold in both generations (Fig. 14B). T1 mutants displayed highly heterogeneous telomeres ranging from 1-6 kb (Fig. 14D and 13D). Since telomere shortening was not observed in ter2-1 mutants and the reduction in telomerase activity and telomere length were similar in plants deficient in TER1 or both TER1/TER2, we conclude that TER1, not TER2, is required for telomere maintenance.

To examine the templating function of TER1 *in vivo*, site-directed mutagenesis was used to mutate the 5'-CUAAACCCUA-3' sequence in TER1 to 5'-CUAAAGGCUA-3' ($TER1_{CC}$). *In vitro* reconstitution reactions revealed that mutant RNA assembled into an active RNP and synthesized TTTAGCC repeats (Fig. 15A and 16A). The $TER1_{CC}$ construct was expressed in wild type *Arabidopsis* under the control of the powerful Cauliflower Mosaic Virus (CaMV) 35S promoter. As expected, TRAP products were generated when the wild type reverse primer was used (Fig. 16B, left panel and 15A). Telomerase activity was

also detected in extracts from 35S::TER1CC transformants, but not from wild type when the TER1CC reverse primer was employed (Fig. 16B, right panel and 15A).

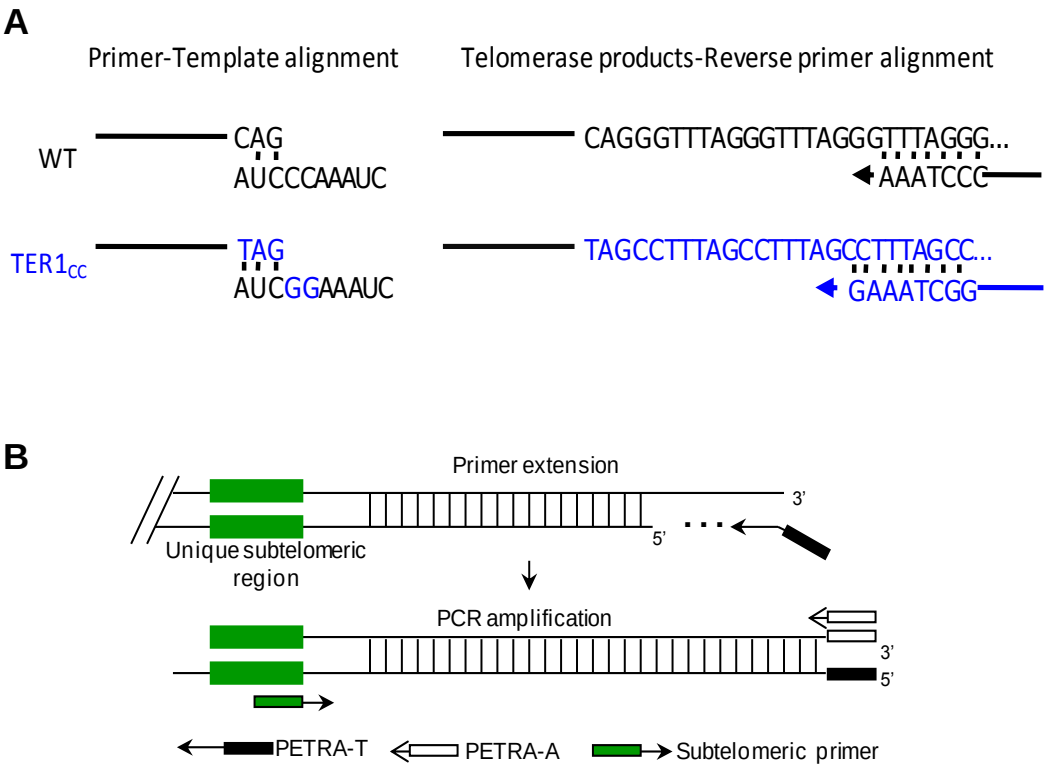
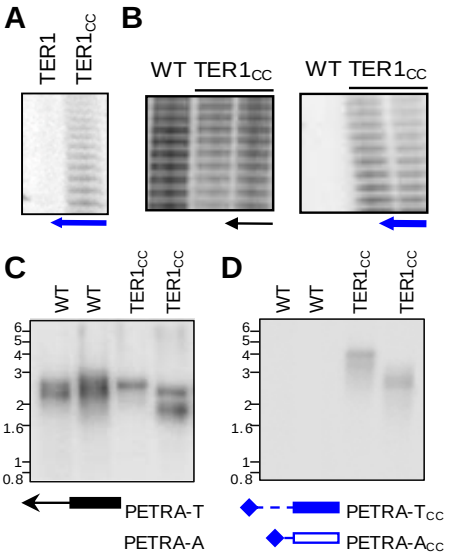


Figure 15. Experimental scheme to test TER1 telomere repeat incorporation *in vivo*. (**A**) Mutation of the telomere sequence in TER1. Wild type plants were transformed with $TER1_{CC}$ and were propagated for two generations before analysis. TRAP was performed with a reverse primer complementary to the mutant repeats (blue arrows). (**B**) Graphic depiction of PETRA assay (290). (**B**). Sequence information of PETRA products.



E

GROUP	# of clones	
WT	85	...TTTAGGGTTTAGGGTTTAGGGTTTAGGGTTTAGGGTTTAGGGTTTAGGG 3'
Mutant single	15	...TTTAGGGTTTAGGGTTTAGGGTTTAGCCTTTAGGGTTTAGGG 3'
		...TTTAGGGTTTAGGGTTTAGCC(TTTAGGG)70-100TTTAGGG 3'
Mutant multiple	1	...TTTAGGGTTTAGCC(TTTAGGG)2TTTAGCC(TTTAGGG)2TTTAGCC(TTTAGGG)3 3'
Mutant/ Misincorp	49	...TTTAGGGTTTAGCC(C)2(TTTAGGG)2-4TTTAGCC(TTTAGGG)2 3'
		...TTTAGGGTTTAGCC(C)2-3 (TTTAGGG)2-4TTTAGCC(G)2-4(TTTAGGG)2-6 3'

Figure 16. TER1 directs telomere repeat incorporation *in vivo*. (**A**) *In vitro* reconstitution of telomerase with $TER1_{CC}$. TRAP was performed with a reverse primer complementary to the mutant repeats. (**B**) *In vivo* reconstitution of telomerase with $TER1_{CC}$. Wild type (WT) plants transformed with $TER1_{CC}$ were propagated for two generations and TRAP was performed with a reverse primer directed at WT (black arrow) or mutant repeats (blue arrow). (**C, D**) Telomere amplification by PETRA. Results for two WT plants and two third generation $TER1_{CC}$ transformants with either WT (black) or mutant (blue) PETRA primers. (**E**) Representative PETRA product sequences depicting single or multiple TTTAGCC (blue) repeats as well as misincorporation events (red) are shown.

Primer Extension Telomere Repeat Amplification (PETRA) (Fig. 15B and ref. (268) was then employed to ask if TER1 specifies the telomere repeat sequence on chromosome ends. Using a subtelomeric primer directed at the left arm of chromosome 1, PETRA products were obtained from wild type plants and transformants expressing 35S::$TER1_{CC}$ (Fig. 16C). More importantly, reactions with PETRA-T_{CC} generated products with 35S::$TER1_{CC}$, but not wild type plants (Fig. 16D).

PETRA products generated with the wild type PETRA-T primer were cloned from 35S::$TER1_{CC}$ transformants. Out of 150 clones, 85 (57%) contained perfect arrays of TTTAGGG repeats, consistent with the presence of wild type TER1 in these plants (Fig. 16E). Strikingly, 65 (43%) clones harbored one or more TTTAGCC repeats and in 49 cases, these repeats were characterized by the misincorporation of additional C or G residues (Fig. 16E), indicating that the $TER1_{CC}$ mutation decreased the fidelity of telomerase *in vivo*. Altogether, our findings indicate that TER1 assembles into an enzymatically active RNP and furthermore that TER1 determines the sequence of telomeres *in vivo*. The biological role of TER2 is currently under investigation and thus the remainder of this study focuses on analysis of TER1.

TER1 associates with POT1a in vitro

To further investigate the composition of *Arabidopsis* telomerase, we examined potential protein binding partners for TER1. T7-tagged KU70 or KU80 was co-expressed with TER1 in RRL (Fig. 17A). IP was conducted with anti-T7 agarose beads, the bound RNA was extracted, cDNA was generated using random pentadecamers and PCR was performed with primers specific for TER1. Only a residual amount of TER1 was detected in the KU70 or KU80 IP that was just above background with unprogrammed RRL or the IP with an unrelated protein, TRFL4 (Fig. 17A).

We next asked whether POT1a or POT1b interact with TER1. Tagged POT1a or POT1b was co-expressed with TER1 in RRL (Fig. 17A), followed by IP and PCR as discussed above. TER1 was enriched in the POT1a IP, but not in the IP with POT1b or TRFL4 (Fig. 17A). Reactions with antisense TER1 or wild type TER2 failed to show an interaction with POT1a (Fig. 17A), indicating that TER1-POT1a binding is specific. Gel shift assays produced results consistent with the IP: TER1 formed an RNP complex with POT1a, but not POT1b (Fig. 17B). Competition experiments confirmed the specificity of TER1-POT1a interaction and furthermore demonstrated that POT1a does not associate with TER2 *in vitro* (Fig. 17B).

Deletional mutagenesis was used to map the POT1a-TER1 interface *in vitro*. A combination of IP and filter binding assays were performed with recombinant POT1a and 5' end labeled TER1 to identify the RNA binding site on

POT1a (Fig. 18A). As expected, TER1 binding was dependent on the OB-folds and not the C-terminal domain of POT1a. OB-1 was sufficient for TER1 binding, but a more robust interaction was observed with constructs containing both OB-1 and OB-2 (Fig. 18A). Filter binding was performed with truncated TER1 constructs to define the POT1a binding site (Fig 18B). POT1a bound within a 5' 242 nt segment unique to TER1 that lies upstream of the telomere template sequence (Fig. 11C and 18B).

Finally, filter binding was used to determine the affinity of POT1a for TER1 *in vitro* (Fig. 18C). In this experiment, recombinant POT1a was incubated with decreasing concentrations of TER1 transcribed *in vitro* with a radiolabeled tracer. The fractions of bound and free RNA were determined to calculate the Kd for the POT1a-TER1 interaction (Fig. 18C). Notably, the dissociation constant for AtPOT1a-TER1 (Kd= 2.1 X 10-7 M) is similar to value obtained for purified mammalian POT1 with single-strand telomeric DNA (10-7-10-8 M) (121, 291).

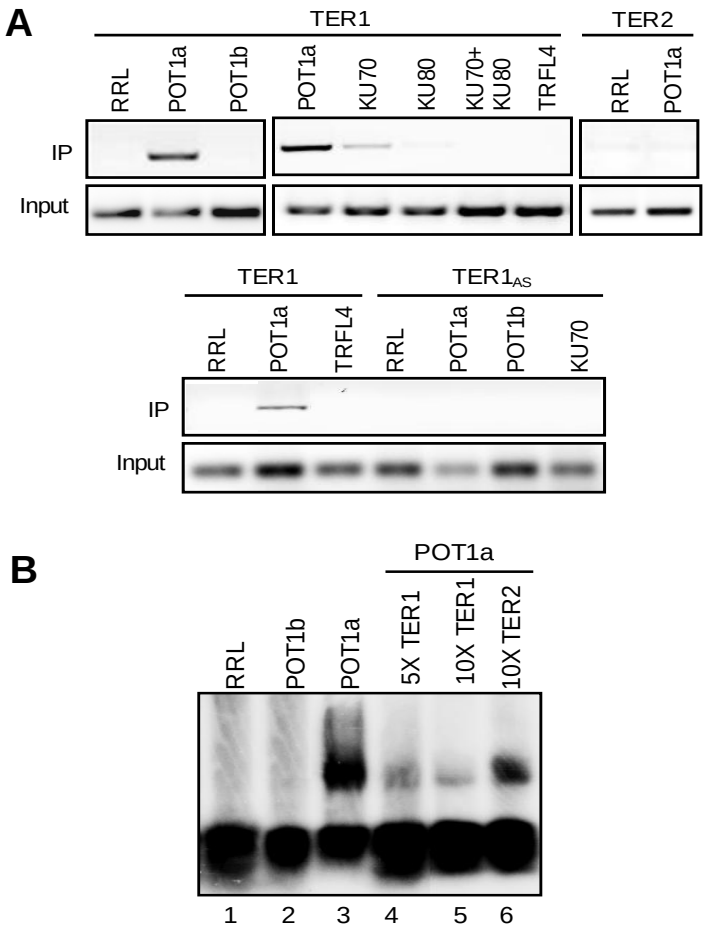


Figure 17. TER1 protein interactions *in vitro.* (**A**) *In vitro* TER1 binding assays with recombinant POT1a, POT1b and KU (top). T7-tagged proteins were co-expressed with TER1 and RT-PCR was carried out after immunoprecipitation (IP). TRFL4, a double-strand telomeric DNA binding protein (292) and $TER1_{AS}$ served as controls. (**B**) Results from electrophoretic mobility shift assays.

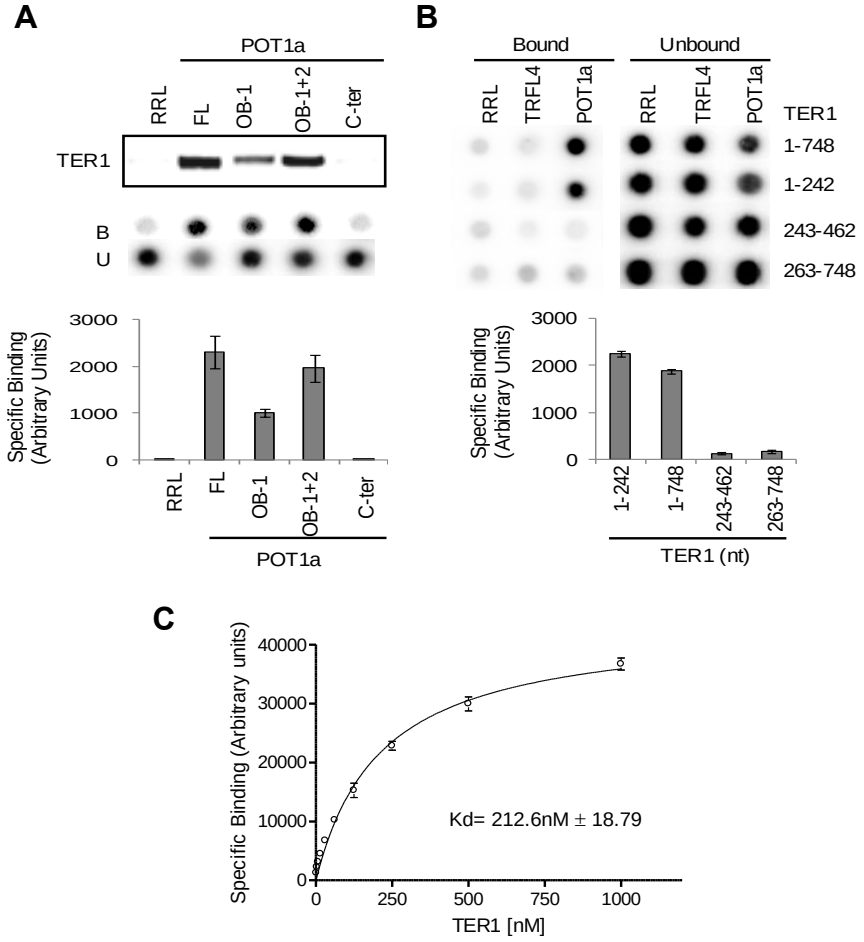


Figure 18. Characterization of the TER1-protein interaction *in vitro.* (A) Mapping the TER1 binding site on POT1a. Top, IP results with POT1a truncation mutants; middle, filter binding data; bottom, quantitation of filter binding data. FL, full length; OB-1, first OB-fold, OB-1+2, first and second OB-folds; C-ter, C terminus; B, bound; U, unbound. (**B**) Mapping the POT1a binding site on TER1. Top, filter binding data. TER1 truncations (in nts) are indicated on right; bottom, quantitation of filter binding data. (**C**) Binding isotherm for the POT1a-TER interaction. Values normalized against nonspecific binding and expressed in arbitrary phosphorimager units. Data were plotted against TER1 concentration (nM) and the Kd value was calculated.

TER1 and POT1a assemble into an enzymatically active telomerase RNP in vivo

We employed IP assays to investigate whether TER1 and POT1a assemble into an RNP *in vivo*. As expected, telomerase activity was recovered from the α-TERT and α-dyskerin IP samples, but not the pre-immune serum or α-histone H3 antibody control (Fig. 19A, top). Western blot analysis confirmed that target proteins were enriched in their corresponding precipitates (Fig. 19A, bottom). Attempts to quantify telomerase activity by Q-TRAP were unsuccessful, likely due to interference by the protein A agarose beads in the fluorescence intensity reading. TER1 was enriched in the α-TERT and α-dyskerin precipitates, but not α-histone H3 (Fig. 19B and 20C). The RNA-protein interaction was specific; U6 snRNA was not detected in any of the IPs (Fig. 20C). A low level of telomerase activity was observed in α-KU70 and α-POT1b pull-downs (Fig. 19A), and consistent with our *in vitro* binding results, only a background level of TER1 was observed (Fig. 19B and 20C). In contrast, abundant telomerase activity and TER1 was detected in the α-POT1a IP (Fig. 19A-B and 20C). Altogether, these data indicate that that TER1 assembles with TERT, dyskerin and POT1a in an enzymatically active RNP *in vivo*.

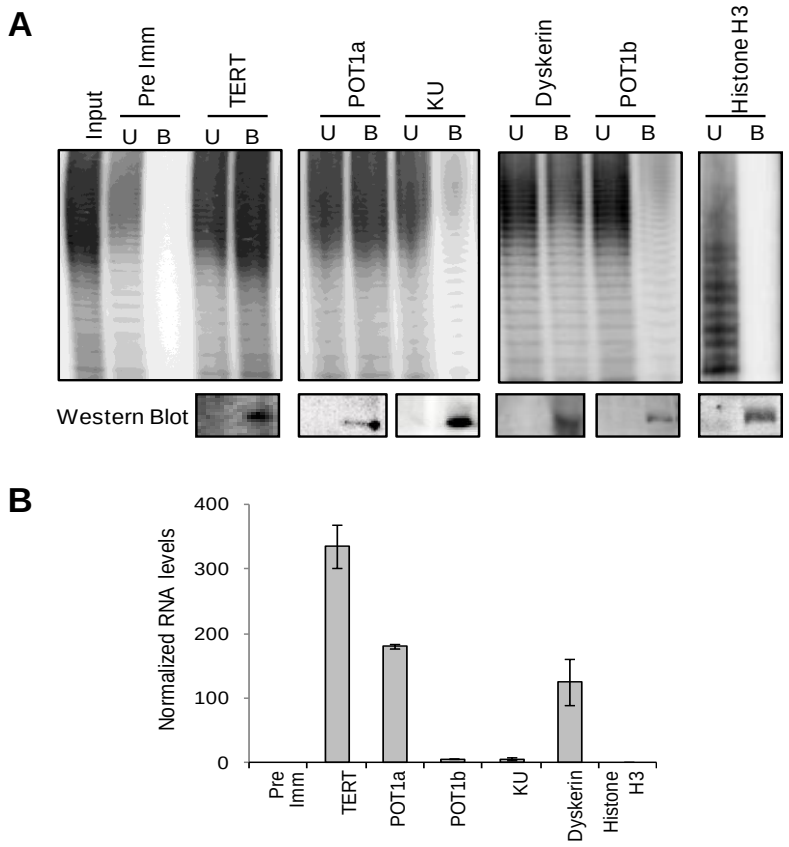


Figure 19. TER1 interactions *in vivo*. Immunoprecipitation was carried out with *Arabidopsis* cell culture extracts using the indicated antibodies. Pre-immune serum, and anti-histone H3 antibodies were used as negative controls. (**A**) Top, TRAP results with unbound (U) and bound (B) IP material. Bottom, Western blot analysis of IP fractions. (**B**) Quantitative RT-PCR results following IP. RNA levels were normalized to the primer efficiency, the levels of U6 snRNA, the pre-immune control and the antibody efficiency.

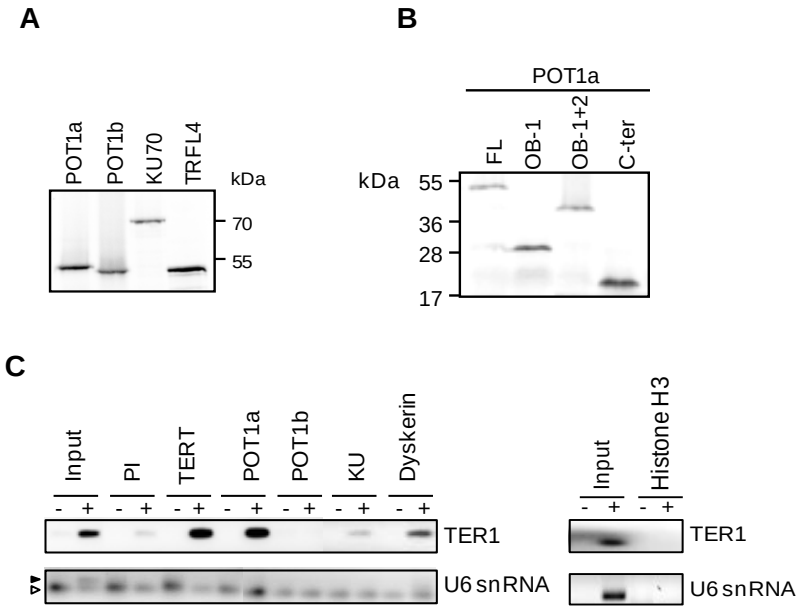


Figure 20. TER1-protein interactions. (A-B) RRL expression of ^{35}S labeled protein. **(A)** Full length proteins. **(B)** POT1a domains. FL, full length; OB-1, first OB-fold; OB-1+2, first and second OB-folds; C-ter, Carboxyl terminus. **(C)** IP was carried out with *Arabidopsis* cell culture extracts using the indicated antibodies. Pre-immune serum and anti-histone H3 antibodies were used as controls. End point RT-PCR from immunoprecipitated material. (▶) Amplification of U6; (▷ Primer dimmers.

Discussion

Arabidopsis *encodes two TERs, but only TER1 is required for telomere maintenance* in vivo

Of the more than 60 organisms, including yeast, ciliates and vertebrates, where telomerase RNA subunits have been reported thus far, only *Arabidopsis* harbors two TER molecules encoded by distinct genetic loci. Both *Arabidopsis* TERs are enriched in purified telomerase fractions; both associate with TERT; and both serve as templates for telomere repeat synthesis *in vitro*. Indeed a 150nt “mini-T” constituting a region conserved in both TER1 and TER2 (R1) is sufficient to promote telomerase activity *in vitro*. Strikingly, this catalytic core is approximately the same size as *Tetrahymena* TER (14) and the minimal TERs defined for human (210nt) (293) and *S. cerevisiae* (170nt) (294).

Despite these similarities, the *Arabidopsis* TERs differ dramatically in overall nucleotide sequence, protein binding partners and roles *in vivo*. TER1 and TER2 are so divergent in regions flanking the conserved catalytic core that they cannot be reliably aligned. Moreover, the telomere template is located in different positions relative to the 5’ terminus. Experiments with circularly permutated *Tetrahymena* TER show that the template along with a few adjacent 3’ residues can be moved to different sites without affecting enzyme properties *in vitro* (295). Thus, the different location of the templating domains in TER1 and TER2 may have little impact on their function *in vivo*. On the other hand, the unique sequences in TER1 and TER2 support distinct protein binding partners.

POT1a, but not POT1b, specifically associates with TER1 and not TER2. Thus, the *Arabidopsis* TERs, like their counterparts in other eukaryotes, are modular and flexible frameworks for protein interactions. Finally, only TER1 functions as a canonical telomerase RNA subunit to promote telomere maintenance *in vivo*.

POT1a is a novel TER1 binding protein

Identification of the templating RNA for *Arabidopsis* telomerase provided insight into the protein composition of this RNP. We found that TER1 associates with both TERT and dyskerin *in vivo*, suggesting that like human telomerase (29), the core telomerase RNP in *Arabidopsis* is comprised of TERT, TER and dyskerin. We did not detect a robust interaction between Ku and TER1. This observation is consistent with a role for TER1 in promoting telomere replication, since *Arabidopsis* Ku is a potent negative regulator of telomere length (296).

Our work also sheds new light on AtPOT1a, a hitherto enigmatic telomeric protein. Repeated attempts to detect telomeric DNA binding by POT1 proteins from *Arabidopsis* and other related plant species have been unsuccessful (130-131), indicating that POT1 likely evolved to bind a different substrate than telomeric DNA in higher plants. Supporting this prediction, AtPOT1a is not required for chromosome end protection and instead physically interacts with telomerase, acting in the same genetic pathway as TERT (129-130). We show that POT1a associates with telomerase via a direct interaction with TER1.

The migration of POT1 from the telomere to telomerase appears to be a relatively recent event. POT1 proteins from algae, moss, maize and Asparagales bind telomeric DNA *in vitro* (132, 274). Moreover, moss POT1, like its counterparts in yeast and vertebrates, is required for chromosome end protection (274). Our data suggest that the first OB-fold in AtPOT1a provides the major contacts for TER1 recognition. Similarly, the crystal structure of human POT1 reveals several residues in OB-1 that lie in close proximity to telomeric DNA (29). One of these, F62, plays a critical role in distinguishing DNA from RNA binding (272). Intriguingly, the corresponding residue is conserved in AtPOT1a, implying that the substrate for POT1a is not simply telomeric RNA, a conclusion supported by our binding studies which show POT1a recognizes a unique region of TER1 upstream of the templating domain. Finally, the Kd we obtained for the AtPOT1a-TER1 interaction is similar to the dissociation constant of purified mammalian POT1 with single-strand telomeric DNA (120-121). Altogether these data suggest that the switch from telomeric DNA to TER may reflect subtle remodeling of the nucleic acid binding pocket in POT1, and involve co-evolution of POT1a with TER1.

Duplication of telomerase subunits

Duplication of telomerase components has driven diversification of function in other eukaryotes. An instructive example is found in the ciliate *Euplotes*, which encodes three differentially expressed TERT proteins (227).

These TERTs assemble with a single TER into distinct RNPs that are postulated to promote a switch in telomerase specificity as the enzyme transitions from de novo telomere formation to maintenance of pre-existing telomeric DNA tracts. The duplication of TER in *Arabidopsis* has likewise led to diversification in function, although the specific role of TER2 is not yet understood. POT1 proteins are also duplicated in *Arabidopsis* and we have now shown that one of these duplicates, POT1a, is associated with the telomerase RNP through interaction with TER1. Altogether, our data indicate that gene duplication has provided the necessary components for assembly of distinct RNPs in *Arabidopsis*, potentially fueling the emergence of novel regulatory mechanisms.

Materials and Methods

Purification of telomerase

Arabidopsis suspension culture cells were maintained as described previously (297). Telomerase was purified from *Arabidopsis* cell culture nuclear extract. Briefly, cells were collected from 4.5 L of cell culture and ~320g of dry tissue were ground in liquid nitrogen. The resulting powder was resuspended in nuclei isolation buffer: 50mM Tris HCl pH 8.0, 5mM EDTA, 10mM KCl, 250mM sucrose, 1.5mM MgCl2, 0.3% Triton X-100, 1mM DTT, 1mM spermine, 1 mM spermidine, 10mM ribonucleoside vanadyl complex, 1X complete protease inhibitors (Roche), 0.4mM Pefablock SC Plus (Roche) and spun at 4000xg for 20 min. The pellet was resuspended in nuclei isolation buffer containing 1%

Triton X-100 and spun at 2000xg for 1 min and then at 4000xg for 1 min. The pellet was resuspended in nuclei extraction buffer: 20mM Hepes pH 8.0, 1.5mM MgCl2, 0.2mM EDTA, 300mM NaCl, 10% glycerol, 1% triton X-100, 0.1% NP40, 5mM DTT, 10mM ribonucleoside vanadyl complex, 1X complete protease inhibitors (Roche), 0.4mM Pefablock SC Plus (Roche), incubated with rotation at 4°C for 30 min and spun for 15 min at 14000xg at 4°C. The supernatant was collected, loaded into a 10 ml Q-sepharose column (Amersham Pharmacia) and eluted using a 100mM-1M NaCl pH 8,4 step gradient at 2ml/min. The fractions were collected and tested for telomerase activity by TRAP. Fractions with peak of activity were dialyzed against 100mM NaCl, loaded in a 5 ml Heparin Agarose column (Amersham Pharmacia) and eluted using a 100mM-1M NaCl pH 8,0 gradient at 2 ml/min. Fractions were collected and tested for telomerase activity by TRAP. Fractions corresponding to the peak of telomerase activity were pooled, concentrated to 500µl and loaded onto a Superose 6® size exclusion column (Amersham Pharmacia) which was calibrated with size markers run under the same conditions (20 ml 100mM NaCl pH 8,0 elution). Molecular weight markers were run under the same conditions: BSA (67kDa), Aldolase (158kDa), Ferritin (440kDa) and Tyroglobulin (669kDa).

RNA isolation

RNA was extracted from heparin column fractions with peak activity using a buffer containing 50mM Tris HCl pH 9.0, 100mM NaCl, 2% SDS, 10mM EDTA

and 20mM β-Mercaptoethanol and standard acid phenol:chloroform extraction was performed. Samples were precipitated at -80°C in ethanol, sodium acetate and glycogen and spun at 14000xg for 30 min at 4°C. The pellet was washed with cold 70% ethanol, air dried at room temperature and resuspended in Tris-EDTA buffer pH8.0. RNA was 3' end labeled with Cytidine-3',5' - bis (phosphate), [5'-^{32}P] (Molecular Probes) in a standard overnight T4 RNA ligation reaction and resolved in a 6% acrylamide gel. The area of the gel was divided in 12 slices based on the PCP-labeled RNA pattern. RNA was eluted from each band and RT-PCR was performed using penta-decamers for first strand synthesis using SuperScript III® reverse transcriptase (Invitrogen). The cDNA was sequenced using seven different primers corresponding to the seven possible permutations of the *Arabidopsis* telomeric repeat. One sequence was obtained and was used in a BLASTN search against the nucleotide collection database (NCBI) and the *Arabidopsis* transcript database (TAIR8 –introns + UTRs) (Fig. S2).

RNA 5' and 3' mapping

To map 5' and 3' ends of TER1 and TER2, a unique DNA linker (miRNA linker-1®, IDT) containing a 5'-5' linked adenylate (rApp) to activate ligation and a 3' dideoxynucleotide (ddCTP) to prevent ligation to the 3' end was ligated onto the 3' end of total RNA extracted from the peak activity fractions from the heparin column using T4 RNA ligase (Epicentre) in the absence of ATP (Fig.

S3). The reaction was incubated overnight at 16°C. After enzyme inactivation, the reaction was purified by gel filtration to reduce non-ligated linker contamination. cDNA was synthesized using an oligonucleotide complementary to the 3'-end linker and then circularized using CircLigase® (Epicentre) following the manufacturer's instructions. PCR was performed using the primers TER1#301fwd: 5'-ACAGAGAACGATGTTCCAACT-3' and TER1 template rev: 5'-CTCC TTGAGAATCTCAGCGAGT-3'; and primer #33: 5'-AACAGAACCAGAGAACGTTG-3' and #39: 5'-TGTAAGCGTAGGGTTTAGT TGTCGTC-3' for TER2. Products were cloned into pDrive (Qiagen) by PCR cloning and sequenced using M13 primers.

Northern blotting

Northern blotting was performed with total RNA extracted from *Arabidopsis* suspension culture. Briefly, RNA was denatured for 2 min at 95ºC in formamide loading buffer containing 5M urea and resolved in a 4% acrylamide gel under denaturing conditions. T7 *in vitro* transcribed TER1 was used as control. Then, RNA was transferred to a Hybond-N+® membrane (Amersham) for 10h at 30min 0.5X TBE. After 1h pre-hybridization the membrane was hybridized for 12h at 40ºC with a pool 5'-end 32P-ATP labeled oligonucleotides complementary to TER1. After washes the membrane was exposed to a phosphor screen.

Primer extension

Primer extension was carried out on total RNA extracted from *Arabidopsis* cell culture. 0.25pmol of 5' end labeled oligonucleotide was incubated with total RNA at 95°C for 5 min and allowed to anneal in two sequential 15 min incubations at 72°C, and 60°C, after which extension mix (50mM Tris-HCl pH 8.3, 15mM KCl, 3mM $MgCl_2$, 5% DMSO, 1mM DTT, 1mM dNTPs, 1.5U RNAseOUT® and 200U SuperScript III® reverse transcriptase (Invitrogen)) was added to a 30µl final volume. The reactions were incubated at 58°C for 3 h. The enzyme was inactivated at 80°C for 10 min. The RNA was hydrolyzed by incubation at 70°C with 15µl of 1N NaOH for 10 min. The reaction was neutralized and precipitated with 15µl of 1N HCl, 20µl 3M NaOAc pH 5.2, ethanol and glycogen. The products were resolved by denaturing PAGE.

In vitro *telomerase reconstitution*

A TERT-pET28a plasmid with an N-terminal T7 tag was used for telomerase reconstitution experiments. Reactions were assembled with 100ng of TERT-pET28a plasmid and 0.5pmol or 0.1pmol of gel purified DNA template encoding TER1 or TER2 respectively, driven by a T7 promoter, in a mix containing Rabbit Reticulocyte Lysate (RRL) (Promega), amino acids, RNase inhibitors, and T7 RNA polymerase. Reactions were incubated for 90 min at 30°C. T7 agarose beads (Novagen) were blocked with buffer W-100 (20mM TrisOAc [pH 7.5], 10% glycerol, 1mM EDTA, 5mM $MgCl_2$, 0.2M NaCl, 1% NP-

40, 0.5mM sodium deoxycholate, and 100mM potassium glutamate) containing 0.5mg/ml BSA, 0.5mg/ml lysozyme, 0.05mg/ml glycogen, 1mM DTT and 1µg/ml yeast tRNA. The reconstitution reaction was mixed with the beads to a 200µl final volume and incubated for 2 h at 4°C with rotation. Beads were washed 6X with 800 µl of W-400 buffer (W-100 containing 400 mM potassium glutamate) and 3X with 800µl of TMG buffer (10mM TrisOAc [pH 7.5], 1mM MgCl2, and 10% glycerol). After the final wash, beads were resuspended in 30µl of TMG. 2µl of beads were used for TRAP assays as previously described (130, 242).

In vitro *binding assays*

For co-IP experiments, POT1a, POT1b, KU70 and KU80 cloned in pET28a with a T7 tag were co-expressed with TERs in RRL as described above. After IP RNA was extracted and RT-PCR was performed. Electrophoretic mobility shift assays used RNA transcribed *in vitro* with T7 RNA polymerase and [γ-32P]-CTP labeled TER1 and TER2. Binding reactions contained 3µl of RRL expressed protein, 0.1pmol of 32P labeled TER and 1X binding buffer (25mM Tris-HCl pH 8.0, 10mM $Mg(OAc)_2$, 25mM KCl, 10mM DTT and 5% Glycerol) in a 30µl final volume. 1µM yeast tRNA and 0.5µM RNA $(U_3AG_3)_4$ were used as nonspecific competitors. After 20 min at 30°C the reaction was loaded onto a 0.8% agarose 0.5X TBE gel and run for 2 h at 70 volts at 4°C. Gels were dried and exposed to phosphorimager screens.

For double-filter binding assays, TERT expressed in RRL was incubated with decreasing concentrations of pre-folded RNA transcribed *in vitro* using 5-end 32P-labeled RNA as a tracer. Binding reactions contained 0.5µl of recombinant protein, pre-folded TER in binding buffer (50mM Tris-HCl pH 7.5, 200mM potassium glutamate, 0.5mg/ml BSA, 0.5mg/ml tRNA, 1mM MgCl2, 1mM DTT and 0.01% NP-40) in a 25µl final volume. After 30 min at 30°C, the reactions were filtered through nitrocellulose and nylon filters using a dot-blot apparatus (BioRad). The membranes were washed with 600µl washing buffer (50mM Tris-HCl pH 7.5, 200mM potassium glutamate, 1mM $MgCl_2$, 1mM DTT and 10% glycerol), dried, exposed to a phosphor storage screen and scanned after 2h. Equilibrium dissociation constants, Kd, were obtained by non-linear regression of the binding data fitted to a one-site binding model using Graphpad Prism® software.

End-point RT-PCR and quantitative RT-PCR

Total RNA was extracted from 0.5 g floral or other tissue using Tri Reagent (Sigma). cDNAs were synthesized from total RNA using Superscript III reverse transcriptase (Invitrogen). Random pentadecamers were incubated with 2 µg of total RNA in the supplied buffer at 65°C for 5 min. Reverse transcription (RT) was carried out with 100U of Superscript III at the following temperatures 37°C for 20 min, 42°C for 20 min and 55°C for 20 min. Enzyme was inactivated

at 80°C for 10 min and RNA was degraded with RNase H (New England Biolabs). 1.5 µl of cDNA was used in PCR.

For real-time RT-PCR, 2 µl of the above cDNA was used at a 1:10 dilution in a 20µl reaction containing 10 µl of SyBr green master mix (NEB) and 2 µl of each primer (2µM). PCR was performed for 40 cycles with 30 s at 95°C and 60 s at 60°C. Threshold cycle values (Ct) were calculated using an iCycler iQ thermal cycler (BIO-RAD) and the supplied Optical System Software.

Quantitative RT-PCR data analysis

Amplification efficiencies were calculated for each primer pair in a 5 point titration curve. The slope was calculated from a standard curve where Ct was on the y-axis and log(cDNA dilution factor) on the x-axis. The corresponding real-time PCR efficiency (E) was calculated according to the equation: E=(10-1/slope)-1. To correct for intra-assay and inter-assay variability, each sample was evaluated by tripicate within one run in at least three different experimental runs. The relative expression level (R) was calculated as follows: R = (Etarget)ΔCttarget(control-sample) / (Ereference)ΔCtreference(control-sample) as previously described (4). U6 snRNA and β-actin were used as reference. Normalization to the pre-immune control and to the efficiency for each antibody was used for RNA quantification in the pull-down samples. Primers used for real-time PCR are as follows: TER1 Q4F:5'-CCCATTTCGTGCCTATCAGACGAC-3'. TER1 Q4R: 5'-TCTCCGACGACCATTCTCTCGATAC-3'; TER2#38: 5'-GACGA

CAACTAAACCCTACGCTTACA-3' and TER2#40: 5'-CAGGATCAATCG GAGAGTTCAATCTC-3'; TER2S: TER2#38 and TER2$_S$# 193: 5'-CCCCATCTC CGACGAGACGAC-3'; TERT Q3F: 5'-ACCGTTGCTTCG TTGTACTTCACG-3' and TERT Q3R: 5'-CGACCCGCTTGAGAAGAAACTCC-3'; U6-1F: 5'-GTCCC TTCGGGGACATCCGA-3' and U6-1R: 5'-AAAATTTGGACCATTTCTCGA-3' β-Actin 2F: 5'-TCCCTCAGCACATTCCAGCAGAT-3' and β-Actin 2R: 5'-AACGATT CCTGGACCTGCCTCATC-3'.

Plant materials and genotyping

Arabidopsis seeds with a T-DNA insertion in TER2 (SAIL_556_A04) were obtained from the *Arabidopsis* Biological Resource Center (Ohio State University, Columbus, OH). Seeds were cold-treated overnight at 4°C, and then placed in an environmental growth chamber and grown under a 16-h light/8-h dark photoperiod at 23°C. Plants were transformed using the in planta method (5). For genotyping, DNA was extracted from flowers and PCR was performed with the following sets of primers: for TER1, LP: 5'-GAAAGACCTCAGCATCAGTGC-3' and RP: 5'-GGACTTTTTGAAAACAATTAC AAATC-3'; for TER2, primer #38: 5'-GACGACAACTAAACCCTACGCTTACA-3' and #45: 5'-CGATGTTGTTTTTCTGCTTAGGACACA-3'. To amplify mutant TER2 alleles containing a T-DNA insertion, the T-DNA specific primer was used along with TER 8526-01 fwd: 5'-GAGACGCAGCGAGCGATAGCCGATAG-3' primer.

Antisense constructs

Full-length TER1 in the antisense orientation (TER1FLAS) and a truncated antisense version of TER1 (nucleotides 470-748) TER1AS were cloned into the gateway destination vector pK7WG2. The constructs were introduced into Agrobacterium tumefaciens strain GV3101. TER1ASFL was transformed into *ter2-1* homozygous plants and TER1AS was transformed into wild type plants using the in planta method. Transformants were selected on Murashige and Skoog (MS) basal medium supplemented with kanamycin (50 µg/ml).

Site-directed mutagenesis and plant transformation

To generate template mutations in the template region of TER1, site-directed mutagenesis was performed with Pfu turbo polymerase (Stratagene) on TER1-pDONR221 using the primers M1: 5'-GCCTATCAGACGACAACTAAAGG CTACACGCTTACA-3' and M2: 5'-TGTAAGCGTGTAGCCTTTAGTTGTCGTCT GATAGGC-3' according to the manufacturer's guidelines. The mutation was confirmed by sequencing. TER1CC was cloned into the destination vector pB7WG2 and transformed into wild type plants. After transformation the seeds were selected in MS agar containing kanamycin at 50µg/ml.

TRF analysis, TRAP and Q-TRAP assays

TRF and TRAP assays were performed as previously described (2-3). TRAP products from mutant TER1CC telomerase were amplified using the primers CC TRAP forward: 5'-CACTATCGACTACGCGATTAG-3' and CC TRAP reverse: 5'- GGCTAAAGGCTAAAGGCTAAAG-3' (Fig. S4A). Q-TRAP was performed as previously described (242, 298-299).

Antibodies, immunoprecipitation and western blotting

AtKU70 antibodies were kindly provided by Dr. Karel Riha (Gregor Mendel Institute, Vienna). AtTERT and AtPOT1a antibodies have been previously described (130, 242). The anti-dyskerin polyclonal antibody was raised in rabbits against recombinant full-length AtNAP57 expressed in E. coli. The POT1b is an affinity-purified peptide antibody (Covance). IP efficiency was calculated for each antibody using 35S labeled protein expressed in rabbit reticulocyte lysate. Western blotting was performed with a 1:2000 dilution of anti-KU70, anti-POT1a, anti-POT1b, anti-TERT and anti-Histone 3 antibodies (Upstate). The anti-dyskerin antibody was used at a 1:5000 dilution. Peroxidase-conjugated light chain-specific mouse anti-rabbit secondary antibody (Jackson Immunoresearch) was used at a 1:20,000 dilution. Following IP, RNA was phenol:chloroform extracted from the beads and subjected to RT using superscript III® reverse transcriptase (Invitrogen) and random pentadecamers.

CHAPTER III

ALTERNATIVE TELOMERASE RNPS IN *ARABIDOPSIS* WITH DISTINCT RNA AND PROTEIN COMPOSITION AND REGULATORY ROLES IN TELOMERE MAINTENANCE

Summary

The telomerase ribonucleoprotein is a highly regulated enzyme that stabilizes eukaryotic genomes by replenishing telomeric repeats on chromosome ends. The essential RNA subunit of telomerase, TER, functions as a template for telomere repeat synthesis by the catalytic subunit TERT. We previously showed that *Arabidopsis* encodes two TER subunits, TER1 and TER2. TER1 is the major template for telomere maintenance. Here we present the functional characterization of TER2. We show that TER2 is processed to form an additional TER isoform termed $TER2_S$. Like TER1 and TER2, $TER2_S$ levels peak in highly proliferative tissues, however the $TER2_S$:TER2 ratio is higher in flowers than in cell culture, indicating that TER2 processing is regulated in different settings. We further show that TER2 and $TER2_S$ assemble into RNP complex with proteins distinct from those associated the TER1 RNP. In particular, TER2, but not TER1 assembles with KU70 and POT1b. Finally, we demonstrate that TER2 and POT1b are both negative regulators of telomerase activity. Null mutations in either gene increase telomerase activity. In contrast, TER2 over-expression reduces the TER1 templating function and results in

shortened telomeres *in vivo*. These findings reveal a complex and novel regulatory pathway for telomerase that involves multiple RNPs composed of different protein subunits and three different TER isoforms.

Introduction

Telomeres are replicated by telomerase, an unusual ribonucleoprotein reverse transcriptase, composed at its core of a reverse transcriptase (TERT) and an RNA template (TER). The function of telomerase is to balance the loss of telomeric DNA due to the end replication problem. Telomerase activity is tightly controlled. Telomerase expression is critical for cell proliferation, but misregulation can lead to stem cell related diseases or carcinogenesis. Telomerase regulation occurs at several levels including intracellular and extracellular signaling pathways (300-301). Transcriptional regulation of hTERT is a major determinant of enzyme activity (193, 302), and in some cell types, hTER expression is controlled (194, 302). The stability and processivity of telomerase are modulated by RNP accessory factors (120, 195-199, 303). Many of these have roles in RNP biogenesis, assembly, recruitment and regulation of telomerase at the end of the chromosome.

The telomerase catalytic subunit TERT contains universally conserved reverse transcriptase motifs, which has allowed identification of this moiety in many organisms. In contrast, the nucleotide sequence of the TER subunit is highly divergent. TERs can be aligned with confidence only among closely

related species making their identification a challenging enterprise. We previously demonstrated that *Arabidopsis thaliana* encodes two TER subunits, TER1 and TER2 (see Chapter II). Both TER1 and TER2 assemble with TERT into active RNP complexes *in vitro*. However, only TER1 is required for telomere maintenance *in vivo*. Depletion of TER1 alone or TER1 plus TER2 leads to a similar decrease in telomerase activity and telomere shortening (Chapter II). In contrast, a null mutation in TER2 does not inhibit telomerase activity or result in telomere shortening. Finally, when plants are transformed with a TER1 construct bearing a mutation in the templating domain, mutant telomere repeats are incorporated onto chromosome ends (Chapter II). Thus, TER1 functions as the major template for telomerase *in vivo*. The role of TER2 has not been explored.

In conjunction with the rapid divergence of TER sequences, telomerase accessory proteins also vary widely. Biochemical purification led to the identification of p43 in the ciliate, *Euplotes (180)* and its homolog p65 in *Tetrahymena* (162, 304). These proteins are implicated in telomerase maturation and activity, but neither has an obvious ortholog in multicellular organisms. Dyskerin associates with telomerase RNP in plants (242) and vertebrates (184, 305-306) and is required for proper RNP biogenesis (307-308) in the nucleolus. In contrast, the budding yeast TER (Tlc1) does not bind dyskerin. Instead it associates with Sm-binding proteins (229). Deletion of the Sm-motif or loss of the members of the Sm complex leads to reduced levels of

TLC1 (229). Thus, the biogenesis pathway of TLC1 seems to parallel that of snRNAs.

The non-catalytic EST proteins promote telomerase activity in budding yeast (98, 303, 309-310). Est1p binds the TER subunit (Tlc1) and recruits telomerase to the chromosome ends through interaction with the ssDNA binding protein Cdc13p (Est4) (98, 246, 310). Est3p is also required for telomerase action *in vivo*, but its precise role is unclear (311). Finally, in both budding yeast and vertebrates TER is bound by the Ku70/80 heterodimer (182, 312), however Ku is not associated with TER in fission yeast (285). Thus, the telomerase enzyme is evolving, as evidenced by the existence distinct pathways for RNP biogenesis, dramatic differences in protein composition and remarkable divergence of TER sequence (147, 313) and Chapter II).

Besides TER, *Arabidopsis* encodes multiple copies of other telomere-related genes including the TRF family of putative double-strand telomere DNA binding proteins (127) and three highly divergent POT1 genes, AtPOT1a, AtPOT1b and AtPOT1c (129-130), A. Nelson and D. Shippen, unpublished). In other organisms, POT1 binds single-strand telomeric DNA and functions in chromosome end protection (28, 122, 314-316). Conversely, none of the POT1 proteins from *A. thaliana* or its close relatives bind telomeric DNA *in vitro* (131). Furthermore, AtPOT1a binds TER1 and physically associates with the TER1 RNP where it serves an essential role in promoting telomere maintenance (130, 313) and Chapter II). The functions of POT1b and POT1c are less clear. Over-

expression of POT1b results in chromosome fusion, implicating this protein in chromosome end protection. However, given the critical role of POT1a in the telomerase RNP and since a null mutation in POT1b has yet to be characterized, it is possible that the telomere deprotection phenotype associated with POT1b over-expression reflects a dominant negative phenotype. The role of POT1c has not been explored.

Here we present a functional analysis of TER2 and its protein binding partners. We show that TER2 undergoes developmentally regulated splicing to generate a novel TER isoform, termed TER2s. We also demonstrate that TER2 assembles with POT1b and Ku70/80 *in vivo*, forming an RNP that is distinct in both RNA and protein composition from the TER1 RNP. Finally, we demonstrate that the TER2 and its binding partner POT1b behave as a negative regulator for telomerase *in vivo*. These findings reveal a complex and novel regulatory pathway for *Arabidopsis* telomerase that may be fueled by the duplication and diversification of key components of the RNP.

Results

TER2 is processed to form TER2$_S$

We previously identified two TERs in *Arabidopsis*: TER1 and TER2 (317) (Fig. 21 and Chapter II). Unexpectedly, end-point RT-PCR analysis of RNA from cell culture amplified a shorter product in addition to TER1 and TER2 (Fig. 21 and 22B). Primer extension was performed with RNA from suspension cell

culture using an oligonucleotide complementary to a conserved region in R2 (Fig. 22A). A similar result was obtained. We detected major products corresponding to TER1 (445nt) and TER2 (740nt) as well as a 220nt product, which we termed $TER2_s$. The size of the TER2s primer extension product suggests that its 5' terminus corresponds to the 5' end of TER2, but additional mapping studies will be required to define the RNA 3' terminus. Cloning and sequencing of the $TER2_s$ RT-PCR product showed that it consists of the two conserved domains in TER2 (R1 and R2) joined together (Fig. 21). R1 contains the telomere template. Sequence complementarity analyses and RNA folding predictions indicate that $TER2_S$ is a bona-fide transcript and not an aberrant reverse transcription product (see Chapter V). In addition, bioinformatic searches did not retrieve any high similarity hit in addition to the TER1 and TER2, arguing that $TER2_S$ is not encoded in the genome. Intriguingly, the characteristic landmarks of consensus mRNA splicing machinery (branch point, 5' and 3' splice sites) are not detected in TER2 (Fig. 21), suggesting that it is generated by an unconventional RNA processing reaction.

Quantification by real time RT-PCR showed that like TERT (318), expression of all three TER isoforms peaks in cells with high telomerase activity (flowers and cell culture) (Fig. 22C). TER1 levels are consistently higher than TER2 in all samples examined. TER2 is more abundant than $TER2_S$ in vegetative organs (leaves and stems) (Fig. 22C). Strikingly, $TER2_S$ is highly enriched in flowers and is much more abundant than TER2. Here the level of

TER2s and TER1 are nearly equivalent (Fig. 22C). Conversely, in cell culture TER2s levels decline and TER2 is more abundant, but in this setting TER1 is more than 60-fold more than TER2. We conclude that TER2 expression and processing are developmentally regulated.

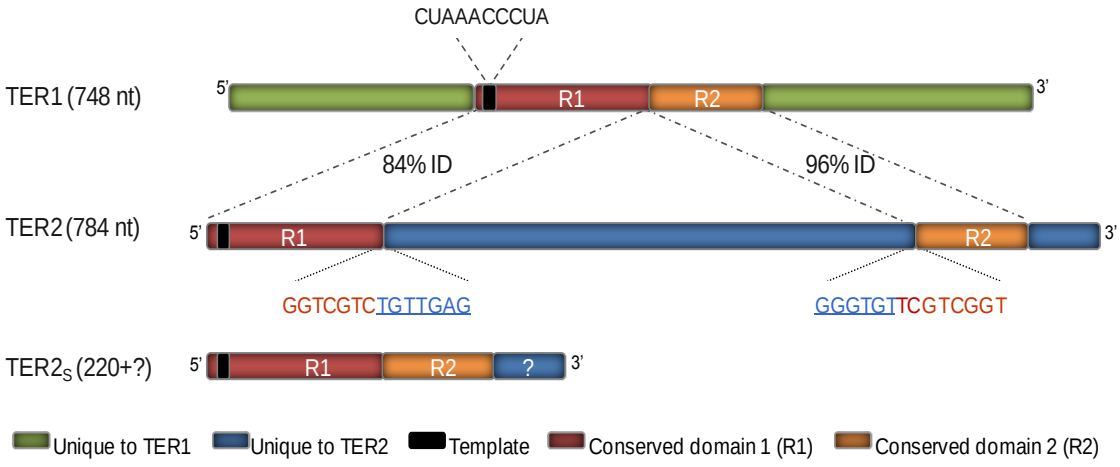


Figure 21. Diagram of TER1, TER2 and TER2$_S$. Templates, unique and conserved domains are shown. Transcript sizes as well as the sequence of the splice junctions are also indicated.

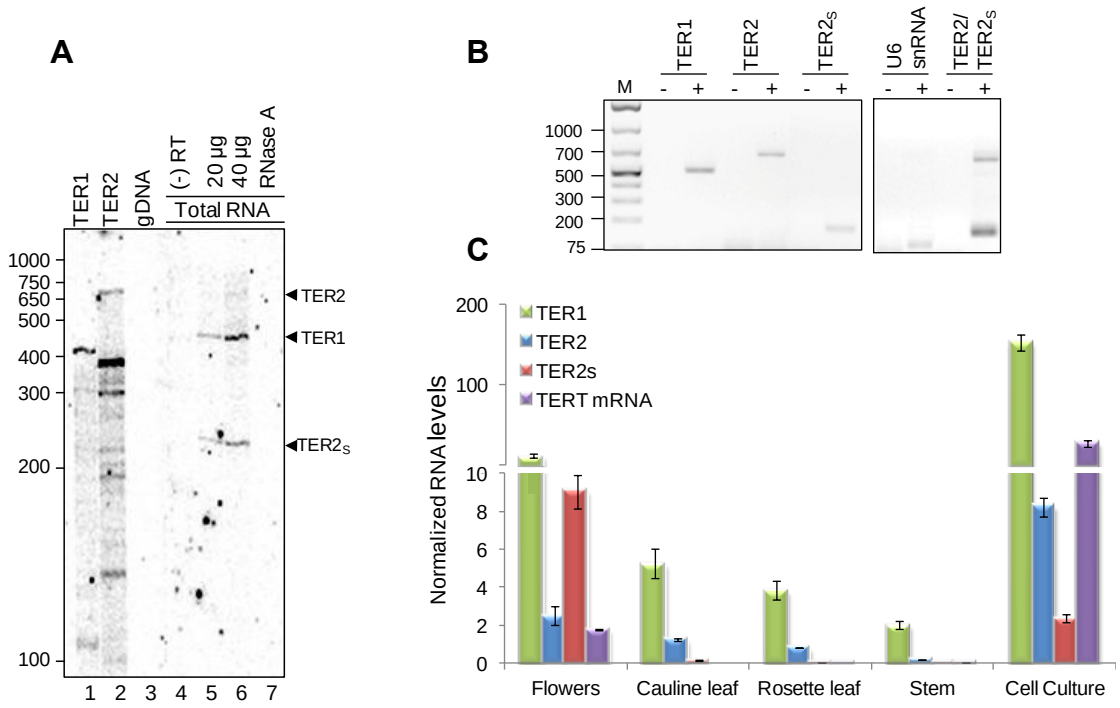


Figure 22. *Arabidopsis* encodes three telomerase-associated RNAs. (A) Primer extension of total RNA from cell culture resolved by PAGE. A primer complementary to a conserved in the R2 was used. Lane 1, *in vitro* transcribed TER1. Lane 2, *in vitro* transcribed TER2. Lane 3, 40 µg genomic DNA (gDNA). Lane 4, minus reverse transcriptase control. Lane 5, 20 µg of total RNA. Lane 6, 40 µg of total RNA. Lane 7, 40 µg of total RNA pre-treated with RNase A. (B) RT-PCR results with total RNA from cell culture. cDNA was generated using random pentadecamers. Odd lanes correspond to minus reverse transcriptase controls. Reactions with primers to detect TER1, TER2 and/or $TER2_S$ are shown in lanes 1-6. U6 snRNA was amplified as a control (lanes 7-8). Lanes 9-10, reactions with a primer pair that detects both TER2 and $TER2_S$. (C) Quantitative RT-PCR results for TER1, TER2, $TER2_S$ and TERT mRNA in different *Arabidopsis* tissues. RNA levels were normalized to U6 snRNA and to the efficiency of each primer pair to allow comparison.

TERT binds TER2 with a higher affinity than TER1 or TER2$_S$ in vitro

Our previous telomerase reconstitution experiments indicate that a construct corresponding to TER2$_S$ can direct telomere repeat synthesis *in vitro*. Thus, all three TER isoforms assemble with TERT into functional RNPs *in vivo*. To further investigate the interaction of TERT with different TER subunits, a double-filter binding assay was employed (Fig. 23). RRL expressed TERT was incubated with decreasing concentrations of *in vitro* transcribed TER1, TER2 and TER2$_S$. A corresponding radiolabeled tracer was used for detection. After incubation and filtration through nitrocellulose and nylon filters, the fractions of bound and free RNA were determined and used to calculate the Kd value for each interaction (Fig. 23A and B). The Kd values for TERT-TER1 and TERT-TER2 were 204.1 ± 11.3 nM and 22.6 ± 2.8 nM, respectively (Fig. 23B). Thus, the affinity of TERT for TER2 is an order of magnitude higher than for TER1. Preliminary experiments suggest that the Kd value for TERT-TER2$_S$ is approximately 1 μM (data not shown), which is more than ~45 and ~5-fold greater than the values for TER2 and TER1, respectively. Hence TERT prefers TER2>TER1>TER2$_S$. These differences in affinity indicate that TERT may preferentially bind TER2 *in vivo*.

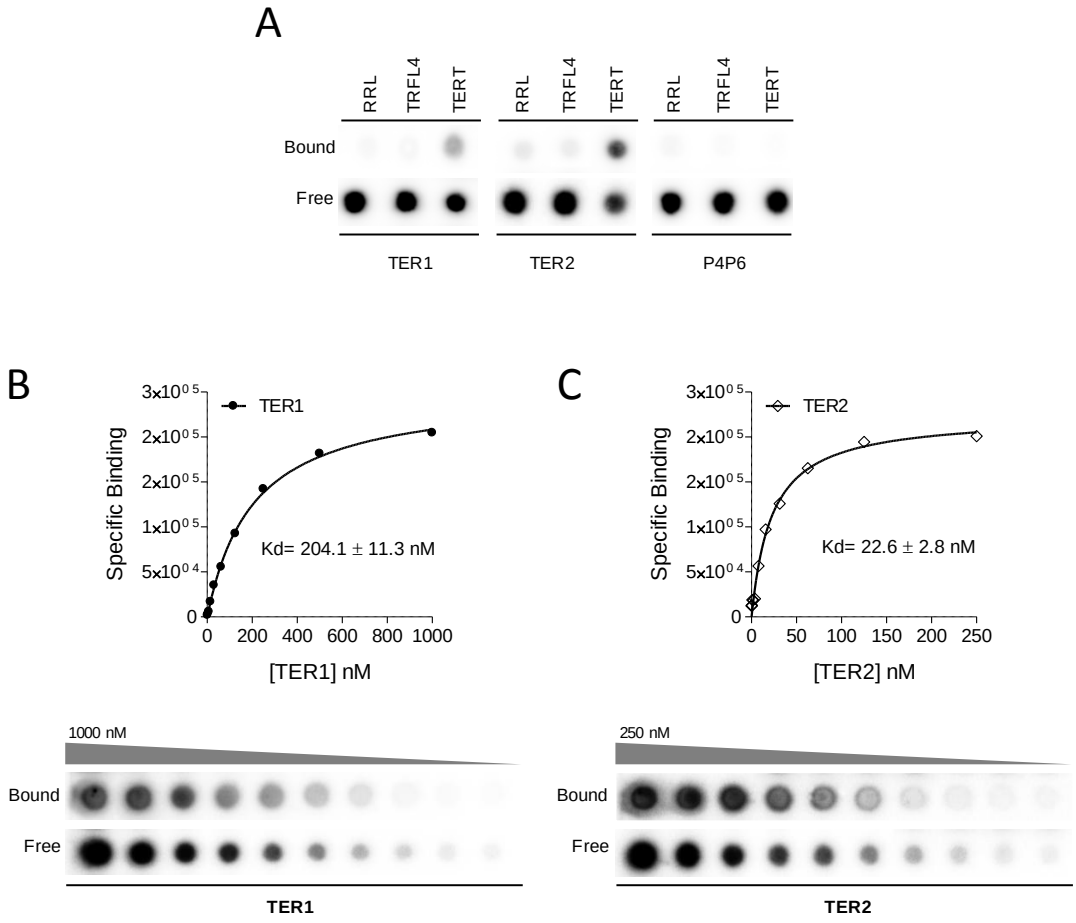


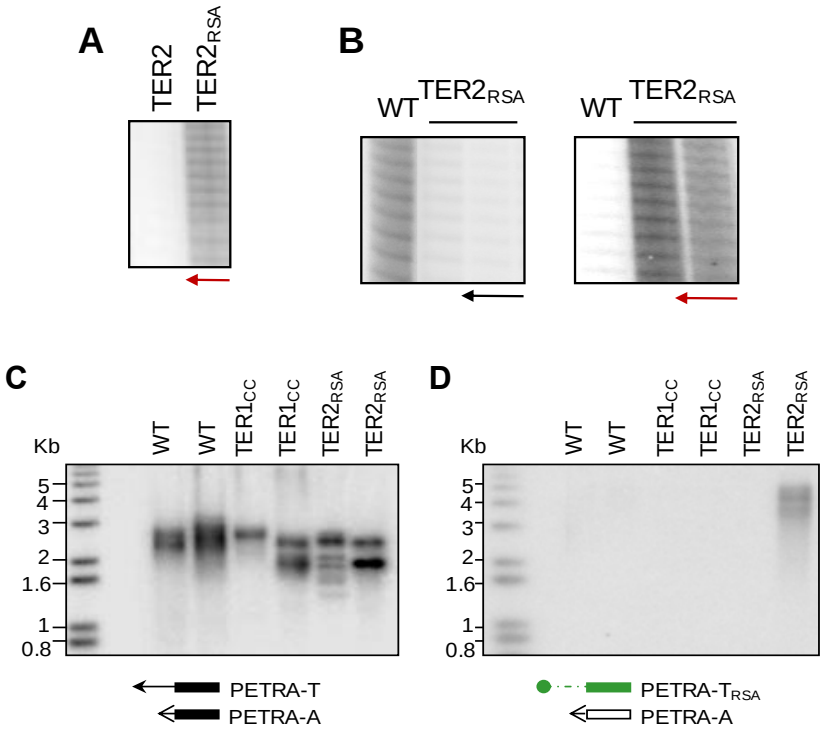
Figure 23. TERT-TER interactions. The affinity of the TERT-TER1 and TERT-TER2 interaction was determined by double filter binding. Top, TERT bound RNA. Bottom, Free RNA. (A) RNA was incubated with RRL alone or TRFL4 (a ds telomeric DNA binding protein) as controls for non-specific binding. Left panel, TER1. Middle panel, TER2. Right panel, the P4P6 domain of the *Tetrahymena thermophila* group I intron was used as control. (B-C) TER titration. TER was added in decreasing concentrations to a reaction containing a constant amount of TERT. (B) Top, Binding isotherm of TERT-TER1. Bottom, Representative filter binding assay. TER1 was titrated by a factor of 2 starting at 1000nM. (C) Top, Binding isotherm of TERT-TER2. Bottom, Representative filter binding assay. TER1 was titrated by a factor of 2 starting at 250nM. Kd values are shown.

TER2 can function as template for telomerase in vivo

Since TER2 can be used as a template for TERT reverse transcription *in vitro* (Ref and Chapter I), mutagenesis was employed to test if TER2 directs telomere repeat addition *in vivo*. The 5'-CUAAACCCUA-3' sequence in TER2 was mutated to 5'-CUAGUACCUA-3' (creating an RsaI restriction site, $TER2_{RSA}$). As expected, mutant ACTAGGT repeats were synthesized using this construct *in vitro* (Fig. 24A). $TER2_{RSA}$ was placed under the control of the powerful Cauliflower Mosaic Virus (CaMV) 35S promoter and transformed into plants null for endogenous TER2 (see Chapter II). Transformants were propagated for two generations before analysis to facilitate assembly of mutant RNA into telomerase complexes and to allow incorporation of mutant telomere repeats onto chromosome ends. As expected, TER1 levels were wild type and the mutant TER2 was over-expressed, as measured by RT-PCR (Fig. 26B). TRAP products were generated with the $TER2_{RSA}$ reverse primer (Fig. 24B, right panel), indicating that the mutant RNA assembled into an enzymatically active RNP *in vivo*. Unexpectedly, however, wild type telomerase activity, derived from endogenous TER1, was dramatically reduced in these plants (Fig. 23B, left panel and see below). This finding suggests that TER2 modulates TER1 function *in vivo*.

Primer Extension Telomere Repeat Amplification (PETRA) (268) was used to determine if TER2 is capable of directing the incorporation of telomere repeats onto chromosome ends. We monitored the telomeric sequence on the

left arm of chromosome 1 in plants expressing 35S::$TER2_{RSA}$ (Fig. 24C-E). As expected, PETRA products were obtained from both wild type and mutant plants using wild type PETRA-T (Fig. 24C). Control reactions with the PETRA-T_{RSA} primer failed to generate products for wild type plants or $TER1_{CC}$ mutants (Fig. 24D). Notably, PETRA products were detected in reactions with one of the two $TER2_{RSA}$ mutants analyzed. Sequence analysis revealed incorporation of a mutant repeat in 14% of the clones (7/50) (Fig. 24E). In all cases only as single mutant repeat was detecte. In one clone the mutant repeat was flanked by the sequence TATA, which is not encoded by the template. In another clone, the mutant repeat was adjacent to another partial mutant repeat (Fig. 24E). These findings indicate although the template mutation $TER2_{RSA}$ reduces fidelity of telomere repeat synthesis, TER2 is capable of directing telomeric DNA incorporation onto chromosome ends *in vivo* although with low efficiency.



E

GROUP	# of clones	SEQUENCE
WT	43	...TTTAGGGTTTAGGGTTTAGGGTTTAGGGTTTAGGGTTTAGGGTTTAGGG 3'
Mutant single	5	...TTTAGGGTTTAGGGTTTAGGGTTCATGGTTTTAGGGTTTAGGG 3'
		...TTTAGGGTTTAGGGTTCATGGT (TTTAGGG)$_{70-100}$TTTAGGG 3'
Mutant/ Misincorp	2	...TTTAGGGTCATGGT TATAGGG(TTTAGGG)$_2$TCATG (TTTAGGG)$_3$ 3'

Figure 24. TER2 serves as a template *for Arabidopsis* telomerase *in vivo.* (A) TRAP from TER2$_{RSA}$ *in vitro* reconstituted complexes. (B) TRAP results for 35S::TER2$_{RSA}$ transformants. Black arrow, reverse primer complementary to WT telomere repeats. Red arrow, reverse primer complementary to TER2$_{RSA}$ mutant repeats. (C-D) Results of amplification are shown with PETRA T and PETRA-A to detect WT telomeric DNA (C), and results with PETRA-T$_{RSA}$ and PETRA-A primer to detect TER2$_{RSA}$ products (D). (E) Representative PETRA product sequences depicting single or multiple ACTAGGT (blue) repeats as well as misincorporation events (red) are shown.

TER2 is a negative regulator of telomerase activity in vivo

In contrast to plants with reduced TER1, *ter2-1* homozygous mutants did not exhibit telomere maintenance defects in first or second generation mutants (Fig. 25B and Chapter II). As part of the characterization of *ter2-1*, Q-TRAP was performed to measure telomerase activity levels. Unexpectedly, we observed a dose-dependent increase in telomerase activity with a 2.1-fold increase in *ter2-1* G1 heterozygotes and 2.8-fold in G1 homozygotes, relative to wild type. A 3.7-fold increase in telomerase activity was detected in G2 *ter2-1* homozygotes (Fig. 25D). Telomerase activity was confined to organs where telomerase is normally expressed (flowers); enzyme levels remained low or undetectable in leaves (Fig. 25D). Hence, depletion of TER2 did not alter the developmental regulation of telomerase expression. These findings indicate that TER2 is a negative regulator of telomerase activity.

TER2 over-expression should decrease enzyme activity if TER2 negatively regulates telomerase. As predicted, an 11-fold decrease in wild type telomerase activity was observed in 35S::$TER2_{RSA}$ plants (Fig. 26A). In addition, although no telomere length change was observed in T1 plants (data not shown), telomeres were markedly shorter and less heterogeneous in T2 and T3 35S::$TER2_{RSA}$ transformants (Fig. 26C), a finding consistent with reduced telomerase activity *in vivo* (143). Altogether, our data argue that TER2 functions as a novel negative regulator of telomerase activity *in vivo*.

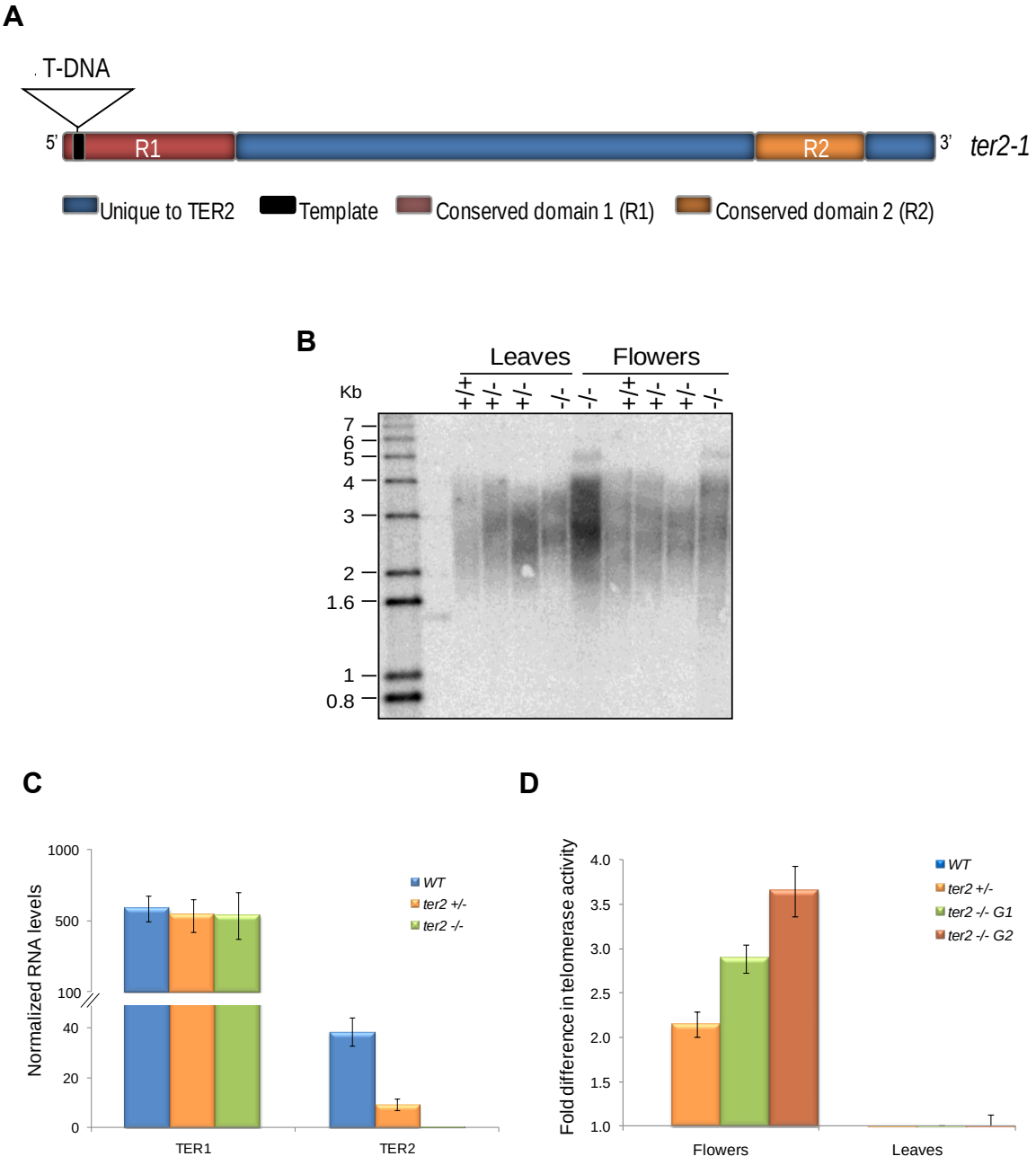


Figure 25. TER2 functions as a negative regulator of telomerase. (A) Position of the T-DNA insertion in *ter2-1*. (B) TRF analysis of *ter2-1* mutants. (C) Q-RT-PCR results for TER1 and TER2 levels in *ter2-1* mutants. (D) Q-TRAP results for WT, *ter2-1+/-* and *ter2-1-/-* plants. Values were normalized to telomerase activity in WT plants.

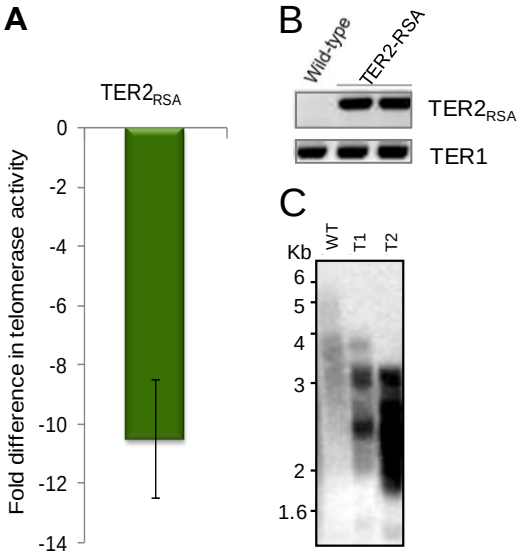


Figure 26. TER2 over-expression results in inhibition of TER1 RNP. (A) Q-TRAP results for 35S::$TER2_{RSA}$ transformants. TRAP was performed using a reverse primer complementary to the wild type repeat. (B) RT-PCR for TER1 and TER2 in WT and TER2 over-expression (35S::$TER2_{RSA}$) plants. (C) TRF analysis of second (T2) and third (T3) generation $TER2_{RSA}$ mutants.

TER2 associates with a set of proteins distinct from TER1 RNP components

We previously reported that TER1 assembles with TERT, POT1a and dyskerin into an active telomerase RNP (317) and Chapter II). Moreover, POT1a does not bind TER2 *in vitro*, implying that TER2 may assemble with a distinct set of proteins *in vivo*. To investigate protein interactions with TER2, co-immunoprecipitation (IP) assays were performed. T7-tagged POT1b was co-expressed with TER1 or TER2 in RRL, followed by IP with anti-T7 agarose beads. The bound RNA was extracted, cDNA was generated using random pentadecamers and PCR was performed with primers specific for TER1 or TER2. We found that POT1b binds TER2, but not TER1 (Fig. 27A). Gel shift analysis using *in vitro* transcribed RNA and recombinant proteins confirmed the interaction between TER2 and POT1b and demonstrated its specificity (Fig. 27B). As part of the analysis of TER2, we asked if it could bind the other *Arabidopsis* POT1 paralog, POT1c. IP reactions showed that POT1c bound both TER1 and TER2, although a slightly stronger interaction was observed with TER2 (Fig 27A and data not shown). Thus, all three POT1 proteins associate with TER, but each protein has a different preference for the TER isoforms.

We detected a weak interaction between KU and AtTER1 (317) and Chapter II). To test whether KU interacts with TER2, IP experiments were conducted as described above. We found a robust interaction with KU70 and TER2 (Fig. 27C). No difference was observed when KU80 was added to the reaction. Gel shift assays revealed RNP complexes containing TER2-KU70 and

TER2-KU70/KU80, but no interaction between TER1 and KU (Fig. 27D). We conclude that TER2 can interact with POT1b, POT1c and KU *in vitro*.

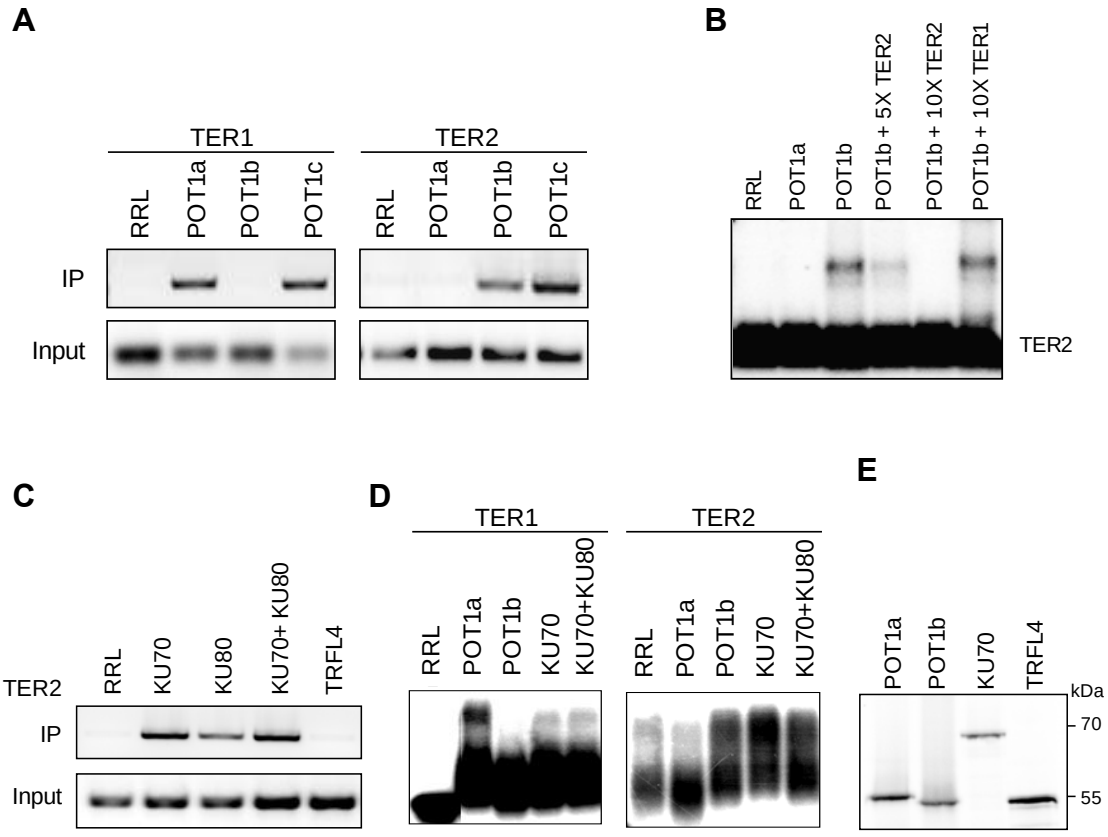


Figure 27. TER2 associates with POT1b, POT1c and KU *in vitro*. (A and C) T7-tagged proteins were co-expressed with TER1 and TER2 in RRL. RT-PCR was carried out after immunoprecipitation (IP). TRFL4, a ds telomeric DNA binding protein, was used as control. (B and D) Results from electrophoretic mobility shift assays. Labeled RNAs are indicated. (B) Competition experiments with TER1 and TER2. (E) ^{35}S-labeled protein expression.

POT1b is a negative regulator of telomerase activity

Given its interaction with TER2, we hypothesized that POT1b might function in telomere maintenance. To test this possibility, we examined the consequences of a null mutation in *POT1b*. The *pot1b-1* line contains a T-DNA insertion in the second exon of the POT1b gene (Fig. 28A). RT-PCR was performed, using primers flanking the insertion sites, to determine if this T-DNA insertion disrupts AtPOT1b gene expression, (Fig. 28A and B). No RT-PCR products were detected using primers either downstream of or across the insertion (Figure 28B, lanes 4 and 6), confirming that expression of the full-length AtPOT1b mRNA was abolished in the mutants. Although AtPOT1b transcripts from the region upstream of the T-DNA can be amplified in both wild-type and the mutant (Fig. 28A, lanes 1 and 2), it encodes only a small (53 amino acid) polypeptide which is likely to be non-functional. As with *ter2* null mutants, no obvious morphological phenotypes were associated with *pot1b-1* (data not shown). TRF analysis revealed no difference in bulk telomere length for homozygous mutants versus wild type siblings (Fig. 28C). Strikingly, however, Q-TRAP showed a 3-fold increase in level of telomerase activity *pot1b-1* flowers relative to wild type. This increase in telomerase activity parallels the increase observed in *ter2-1* mutants (Fig. 28D) and implies that POT1b, like TER2, functions as a negative regulator of telomerase activity.

To gain further insight into the function of POT1b, we measured TER1, TER2 and $TER2_S$ transcript levels in *pot1b-1* null plants. As expected, no

change in TER1 was observed in *pot1b-1* flowers compared to wild-type. However, TER2 and $TER2_S$ transcripts were undetectable in *pot1b-1* mutants (Fig. 28E), suggesting that POT1b is required for stability and or accumulation of TER2. Notably, TER1, TER2 and $TER2_S$ levels are unperturbed in plants null for POT1a (Fig. 28E). We conclude that POT1a and POT1b not only interact with different TER isoforms, but also they make distinct contributions to TER1 and TER2 metabolism.

To further investigate the TER2-POT1b interaction, filter binding assays were performed with TER2 deletion constructs to map the POT1b binding site. POT1b bound a construct lacking the TER2 3' unique region (TER2 Δ 3') (Fig. 29 A-C). POT1b binding was also detected with $TER2_s$ and with a construct lacking the R2 domain (TER2 R1/IV). In contrast, no binding to the TER2 intervening sequence (TER2 IV) was detected. TERT binding to TER2 followed a similar pattern of association and hence both the TERT and POT1b binding sites are contained within the 145 nt R1 domain of TER2. Further mapping studies are needed to assess whether POT1b and TERT could compete for the same site on TER2.

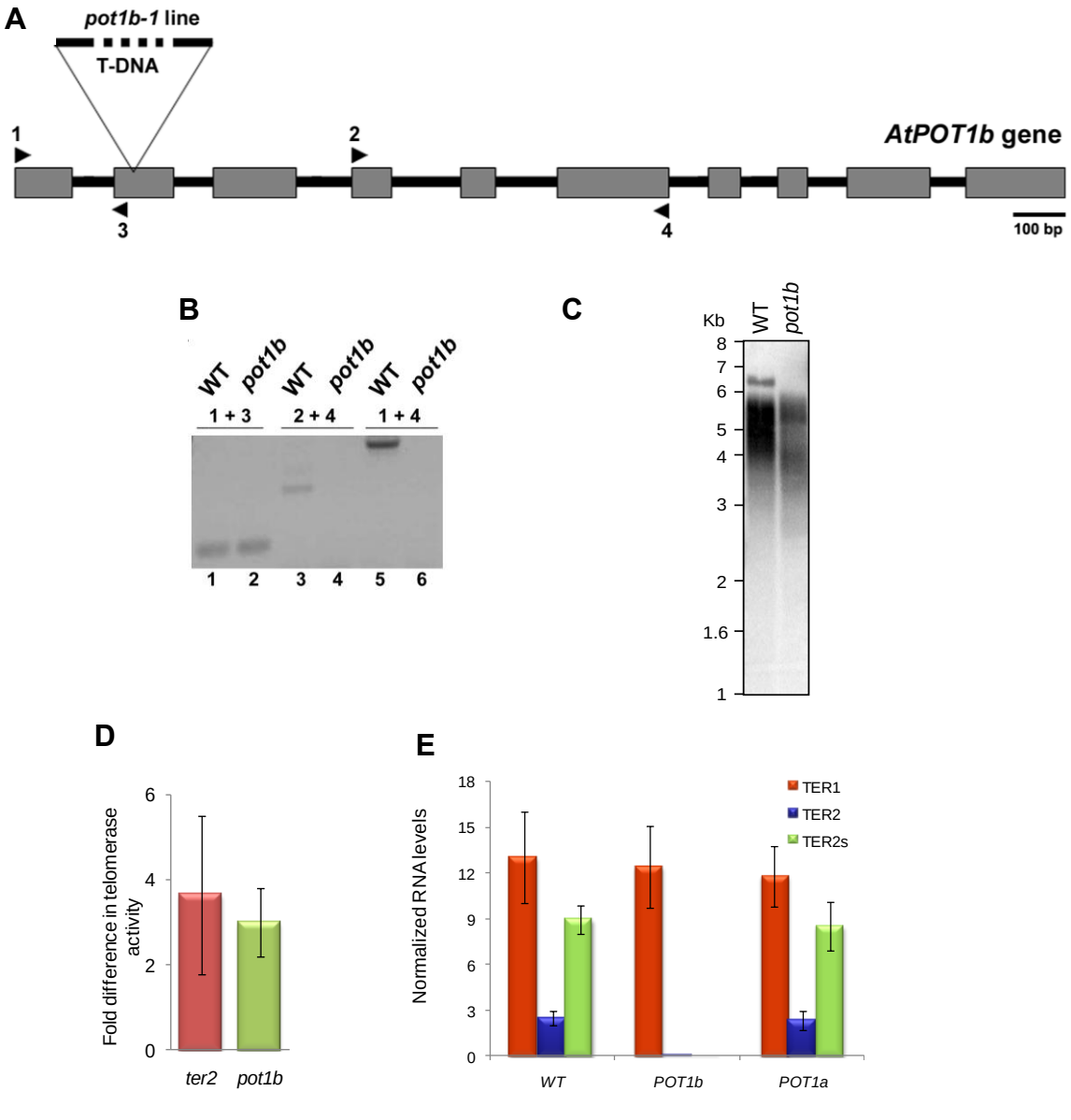


Figure 28. POT1b functions as a negative regulator of telomerase. (A) Position of the T-DNA insertion in *pot1b-1*. (B) RT-PCR results for POT1b mRNA in wild type and *pot1b* mutants. 1 and 4 are primer pairs are indicated in (A). (C) TRF analysis of WT and *pot1-b* mutants. (D) Q-TRAP results for WT and *pot1b* plants. Values were normalized to telomerase activity in WT plants. (E) Q-RT-PCR results for TER1 TER2 and $TER2_S$ levels in WT, *pot1b* and *pot1a* plants. Panels A, B and C were generated by Dr. Eugene Shakirov.

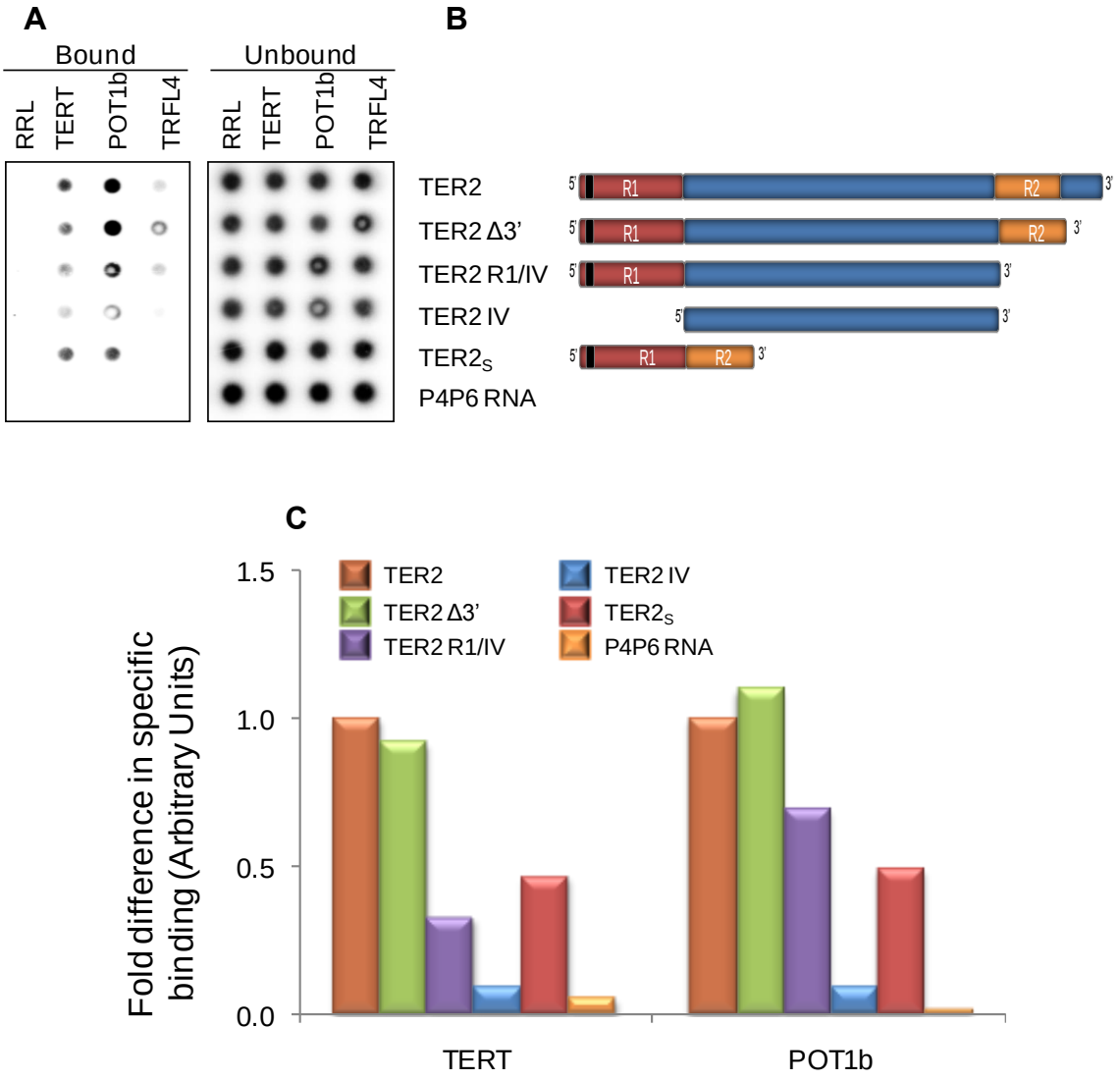


Figure 29. POT1b associates with TER2$_S$. TER2 deletion constructs were tested for association with POT1b and TERT by filter binding. (A) Representative filter binging assay. Left, bound. Right, unbound. TRFL4, a double-strand telomeric DNA binding protein, was used as control. (B) Graphic depiction of the TER2 deletion transcripts used for binding experiments. (C) Quantification of data from (B). The specific binding was calculated.

TER2 and TER2$_S$ assemble into RNPs distinct from TER1 in vivo

We investigated whether the three TER isoforms assemble into different RNP complexes *in vivo* using glycerol gradient fractionation (Fig. 30A). Telomerase activity was monitored by TRAP. TER1 and TER2 were amplified by RT-PCR and western blot analysis was employed to detect TERT and KU (Fig. 30A). Free TER1 and TER2 were found in the least dense fractions of the gradient (Fig. 30A, lanes 1-2). Fractions with telomerase activity contained TER1, TER2 and TERT but not KU70 (Fig. 30A, lanes 8-9). TERT and TER1 or TER2 were also found in inactive fractions (Fig. 30A). Interestingly, KU70 was only detected in non-active fractions containing TERT and TER2. This result supports our *in vitro* binding data and indicates that KU70 may assemble into an inactive TER2 RNP *in vivo* (Fig. 30A, lanes 4-5). Altogether, these results imply that *Arabidopsis* assembles several telomerase RNPs, which differ in TER and protein composition and in enzymatic activity.

IP with *Arabidopsis* cell culture extracts was performed to further evaluate telomerase RNP interactions *in vivo*. Following IP, end-point and quantitative RT-PCR were used to monitor the three TER isoforms and the data were normalized to account for differences in IP efficiency and primer usage (Fig. 30B). As expected, none of the TER isoforms were detected in Histone H3 or pre-immune IP control reactions (Fig. 30B). Like TER1 (ref and Chapter II), TER2 was enriched in the TERT precipitate. Strikingly, only two-fold more TER1 was precipitated relative to TER2 (Fig. 30B). Since TER1 is approximately 19-

fold more abundant than TER2 in cell culture (Chapter II), we conclude that TERT preferentially associates with TER2 *in vivo*. $TER2_s$ was also associated with TERT *in vivo*, but it was less enriched than TER2 by approximately 8-fold. Since TER2 is 4-fold more abundant than $TER2_S$ in cell culture, these findings are consistent with our *in vitro* binding data showing that TERT has a lower affinity for $TER2_s$ than TER2. TER2, like TER1 was enriched in the dyskerin IP (Fig. 30B; Chapter II), in agreement with the presence of putative H/ACA boxes at its 3' terminus (data not shown). In contrast, only a trace amount of $TER2_s$ was seen in this IP, suggesting that the processed form of TER2 may lack the appropriate recognition site for dyskerin.

In addition, KU70 antibody pulled down TER2, but only trace amounts of TER1. Notably, the amount of TER2 precipitated was significantly less than with anti-TERT antibody (Fig. 30B), suggesting that KU associates with only a subset of TER2 RNPs. Finally, TER2 was strongly enriched over TER1 in the anti-POT1b pull-down (Fig. 30B), while a only small fraction of TER2 was detected in the anti-POT1a IP (Fig. 30B).

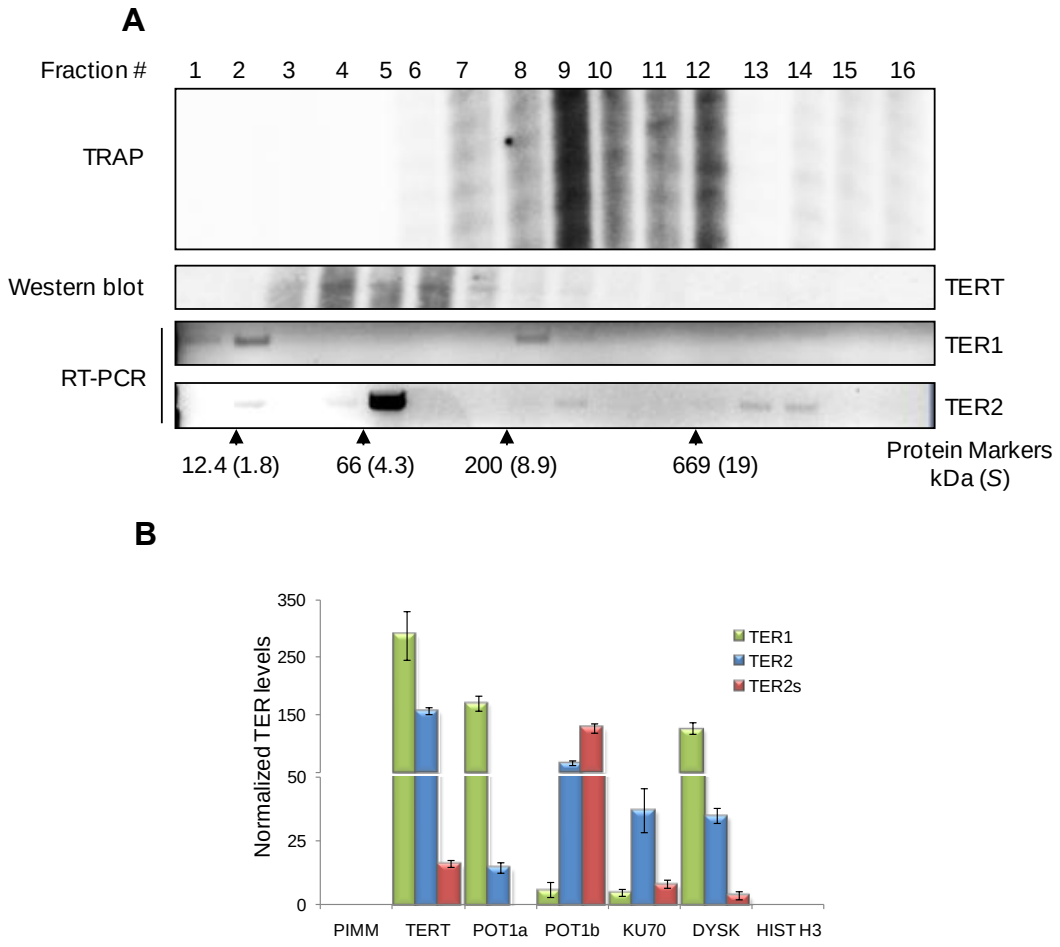


Figure 30. Distinct TER1, TER2 and TER2$_S$ RNPs *in vivo*. (A) Glycerol gradient fractionation of *Arabidopsis* cell culture nuclear extracts. Top, telomerase activity assay. Middle, TERT and KU were visualized by western blot. Bottom, TER1 and TER2 were detected by RT-PCR. The position of size markers is indicated below the bottom panel. the TRAP panel. (B) IP was carried out with *Arabidopsis* cell culture extracts using the indicated antibodies. Pre-immune serum, anti-histone H3 and anti-enolase antibodies were used as controls. RNA levels were normalized to the primer efficiency, the levels of U6 snRNA, the pre-immune control and the antibody efficiency.

TER2s displayed a different pattern of protein associations than TER2. Although a residual amount of $TER2_S$ was detected in the KU70 IP, 2-fold more $TER2_S$ than TER2 was detected in the anti-POT1b IP. This finding is noteworthy since 7-fold more $TER2_S$ was enriched in the POT1b IP than in the TERT IP. Thus, $TER2_S$ appears to be primarily associated with POT1b in an RNP lacking TERT, dyskerin and KU. We conclude that *Arabidopsis* harbors multiple RNP complexes with distinct RNA and protein composition and different roles in telomere maintenance.

Discussion

All organisms studied thus far encode one TER subunit (14, 170-172, 319-321). *Arabidopsis* is unusual as it encodes multiple telomerase associated RNAs. We previously showed that TER1 is the major template for telomere maintenance *in vivo*, leaving unresolved the role of TER2 in telomere biology. Here we demonstrate that TER2 assembles into an RNP that is distinct in protein composition from the TER1 RNP, and further that the TER2 RNP negatively regulates telomerase *in vivo*.

Three TER isoforms in Arabidopsis

As part of our analysis of TER2, we unexpectedly discovered that it is processed to generate another TER isoform, TER2s. $TER2_S$ is not encoded in the *Arabidopsis* genome and sequence analysis indicates that $TER2_S$ sequence

is 100% identical to the TER2 conserved regions (R1 and R2), suggesting that $TER2_S$ is the product of TER2 processing. In addition, $TER2_S$ is detected by northern blotting of total RNA from wild type flowers and tissue culture (Chapter V) which indicates that $TER2_S$ is a bona fide transcript and not an RT-PCR artifact. Consistent with this conclusion, computer prediction analysis did not predict any high score branch points, nor donor or acceptor splice sites (data not shown).

Although the splicing machinery is conserved throughout eukaryotes, plant consensus splice sites are detected in a low fraction of mRNAs (322-323), which suggests plant splicing sites are highly variable. Thus, TER2 may be processed by a non-canonical splicing mechanism. We think it is likely that *Arabidopsis* TER1 and TER2 are synthesized by RNA polymerase II like their counterpart in vertebrates, but we do not know whether the two RNAs contain a 5' cap. Furthermore, we could not detect a poly(A) tail on TER1 or TER2, but it is possible that a subset of TER2 molecules contain a poly(A) tail and these are processed to generate TER2.

Processing of *Saccharomyces pombe* TER1 3' end of TER1 is required for telomerase function (324). Although, SpTER1 transcripts bearing a poly(A) tail are detected in fission yeast cells (171, 324), SpTER1 transcripts associated with telomerase do not contain a poly(A) tail. The mechanism of processing is via slicing (324), a process in which only the first transesterification reaction is carried out by the spliceosome. Failure to process SpTER1 results in telomerase

inhibition. Thus, RNA processing may play a key role in telomerase biogenesis and function.

The production of $TER2_s$ is developmentally regulated in different plant cells and inversely correlated with the abundance of TER2. In organs lacking telomerase (leaves and stem) only a small amount of TER2 is observed, and TER2s is undetectable. In the telomerase rich setting of flowers and cell culture, both TER2 and TER2s levels rise. Curiously, TER2s becomes the predominate isoform in flowers, but in cell culture the profile flips and TER2 is more abundant than TER2s. While the level of TER2/TER2s does not directly correlate with enzyme activity, the ratio of TER1:TER2 is similar in both settings. In cell culture, TER1 is approximately 19 fold more abundant than TER2 and in flowers. Thus, it is possible that TER2 splicing decreases the abundance of this negative regulator for TER1 and as a consequence results in an increase in telomerase activity (see below).

TER1 and TER2 interact with an overlapping, but distinct set of proteins

Our data suggest that TER1 and TER2 RNPs consist of the same core components, TERT and dyskerin, but then diverge with respect to their interactions with KU and the POT1 family of proteins. Ku70/80 exhibits a strong preference for TER2 over TER1. Intriguingly, both IP and glycerol gradient data indicate that Ku70 co-purifies with TER2 in fractions lacking telomerase activity. Hence not only is the KU-TER interaction dispensable for telomerase function in

Arabidopsis, the interaction of KU with TER2 may reduce the intrinsic catalytic function of the TER2 RNP and diminish with its ability to direct telomere repeat incorporation *in vivo* (see below).

The POT1 paralogs also exhibit dramatically different preferences for the TER isoforms. POT1b interacts with TER2, but not with TER1 (Fig. 27 and 30), while POT1a displays the opposite specificity (Chapter II), interacting with TER1 but not TER2 (313) and Chapter II]. POT1a binds to a CU-rich sequence within the highly divergent 5' region of TER1 that is absent from TER2 (see Chapter V). In contrast, POT1b recognizes the R1 domain in TER2 and $TER2_s$, which contains the telomere template and displays 84% identity with TER1 (data not shown). This finding implies that subtle differences in TER1 and TER2 sequence/structure will dictate the specificity of POT1b for TER2. Interestingly, POT1c, a product of a POT1a gene duplication composed by a single OB-fold, binds both TER1 and TER2 at a similar level. These results suggest that specificity of the POT1a-TER1 and POT1b-TER2 interaction may reside within the second OB-fold domain.

Duplication of TER and POT1 may fuel the evolution telomerase RNP complexes with distinct functions

Although POT1 family genes have undergone duplication events in a wide variety of organisms (118, 325-326). Among mammals, there is a single POT1 gene in humans, while mice have two (118). Similarly, in *Arabidopsis*,

there are three POT1 paralogs but in papaya, a member of the same family, only a single POT1 gene is evident (132). In a few cases the POT1 orthologs have similar functions, (e.g mouse POT1a and POT1b), but frequently POT1 duplication has resulted in different function (327). For example, *Tetrahymena* POT1a is expressed in vegetative cells and functions in telomere length regulation and protecting the telomere end (326). POT1b, on the other hand, is only expressed in mated cells where it is involved in de novo telomere formation (327).

The functions and interactions of *Arabidopsis* POT1 have diverged radically from their counterparts in other organisms. Null mutations in *pot1a* and *pot1b* mutants do not lead to chromosome capping defects (129-130, 317) and this study). In addition, both POT1a and POT1b have evolved to bind TER instead of ssDNA. Furthermore, the function of POT1b has diverged from the POT1a. POT1b not only binds TER2 instead of TER1, but this interaction appears to be required for TER2 stability. In contrast, the POT1a association with TER1 is not critical for TER1 stability. How POT1a promotes telomere maintenance by TER1 remains unknown.

We hypothesize that duplication of POT1 in *Arabidopsis* fueled the evolution of the nucleic acid pocket in POT1 allowing it to recognize TER instead of telomeric DNA. This in turn stimulated the migration of POT1 from the chromosome terminus to the telomerase RNP. The duplication of TER appears to be a much more recent event in plant evolution since TER2 has not been

identified outside *A. thaliana*. With a second copy of TER and additional TER binding proteins, the groundwork was laid for the emergence of a novel regulatory mechanism for telomerase.

The TER2 RNP functions as a negative regulator of telomerase

Our data indicate that TER2 is a negative regulator for telomerase activity. Although at least a subset of TER2 molecules are assembled into enzymatically active RNP complexes *in vivo*, TER2 does not play a significant role in telomere maintenance. A low level of TER2 directed telomere repeat incorporation was detected when the $TER2_{RSA}$ construct was over-expressed. In addition, we observed a lower frequency of $TER2_{RSA}$ directed incorporation of mutant repeats, compared to $TER1_{CC}$, despite the presence of wild type TER1 in these plants. Therefore, we suspect that low level of repeat incorporation by TER2 reflects the gross over-expression of this RNA and the inhibition of TER1, the preferred telomerase template.

Precisely how TER2 inhibits telomerase is unknown, but the current data suggest that telomerase regulation is facilitated by TER2 interacting proteins. TERT has a ~9-fold higher affinity for TER2 than for TER1 *in vitro*, and although additional proteins likely impact telomerase biogenesis *in vivo* (147, 162, 308, 328), TER2 is over-represented in TERT-containing RNP complexes. Thus, TER2 may inhibit telomerase by outcompeting TER1 for the telomerase catalytic subunit.

The telomerase accessory proteins POT1b and KU appear to reinforce TER2-mediated negative regulation. Preliminary studies show that *pot1b-1*l mutants exhibit a three-fold increase in telomerase activity (A. Nelson and D. Shippen, unpublished data), similar to the increase observed in *ter2-1* mutants. In addition, the TER2-interacting factor KU is a strong negative regulator of telomerase-mediated telomere elongation in *Arabidopsis* (296). Since telomerase enzyme activity levels are not increased in *ku70-/-* mutants (251), Nelson A. and D. Shippen unpublished data), KU must regulate telomere elongation rather than enzyme activity. One intriguing possibility is that KU is involved in recruitment of telomerase to the telomere as in budding yeast (182, 252). However, because KU is associated with the TER2 RNP, this would result in recruitment of the negative regulatory complex, decreasing telomere extension. The discovery of a novel negative regulatory mechanism for plant telomerase argues that additional modes of restraining telomerase remain to be elucidated in mammals where misregulation of this enzyme can lead to carcinogenesis.

Materials and Methods

Primer extension

Primer extension was carried out on total RNA extracted from *Arabidopsis* cell culture. 0.25pmol of 5' end labeled oligonucleotide was incubated with total RNA at 95°C for 5 min and allowed to anneal in two sequential 15 min

incubations at 72°C, and 60°C, after which extension mix (50mM Tris-HCl pH 8.3, 15mM KCl, 3mM MgCl2, 5% DMSO, 1mM DTT, 1mM dNTPs, 1.5U RNAseOUT® and 200U SuperScript III® reverse transcriptase (Invitrogen) was added to a 30µl final volume. The reactions were incubated at 58°C for 3 h. The enzyme was inactivated at 80°C for 10 min. The RNA was hydrolyzed by incubation at 70°C with 15µl of 1N NaOH for 10 min. The reaction was neutralized and precipitated with 15µl of 1N HCl, 20µl 3M NaOAc pH 5.2, ethanol and glycogen. The products were resolved by denaturing PAGE.

In vitro telomerase reconstitution

A TERT-pET28a plasmid with an N-terminal T7 tag was used for telomerase reconstitution experiments. Reactions were assembled with 100ng of TERT-pET28a plasmid and 0.5pmol or 0.1pmol of gel purified DNA template encoding TER1 or TER2 respectively, driven by a T7 promoter, in a mix containing Rabbit Reticulocyte Lysate (RRL) (Promega), amino acids, RNase inhibitors, and T7 RNA polymerase. Reactions were incubated for 90 min at 30°C. T7 agarose beads (Novagen) were blocked with buffer W-100 (20mM TrisOAc [pH 7.5], 10% glycerol, 1mM EDTA, 5mM MgCl2, 0.2M NaCl, 1% NP-40, 0.5mM sodium deoxycholate, and 100mM potassium glutamate) containing 0.5mg/ml BSA, 0.5mg/ml lysozyme, 0.05mg/ml glycogen, 1mM DTT and 1µg/ml yeast tRNA. The reconstitution reaction was mixed with the beads to a 200µl final volume and incubated for 2 h at 4°C with rotation. Beads were washed 6X

with 800 µl of W-400 buffer (W-100 containing 400 mM potassium glutamate) and 3X with 800µl of TMG buffer (10mM TrisOAc pH 7.5, 1mM $MgCl_2$, and 10% glycerol). After the final wash, beads were resuspended in 30µl of TMG. 2µl of beads were used for TRAP assays as previously described (130, 242).

In vitro *binding assays*

For co-IP experiments, POT1a, POT1b, KU70 and KU80 cloned in pET28a with a T7 tag were co-expressed with TERs in RRL as described above. After IP RNA was extracted and RT-PCR was performed (see Supplementary Methods).

Electrophoretic mobility shift assays were performed with RNA transcribed *in vitro* with T7 RNA polymerase and [α-32P]-CTP labeled TER1 and TER2. Binding reactions contained 3µl of RRL expressed protein, 0.1pmol of 32P labeled TER and 1X binding buffer (25mM Tris-HCl pH 8.0, 10mM $Mg(OAc)_2$, 25mM KCl, 10mM DTT and 5% Glycerol) in a 30µl final volume. 1µM yeast tRNA and 0.5µM RNA $(U_3AG_3)_4$ were used as nonspecific competitors. After 20 min at 30°C the reaction was loaded onto a 0.8% agarose 0.5X TBE gel and run for 2 h at 70 volts at 4°C. Gels were dried and exposed to phosphorimager screens.

For double-filter binding assays, TERT expressed in RRL was incubated with decreasing concentrations of pre-folded RNA transcribed *in vitro* using 5-end 32P-labeled RNA as a tracer. Binding reactions contained 0.5µl of

recombinant protein, pre-folded TER in binding buffer (50mM Tris-HCl pH 7.5, 200mM potassium glutamate, 0.5mg/ml BSA, 0.5mg/ml tRNA, 1mM MgCl2, 1mM DTT and 0.01% NP-40) in a 25µl final volume. After 30 min at 30°C, the reactions were filtered through nitrocellulose and nylon filters using a dot-blot apparatus (BioRad). The membranes were washed with 600µl washing buffer (50mM Tris-HCl pH 7.5, 200mM potassium glutamate, 1mM $MgCl_2$, 1mM DTT and 10% glycerol), dried, exposed to a phosphor storage screen and scanned after 2h. Equilibrium dissociation constants, Kd, were obtained by non-linear regression of the binding data fitted to a one-site binding model using Graphpad Prism software.

Plant materials and genotyping

Arabidopsis seeds with a T-DNA insertion in TER2 (SAIL_556_A04) were obtained from the *Arabidopsis* Biological Resource Center (Ohio State University, Columbus, OH). Seeds were cold-treated overnight at 4°C, and then placed in an environmental growth chamber and grown under a 16-h light/8-h dark photoperiod at 23°C. Plants were transformed using the in planta method [88]. For genotyping, DNA was extracted from flowers and PCR was performed with the following sets of primers: #38: 5'-GACGACAACTAA ACCCTACGCTTACA-3' and #45: 5’-CGATGTTGTTTTTCTGCTTAGGACACA-3’. To amplify mutant TER2 alleles containing a T-DNA insertion, the T-DNA

specific primer was used along with TER 8526-01 fwd: 5'-GAGACGCAGCGAGCGATAGCCGATAG-3' primer.

Template mutation and plant transformation

To generate template mutations in the template region of TER2, a PCR product containing the desired template mutation was generated with the primers TER2RSA fwd (5'-CACCGACGACAACTAGTACCTACG CTTACA-3') and TER2 end reverse (5'-AATTCTGTGTAGCTATGATCTTGTGGCA-3'). The mutation was confirmed by sequencing. $TER2_{RSA}$ was cloned into the destination vector pB7WG2 and transformed into plants homozygous for the T-DNA insertion in *ter2-1.* After transformation the seeds were selected in MS agar containing kanamycin at 50µg/ml.

End-point RT-PCR and quantitative RT-PCR

Total RNA was extracted from 0.5 g floral or other tissue using Tri Reagent (Sigma). cDNAs were synthesized from total RNA using Superscript III reverse transcriptase (Invitrogen). Random pentadecamers were incubated with 2 µg of total RNA in the supplied buffer at 65°C for 5 min. Reverse transcription (RT) was carried out with 100U of Superscript III at the following temperatures 37°C for 20 min, 42°C for 20 min and 55°C for 20 min. Enzyme was inactivated at 80°C for 10 min and RNA was degraded with RNase H (New England Biolabs). 1.5 µl of cDNA was used in PCR.

For real-time RT-PCR, 2 µl of the above cDNA was used at a 1:10 dilution in a 20µl reaction containing 10 µl of SyBr green master mix (NEB) and 2 µl of each primer (2µM). PCR was performed for 40 cycles with 30 s at 95°C and 60 s at 60°C. Threshold cycle values (Ct) were calculated using an iCycler iQ thermal cycler (BIO-RAD) and the supplied Optical System Software.

Quantitative RT-PCR data analysis

Amplification efficiencies were calculated for each primer pair in a 5 point titration curve. The slope was calculated from a standard curve where Ct was on the y-axis and log(cDNA dilution factor) on the x-axis. The corresponding real-time PCR efficiency (E) was calculated according to the equation: E=(10-1/slope)-1. To correct for intra-assay and inter-assay variability, each sample was evaluated by tripicate within one run in at least three different experimental runs.

The relative expression level (R) was calculated as follows: R = (Etarget)ΔCttarget(control-sample) / (Ereference)ΔCtreference(control-sample) as previously described (329). U6 snRNA and β-actin were used as reference. Normalization to the pre-immune control and to the efficiency for each antibody was used for RNA quantification in the pull-down samples. Primers used for real-time PCR are as follows: TER1 Q4F:5'-CCCATTTCGTGCCTATCAGACGAC-3'. TER1 Q4R: 5'-TCTCCGACGACCATTCTCTCGATAC-3'; TER2#38: 5'-GACGA CAACTAAACCCTACGCTTACA-3' and TER2#40: 5'-CAGGATCAATCGGAG

AGTTCAATCTC-3'; TER2S: TER2#38 and TER2S# 193: 5'-CCCCATCTCCGA CGAGACGAC-3'; TERT Q3F: 5'-ACCGTTGCTTCGTTGTACTTCACG-3' and TERT Q3R: 5'-CGACCCGCTTGAGAAGAAACTCC-3'; U6-1F: 5'-GTCCCTTCG GGGACATCCGA-3' and U6-1R: 5'-AAAATTTGGACCATTTCTCG A-3' β-Actin 2F: 5'-TCCCTCAGCACATTCCAGCAGAT-3' and β-Actin 2R: 5'-AACGATTCCT GGACCTGCCTCATC-3'.

TRF analysis, TRAP and Q-TRAP assays

TRF and TRAP assays were performed as previously described [29,64]. TRAP products from mutant TER2RSA telomerase, a specific mutant reverse primer 5'-CCTAGTACCTAGTACCTAGTACCTA-3' was used. Q-TRAP was performed as previously described (242).

Antibodies, immunoprecipitation and western blotting

AtKU70 antibodies were kindly provided by Dr. Karel Riha (Gregor Mendel Institute, Vienna). AtTERT and AtPOT1a antibodies have been previously described (130, 242). The anti-dyskerin polyclonal antibody was raised in rabbits against recombinant full-length AtNAP57 expressed in E. coli. The POT1b is an affinity-purified peptide antibody (Covance). IP efficiency was calculated for each antibody using ^{35}S labeled protein expressed in rabbit reticulocyte lysate. Western blotting was performed with a 1:2000 dilution of anti-KU70, anti-POT1a, anti-POT1b, anti-TERT and anti-Histone H3 antibodies

(Upstate). The anti-dyskerin antibody was used at a 1:5000 dilution. Peroxidase-conjugated light chain-specific mouse anti-rabbit secondary antibody (Jackson Immunoresearch) was used at a 1:20,000 dilution. Following IP, RNA was phenol:chloroform extracted from the beads and subjected to RT using superscript III® reverse transcriptase (Invitrogen) and random pentadecamers.

Glycerol gradient fractionation

In vitro reconstitution reactions were scaled-up to 400µl final volume and fractionated on a 10 mL 10-50% glycerol gradient in 20mM HEPES, pH 8.4, 100mM NaCl, 10mM MgCl2, 1mM EDTA, 1mM DTT, 0.1% NP40 for 20h at 35000 rpm in a SW-40 rotor (Beckman) at 4C. Fractions of 500ul were collected manually from the top of the gradient. The following size standards were run under the same conditions: Cytochrome c (12.4 kDa, 1.8S), bovine serum albumin (66 kDa, 4.3S), β-amylase (200kDa, 8.9S) and Thyroglobulin (669 kDa, 19S).

CHAPTER IV

CHARACTERIZATION OF THE POT1a-TER1 INTERACTION

Summary

In most organisms the proteins that bind telomeric single-stranded DNA do so via OB-fold domains. In *Arabidopsis*, the OB-fold containing protein POT1a has evolved to bind TER1 and to work as a positive regulator of telomerase. Here we report the structural and sequence requirements of the POT1a-TER1 association. Using a combination of deletional mutagenesis, nucleotide substitutions and RNA structure predictions we identify a minimal POT1a binding region of 38 nt within the hypervariable 5'-end of TER1. This region is characterized by a 6nt U-C rich internal loop flanked by two base-paired stems. Notably, structures with similar characteristics were found in putative TER subunits from *Arabidopsis* close relatives. In particular, the putative TER1 of *Brassica* is highly similar in sequence composition and architecture to the AtTER1-POT1a interacting region and shows co-variation in the base paired stem required for POT1a recognition. We show that the two AtPO1 OB-folds act as nucleic acid recognition domains. TER1 binding is supported by the first OB-fold, although a more robust interaction is observed with both OB-folds. Our data suggest RNA binding activity is not unique to *Arabidopsis* POT1, but rather is a widespread phenomenon in plants.

Introduction

In most eukaryotes, the terminal portion of telomeric DNA is characterized by a single-stranded extension of the 3'-end of the G-rich strand. This G-rich extension or G-overhang is required for telomere maintenance by telomerase, as this enzyme does not extend double-stranded blunt ends (330). The reverse transcriptase subunit of telomerase (TERT) employs a telomeric sequence within the telomerase-associated RNA (TER) as a template for the addition of G-rich telomeric DNA repeats (14-15). The 3' terminus of the G-overhang is aligned on TER and extended one nucleotide at a time until the end of the template has been copied into DNA. The DNA 3' end is subsequently repositioned back to the beginning of the TER templating domain for another round of synthesis.

TER molecules are highly variable in nucleotide sequence and size: 159 nt in *Tetrahymena* (14), 380-560 nt in vertebrates (231, 320) and ~930-1540 nt in budding and fission yeast (50, 171, 331). Phylogenetic and mutational analysis has uncovered several functionally conserved secondary structure domains (Fig. 7, Chapter I). One element is the boundary domain, which is located immediately 5' of the template and defines the margins for reverse transcription. In addition, TER subunits are proposed to form a pseudoknot close to the template as well as a trans-activation domain. Vertebrate TER subunits contain an additional CR7 domain and a box H/ACA motif, while yeast TER transcripts contain sites for Est1 and Sm binding. Remarkably, telomerase

activity can be reconstituted *in vitro* with human TERT and two sub-domains within TER, the pseudoknot/ template region and the CR4/CR5 trans-activation domain (332). A truncated or “Mini T” version of yeast TER has also been reported (294).

TER is modular and is a highly flexible scaffold for telomerase-associated proteins. These accessory factors diverge among different organisms and participate in biogenesis, regulation and recruitment of telomerase to the telomere. For instance, *Tetrahymena* telomerase associates with factors that both positively and negatively regulate enzyme function (333). Human and budding yeast TER subunits are bound by the Ku70/80 heterodimer (182, 312) and in yeast Ku acts a positive regulator of telomerase (182, 252). Additionally, the yeast TER binds Est1, a key regulatory protein necessary for telomerase recruitment to the chromosome terminus (98, 100, 247).

In *Arabidopsis*, TER1 and TER2 associate with different sets of proteins that are proposed to influence the RNP function (see Chapters II and III). KU70, a powerful negative regulator of telomere length, associates preferentially with TER2. Interestingly, POT1 family proteins distribute to both TER subunits and are implicated in the positive and negative regulation of telomerase (see Chapters II and III).

POT1 (Protection of Telomeres) was first identified in fission yeast as an essential oligosaccharide/oligonucleotide binding fold (OB-fold) containing protein that binds the ss 3’-end G-overhang. A POT1 deficiency results in rapid

loss of telomeric and subtelomeric DNA (28, 334). POT1 homologs have been found in most eukaryotes (28, 118, 129, 132, 314, 335-337) and the vast majority of those characterized so far are implicated in chromosome end protection (Ref). Vertebrate POT1, functions by preventing the activation of the ATR-mediated DNA damage response (118, 336, 338). In addition, POT1 forms a heterodimer with TPP1, a homolog of the ciliate TEBP-β, which is proposed to regulate telomerase activity (121, 316) and increase telomerase processivity (120, 197).

While most mammals encode a single POT1, rodents express two POT1 paralogs that share 72% identity: mPOT1a and mPOT1b (118). Both POT1a and POT1b associate with TPP1 and their combined function results in protection of the chromosome end from DNA damage response (118, 316). Notably, the prevention of ATR activation relies on POT1a, whereas POT1b prevents C-strand resection (118, 339).

In several other organisms, including ciliates, worms, and *Arabidopsis* (40, 119, 129, 326), more than one POT1 paralog is present. These paralogs play different roles in telomere metabolism. *Tetrahymena* encodes two POT1 homologs. In *Tetrahymena*, TtPOT1a is a negative regulator of telomere length and prevents activation of a cell-cycle checkpoint (326). The function of TtPOT1b is currently unknown. *C. elegans* encodes at least four POT1 proteins bearing sequence homology to the OB-fold of vertebrate POT1 (40). Notably, these proteins show different ligand specificities. CeOB1 binds single-stranded

G-rich telomeric DNA, while CeOB2 binds single-stranded C-rich telomeric DNA. Depletion of either protein results in telomere length defects and increased recombination, although the extent of the phenotype is slightly different for each mutant (40).

Like POT1 proteins from vertebrates, *Arabidopsis* encodes more than one POT1 homologs; In this case, POT1a, POT1b and POT1c. POT1a and POT1b contain two N-terminal OB-folds and a C-terminal domain (129), Nelson A. and Shippen DE unpublished data). POT1c, though, only harbors a single OB-fold. POT1a and POT1b are highly divergent, displaying only 49% protein sequence similarity. As discussed in Chapters II and III, *Arabidopsis* POT1 proteins do not associate with single-stranded DNA, but instead have evolved to bind TER and to regulate telomerase activity.

POT1a is an accessory factor of the telomerase RNP (130, 313). Specifically, it associates with TER1 and functions as a positive regulator of telomerase (130, 313). *pot1a* null mutants display an ever-shorter-telomere phenotype, yet chromosome ends are fully protected in these mutants (130). POT1b, on the other hand, associates with TER2 and negatively regulates telomerase (Chapter III). Strikingly, POT1c associates with both TER1 and TER2. Finally, over-expression of dominant negative alleles of POT1b or POT1c result in genome instability, suggesting that Pot1b and POT1c contact a critical component of the telomere capping complex in addition to their interaction with telomerase (129).

The differences observed in nucleic acid binding specificity between *Arabidopsis* POT1 proteins and their homologs led us to dissect the fundamental aspects of the POT1a-TER1 interaction. Here we define a minimal POT1a binding site within the TER1 5'-end hypervariable region, upstream of the templating domain. This region is conserved in TER1 homologs from *Arabidopsis* close relatives. Additionally we demonstrate that POT1a, like POT1 proteins from other organisms, binds nucleic acid through its OB-folds. Finally, we present evidence for functional homology between *Arabidopsis* POT1a and yeast Est1. Altogether, our results suggest that AtPOT1a has evolved to play multiple roles in telomeres biology.

Results

Pot1a interacts with the 5' unique region of TER1

With the goal of obtaining a detailed structural model of the POT1a-TER1 complex, we sought to identify the POT1a-binding site of TER1 as the initial step in the characterization of their interaction. TER is so highly divergent at the nucleotide level that sequence alignments can be performed with confidence only among very closely related organisms (169-170, 172, 340-341). This suggests that TER molecules from different species share conserved secondary structural elements (169-170, 172, 340-341). We began by generating a preliminary structural model of TER1 from *A. thaliana*.

Our collaborator, Dr. Yehuda Tzfati (Hebrew University) used Mfold to generate a putative partial secondary structure for TER1 (Fig. 31). He defined common elements in all the Mfold the predicted secondary structures. Secondary structural elements common to TER in other organisms were observed in the predicted AtTER1 structure. A region forming a putative template boundary was found immediately 5' of the template. Secondary structure elements were also found that are consistent with the formation of a pseudoknot with three U-A:U base triples. To predict the rest of the secondary structure, Mfold was constrained to form the predicted template boundary region and pseudoknot, while keeping the template single-stranded. A small stem resembling p6.1, a key component of the CR4/CR5 domain in vertebrate TER (169), as well as a perfect Sm consensus sequence (AU_4G), was identified near the 3' end of the RNA. These observations are consistent with modular folding of for TER subdomains.

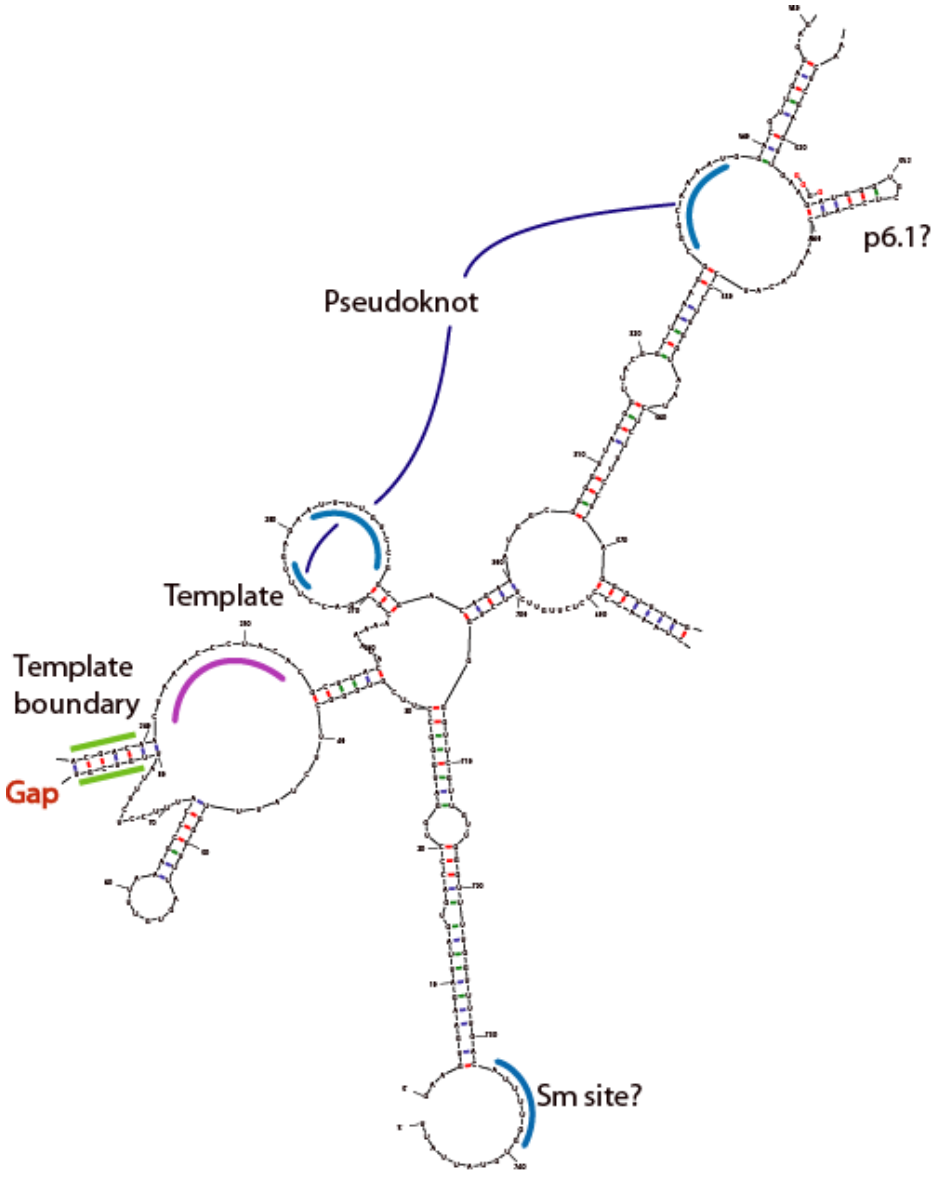


Figure 31. Preliminary secondary structure model of *A. thaliana* TER1. This structure includes the core conserved TER domains. The single-stranded template region (purple line) and the template boundary element (green lines) are indicated, as well as a putative Sm binding site and a predicted pseudoknot. (Model was generated by Dr. Yehuda Tzfati, Hebrew University).

We reasoned that the POT1a binding site must reside in a region of TER1 that is not found in TER2, since POT1a binds TER1 and not TER2. Also, since the highly conserved R1 region of both RNAs is sufficient to reconstitute telomerase activity *in vitro* (Chapter II), this region is bound by TERT and is likely to be inaccessible to other proteins. To test this hypothesis, we designed three constructs to test binding by electrophoretic mobility shift assay (EMSA). We did not observe association of POT1a with a construct containing the TER1 R1 and R2 conserved regions nor with the construct corresponding to the TER1 3'unique region (Fig. 32). This finding suggests that POT1a binds within the 5' non-conserved region upstream of the template.

The folding prediction for AtTER1 showed a gap upstream of the template domain that was not predicted to be part of the core functional elements of TER1 required for telomerase activity (Fig. 31). This 160 nt gap (nucleotides 87-234) is contained within the 5' region that is unique to TER1 and not present in TER2. We divided the TER1 5' unique region into three overlapping constructs of ~120nt and tested each one for binding to POT1a. Binding was observed with constructs 57-176 and 117-233, suggesting that the POT1a binding site may reside in the overlapping region. No association was detected with construct 1-117, which corresponds to the most 5' region of TER1 and includes the 5'-most 86nt that are predicted to form part of the TER1 core structure (Fig. 33).

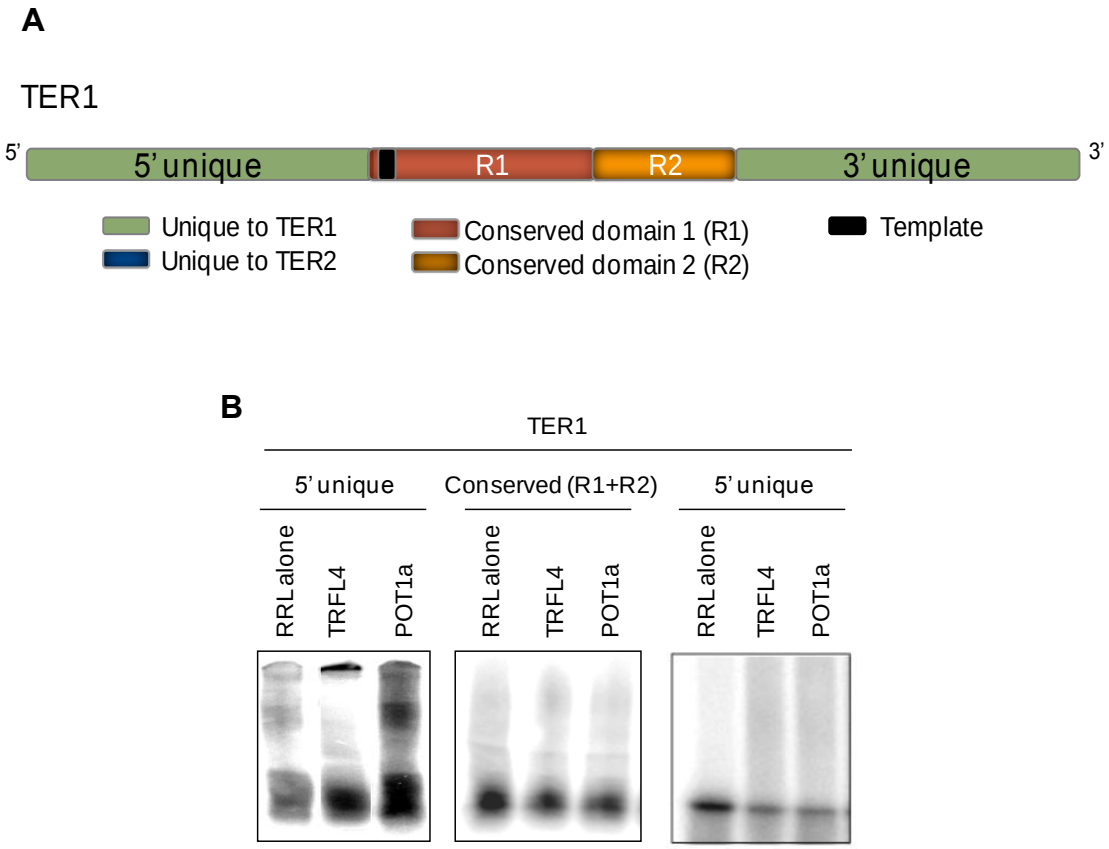


Figure 32. POT1a associates with the 5' unique region of TER1. (A) Diagram of TER1 indicating the template, 5'-unique, 3'-unique and conserved regions. (B) Results from electrophoretic mobility shift assays. Labeled RNAs are indicated. RRL alone and TRFL4, a ds telomeric DNA binding protein, were used as controls.

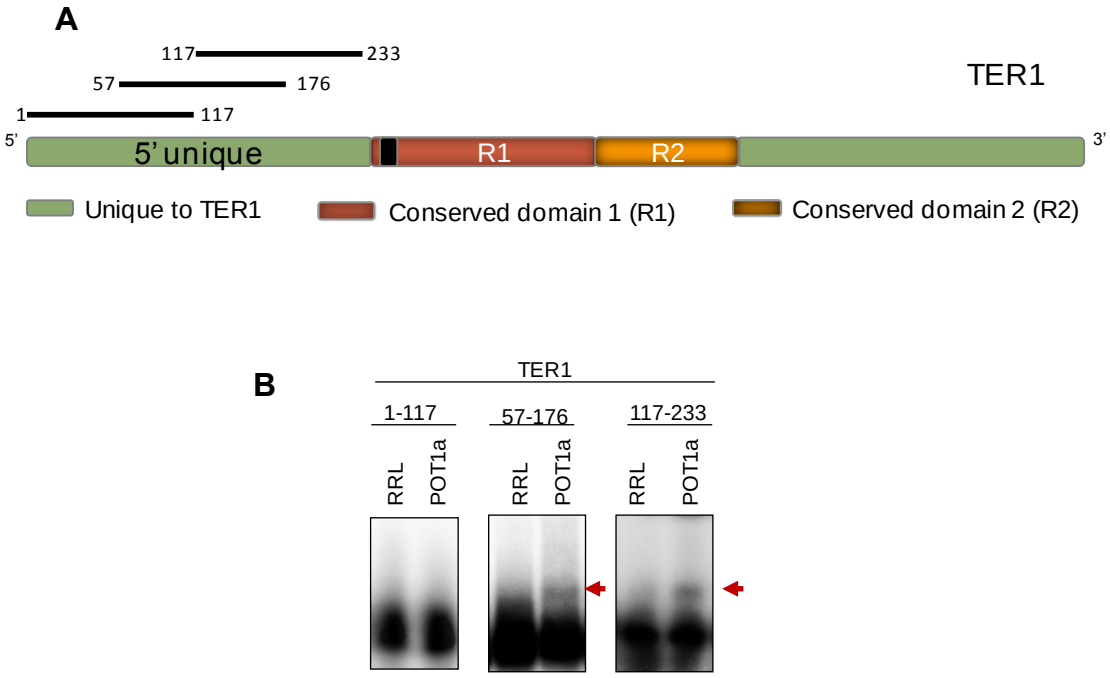


Figure 33. POT1a associates with a region upstream of TER1 template. (A) Diagram of TER1 indicating overlapping constructs within the TER1 5' unique region. (B) Results from electrophoretic mobility shift assays. Labeled RNAs are indicated. RRL alone and TRFL4, a ds telomeric DNA binding protein, were used as controls.

We next examined 60nt overlapping constructs covering the 117-233 region. No tested binding to POT1a was observed (data not shown), suggesting that the structure of TER1 may be important for POT1a binding. Since our secondary structure model predicts that TER1 folds into modular independent domains, Dr. Tzfati used Mfold to generate a secondary structure prediction for the TER1 5' unique region (http://mfold.bioinfo.rpi.edu/cgi-bin/rna-form1.cgi) (342).

Nine structures were predicted for this region. Secondary structure elements within these predictions were analyzed. A smaller fragment (117-242) was fed into MC-fold program (http://www.major.iric.ca/MC-Pipeline/), which provides secondary structure prediction, but also include non-canonical base pairing patterns and tertiary structure information (343). The resulting structure is shown in Fig. 34. In this structure there are two main stem-loop elements separated by a spacer, part of which is predicted to pair with one of the loops to form a pseudoknot. The templating domain is located immediately 3' of this fragment and it is predicted to be in a single-stranded conformation (Fig. 34). We divided the structure in three domains: a 5' stem-loop, a ss linker and a 3' stem loop and tested these domains for binding to POT1a.

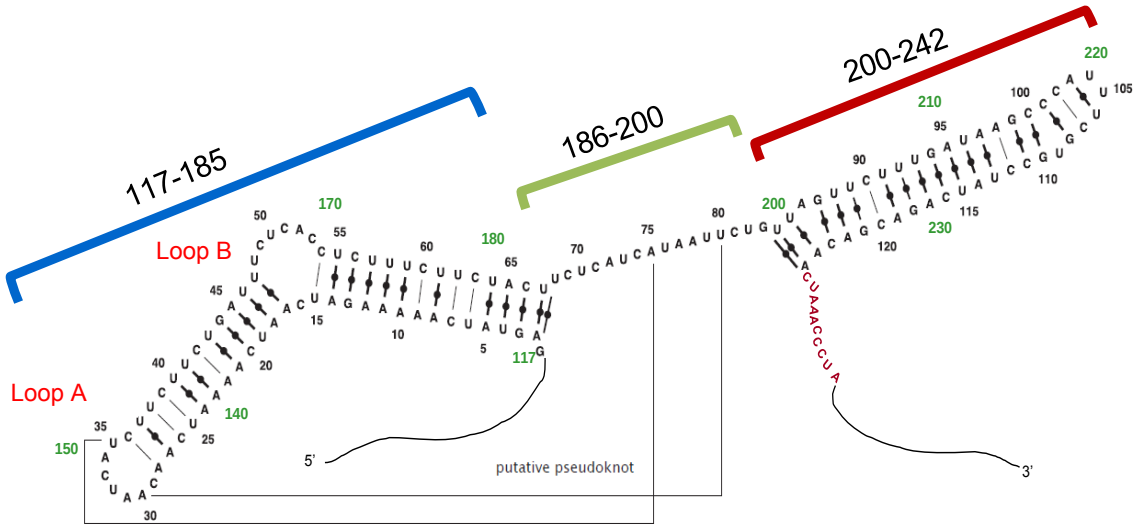


Figure 34. Secondary structure prediction the TER1 region bound by POT1a. This structure was generated the MC-fold software, which gives additional information about RNA tertiary structure (343). The template region is indicated in red font.

Interestingly, Pot1a binding was only observed with those constructs containing the 5' stem-loop (Fig. 35C). The 3'-stem is very stable and predicted in most of the nine alternative structures. The failure of AtPOT1a to bind this 3'-stem is consistent with binding analysis of a fragment comprising nucleotides 184-233 (data not shown).

Since OB-fold domain proteins bind ss nucleic acids, we hypothesized that POT1a was likely to associate with a ss stretch in TER1. There are three potential POT1a binding sites in the 117-242 structure: two loops within the 5' stem-loop (nucleotides 117-185) and the ss linker (186-200) (Fig. 34). *In vitro* binding experiments showed that although the nucleotide composition of these three regions is highly similar, POT1a does not associate with the linker domain (Fig. 35D). Thus, the requirements for POT1a ligand recognition appear to rely on the RNA structure or a combination of structure and sequence. This finding also raises the possibility that the linker is required to form the predicted pseudoknot structure and ensure a proper conformation.

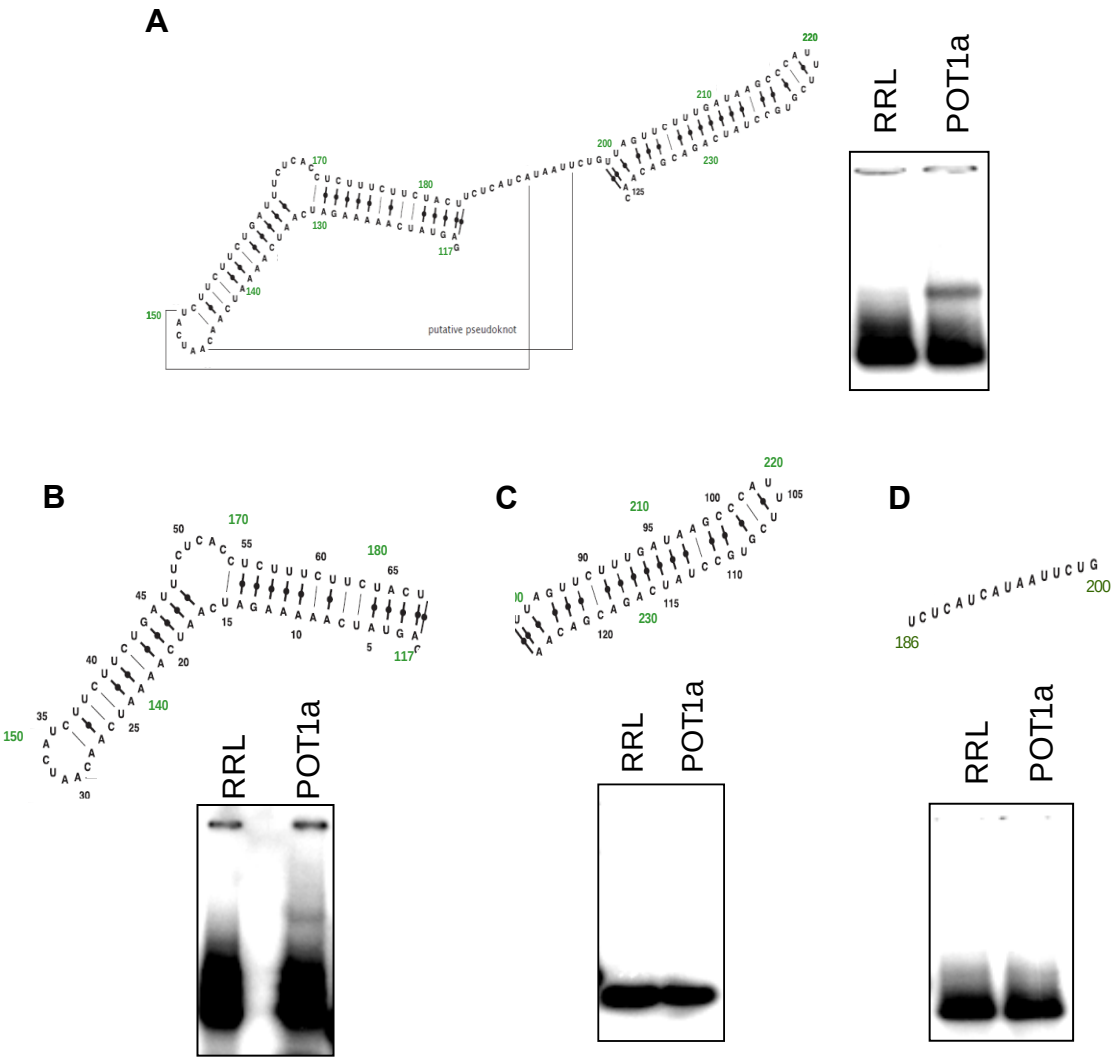


Figure 35. POT1a binds to a loop-stem-loop in the 5' end of TER1. The POT1a binding site in TER1 was mapped using a combination of deletional mutagenesis and EMSA. (A-D) Top, predicted secondary structure of the TER1 constructs. Bottom, EMSA results with the corresponding radiolabeled RNA and POT1a. RRL alone was used as a negative control.

We designed constructs containing nucleotide substitutions or deletions in one or both loops within the 5'-stem-loop structure (Fig. 34 and 36A). Loop A corresponds to nucleotides 144-152 and loop B to nucleotides 130-133 and 163-170. The effect of these changes was monitored by Mfold and the constructs were tested for POT1a association by filter binding (Fig. 36A). POT1a only associated with a construct that retained the wild type sequence and structure of the B loop (Fig. 34 and 36A) and which lacked loop A. No binding was detected for the other constructs, suggesting that POT1a associates with the B loop in TER1. The only nucleotide change within the single-stranded region of the loop that affected binding was cytidine 167 to adenine, indicating that this residue may be a major contact for POT1a (Fig. 36A). Finally, mutation of the uracils 162-164 to adenine in the duplex region upstream of the C-U rich loop, to expand the loop abolished binding. Thus, binding site architecture is important for POT1a association.

In summary, we obtained a minimal POT1a ligand of 42 nt (including three uridils to favor formation of the structure) derived from the 150nt immediately upstream of the template in TER1, which is composed of a C-U rich loop flanked by two 10-base pair stems (Fig. 36B). Although, further experiments are required to dissect the absolute minimal requirements for the POT1a-TER1 interaction, this 42 nt ligand is suitable for co-crystallization studies.

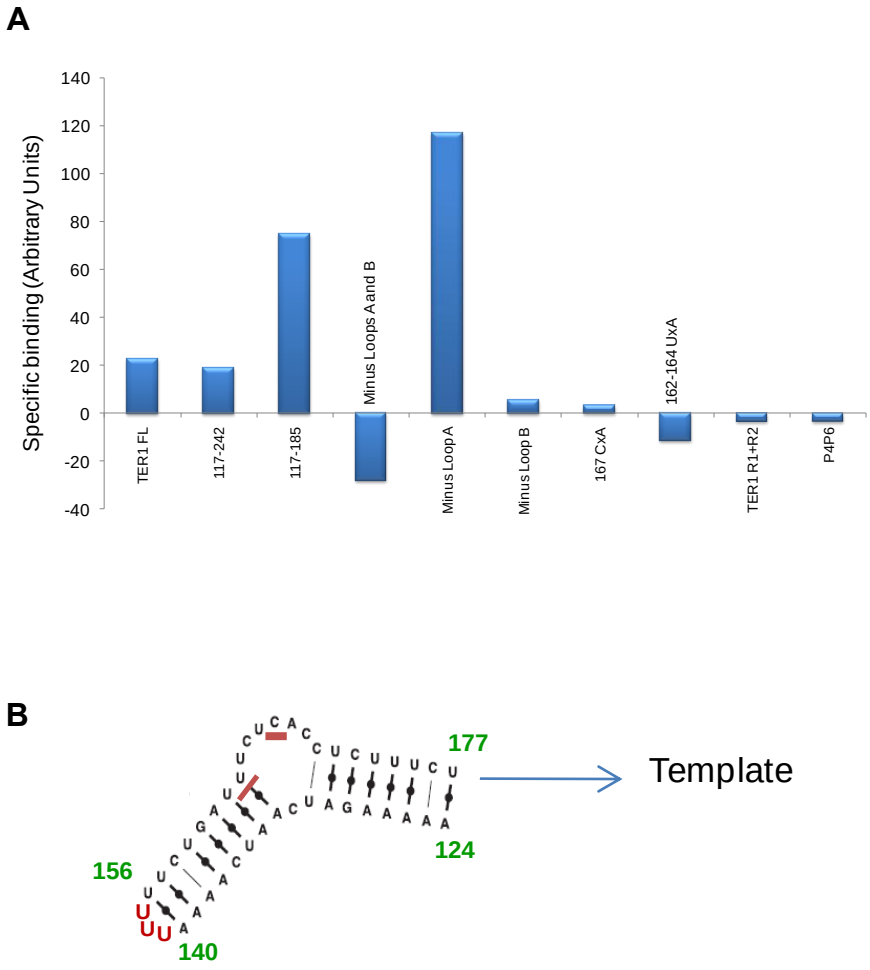


Figure 36. POT1a binds to a bulged stem in TER1. Additional constructs were tested by filter binding assay to establish the POT1a binding site in TER1. (A) Quantification of filter binding data. (B) Minimal POT1a-TER1 binding site. Three Uridils (red font) were added to facilitate secondary structure formation. Cytidine 167, as well as the three Uridils that disrupted POT1a association are underlined in red.

The POT1a binding site is conserved in TER1 from other plant species

Although POT1a orthologs from several plant species have been identified, their nucleic acid binding specificity is unknown. Notably, analysis of POT1a from a wide range of eudicots failed to reveal DNA binding activity *in vitro* (131) (Fig. 37). Thus we hypothesize that RNA binding activity may be a conserved property of POT1a from eudicots. In collaboration with Dr. Mark Beilstein, putative TER1 sequences from eight close relatives of *Arabidopsis* were obtained by High-efficiency thermal asymmetric interlaced PCR (TAIL-PCR), using degenerate primers designed based on the TER1 genomic region (see Materials and Methods). If TER1 binding is a conserved function of plant POT1a, we would expect similarities in the secondary structure and location of the POT1a binding site. The putative TER1 sequences from *Arabidopsis* close relatives (Fig. 38) were aligned using Clustal 2.0.10 software (EMBL-EBI). All sequences showed conservation from the template region through the TER1 region overlapping the putative SnRK1-interacting protein 1 coding sequence. However alignment of the region upstream of the template is much less robust (Fig. 39) and suggests that this region of TER1 is hypervariable.

Based on the location of the AtPOT1a binding site in TER1, we performed secondary structure predictions of the 150nt upstream of the template of putative TER1 subunits using the MC-fold software (343). Several alternative predictions were generated for each TER1. Among these predictions, we found a consensus structure in three out of eight species (Fig. 40).

Overall the structure can be described as two duplex regions flanking a C-U rich loop (Fig. 40 A-B). In two cases the C-U rich loop is located at the end of a stem insertion (Fig. C-D). Interestingly, the prediction for *Brassica* TER1 is highly similar to its *Arabidopsis* counterpart. Moreover, the three U-A base pairs closest to the loop, which are important for interaction of AtPOT1a with AtTER1 show co-variation in *Brassica* TER1.

In summary, these findings suggest that the stem is conserved in several plant species and hence may represent a functional site in TER1 for binding of POT1a.

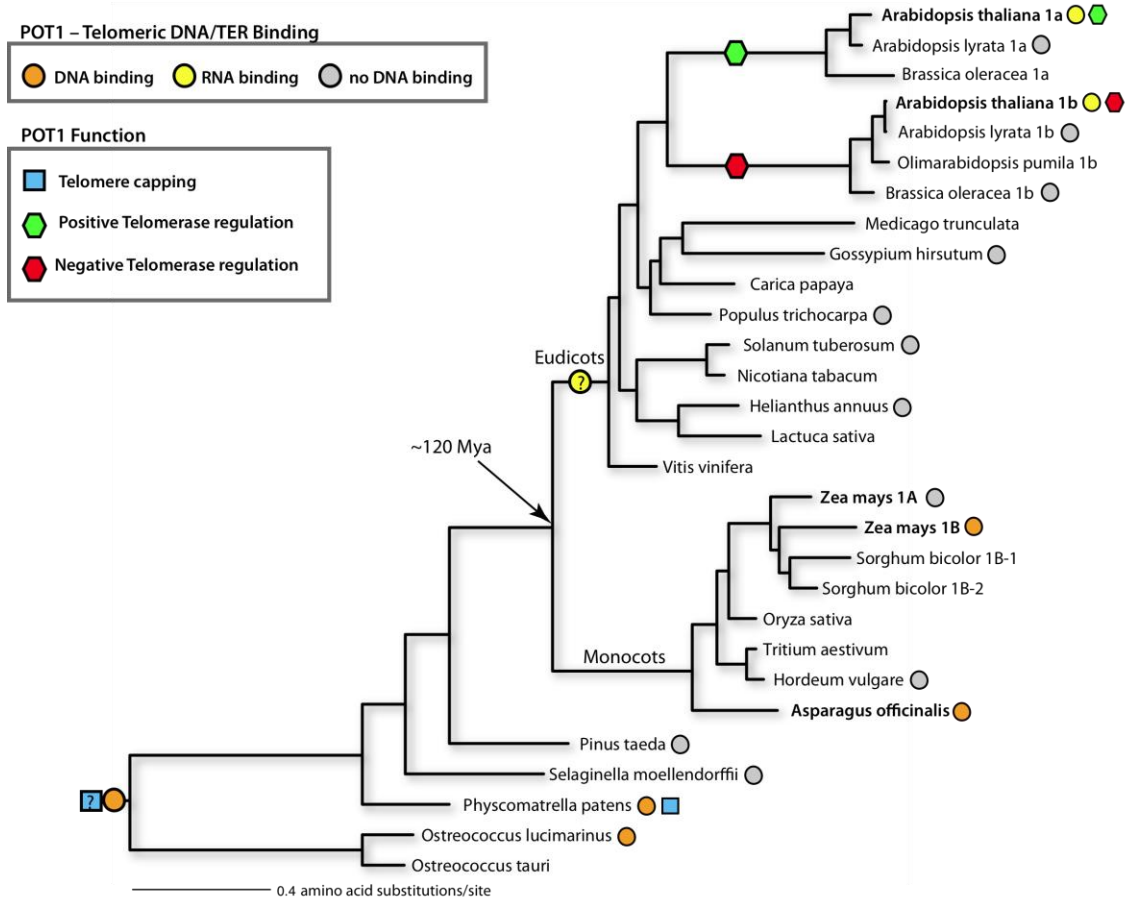


Figure 37. Phylogenetic tree of plant POT1 proteins. DNA and RNA binding activities, as well as capping function of known plant POT1 proteins are indicated. 1a, POT1a. 1b, POT1b. (Generated by Dr. Mark Beilstein).

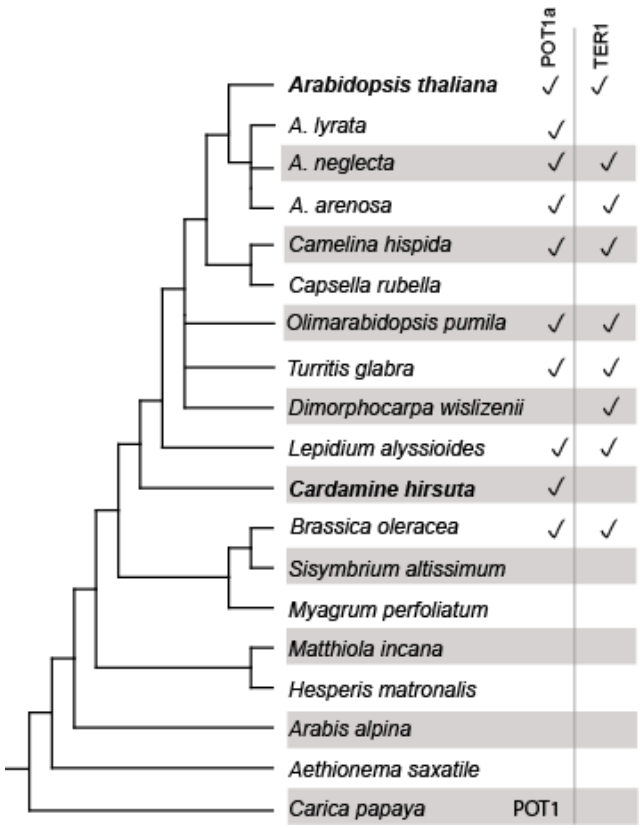


Figure 38. Phylogenetic tree of the Brassicaceae family. Composite tree based on *Beilstein* et al. 2006 & *Beilstein* et al. 2008 that approximates the organismal phylogeny of *Brassicaceae.* ✓, indicates the positive identification of putative TER1 or POT1a orthologs. (Generated by Dr. Mark Beilstein)

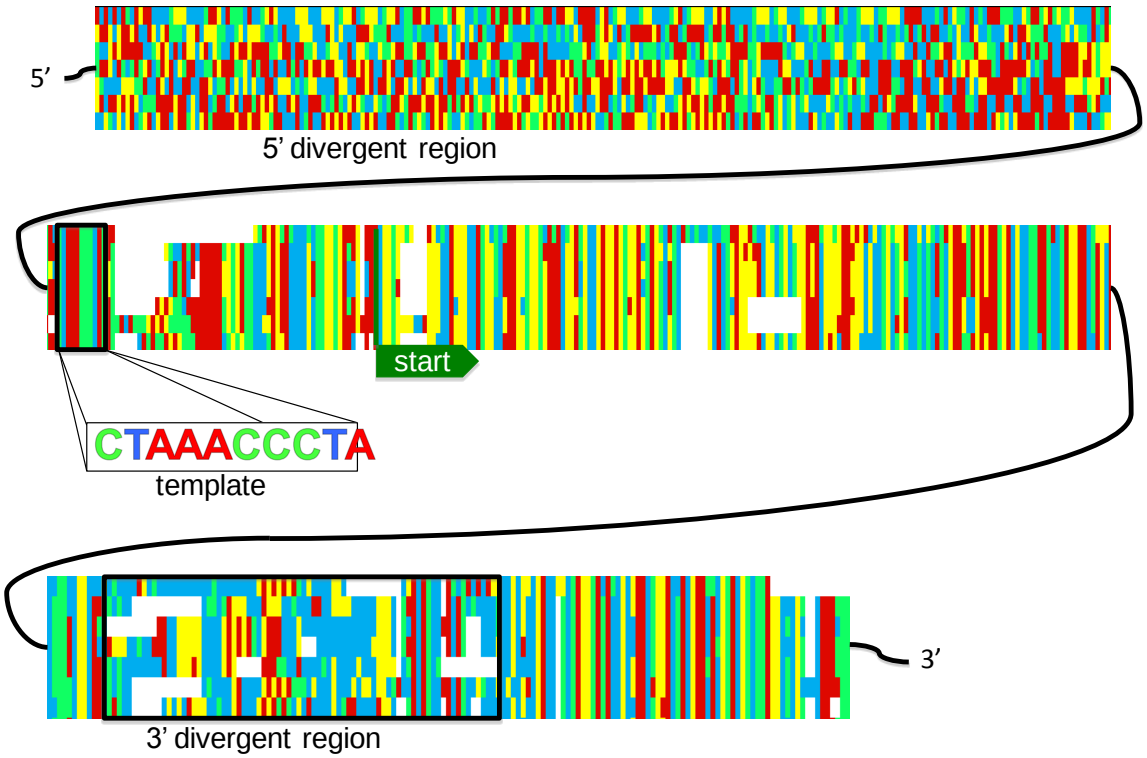


Figure 39. Sequence alignment of putative TER1 sequences of *Arabidopsis* close relatives. Color codes are assigned to nucleotides to indicate sequence conservation: Green, C; Blue, T; Red, A; Yellow, G. TER1 sequences from seven different plant species are aligned. The conserved template domain, as well as the hypervariable 5’ and 3’ regions are indicated. Start corresponds to the start codon of the overlapping SnRK1 protein coding gene (Generated by Dr. Mark Beilstein).

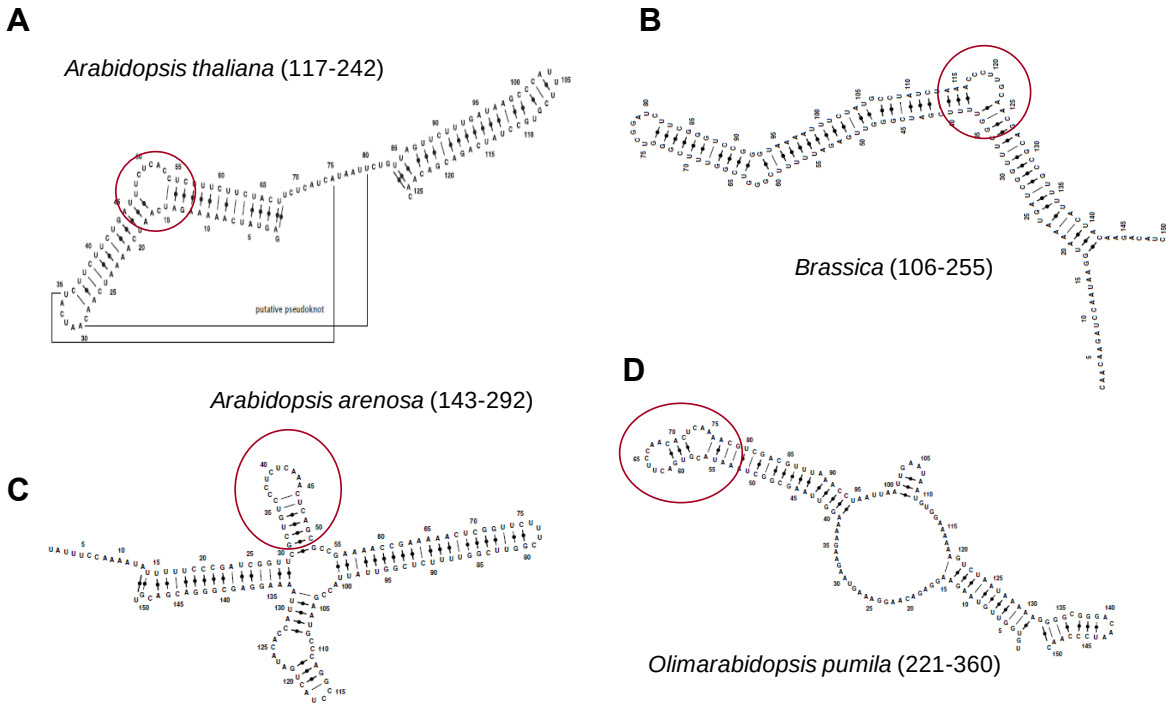


Figure 40. Secondary structure prediction for *Arabidopsis* close relatives. C-U rich loops were detected in the 5' divergent region of putative TER1 (red circle). MC-fold software was used to predict the structures (343).

POT1a OB-folds are responsible for TER1 binding

The crystal structure of human POT1a bound to its ss DNA substrate shows that the OB-fold domains are responsible for ligand recognition (122). We conducted binding studies with TER1 and a limited series of POT1a truncation mutants to map the RNA binding site in POT1a (see Chapter II, Fig. 18A). POT1a association with TER1 was dependent on the N-terminal half of the protein, which is predicted to encode two OB-folds. No signal above background level was detected with the C-terminal portion of the protein, although binding was detected with the first OB-fold alone, binding detected with both OB-1 and OB-2 was about two-fold higher than with OB-1 alone. We conclude that both OB-1 and OB-2 contribute to recognition of TER1.

A phenyalanine residue, F62, located within OB-1 plays a critical role in the telomeric DNA interaction with human POT1 (119, 122). This residue is conserved in POT1 proteins from different organisms and in *Arabidopsis* POT1a it corresponds to F65 (Fig. 41A). Superimposition of AtPOT1a OB1 onto the crystal structure of hPOT1 predicts F65 resides within the recognition pocket (Croy, J. and Wuttke, D., University of Colorado, Fig 41B). Therefore, we tested if AtPOT1a F65 is important for TER1 recognition. A point mutation in OB-1, F65A, reduced the affinity of POT1a for TER1 *in vitro* (data not shown). In contrast no effect on affinity was observed when another highly conserved phenylalanine also located within OB-1 (data not shwon), but predicted outside

the recognition pocket, was mutated. We conclude that AtPOT1a recognition of TER1 may share characteristics of the POT1-ssDNA recognition.

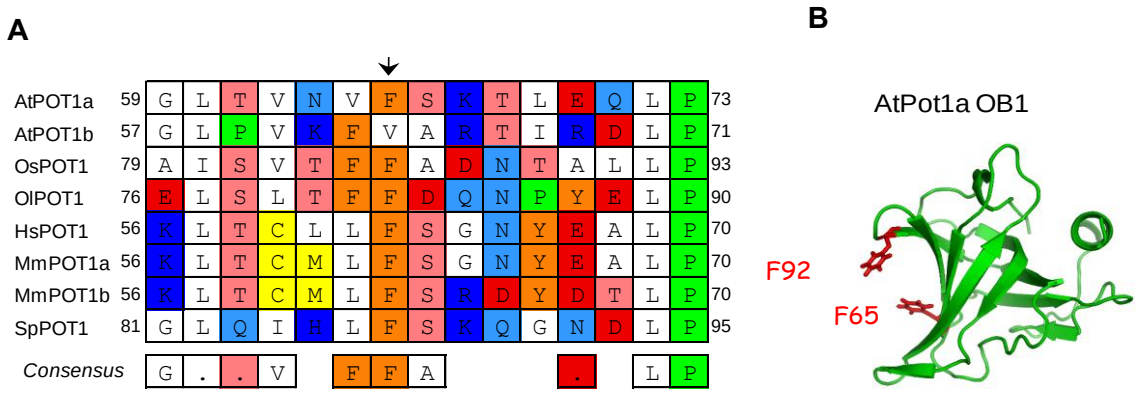


Figure 41. AtPOT1a Phe65 may be important for TER1 recognition. Alignment of POT1 proteins from different species. At, *Arabidopsis thaliana*; Os, *Oryza sativa*; Hs, *Homo sapiens*; Mm, *Mus musculus*; Sp, *Schizosaccharomyces pombe*. Arrow indicates the position of the highly conserved phenylalanine corresponding to F65 in AtPOT1a. (B) Structural docking of AtPOT1a OB-fold onto hPOT1 OB-fold. (Generated by Croy, J. and Wuttke, D., University of Colorado). The position of F65 and F92 are indicated.

Discussion

POT1 proteins are a conserved group of OB-fold-containing proteins that bind ss telomeric DNA and play essential roles in chromosome-end protection and telomere length regulation. The main functions of POT1 appear to be conserved among most eukaryotes. *Arabidopsis* POT1 is a notable exception. None of the three POT1 paralogs in this organism binds telomeric DNA (131). Instead, these proteins have evolved a capacity for RNA binding as well as functions in telomerase regulation (129-130, 313).. To further characterize the POT1a-TER1 interaction, and to obtain a protein-RNA complex amenable for co-crystallization experiments, we sought to determine the minimal POT1a binding site in TER.

Functional analogies between POT1a and EST1

We determined that the minimal POT1a binding site is located in the 5'-unique region of TER1. The binding site consists of a C/U-rich loop flanked by two 10 base-pair stems, approximately 70 nt immediately upstream of the template sequence. Mutation of a single C residue in the loop region or base-paired U residues in the upstream stem leads to loss of POT1a binding. This suggests that POT1a makes both structure and sequence-specific contacts with TER1.

The architecture of the POT1a binding site resembles the yeast Est1 binding site in Tlc1 (Fig 42). It is a highly conserved bulged stem structure that

was initially identified by sequence alignment of telomerase RNAs of seven yeast species. Mutagenesis studies indicate that this bulged stem structure is essential for telomerase function *in vivo*, by serving facilitating the association of Est1p, a telomerase recruitment factor (181). Like POT1a, Est1 is a component of yeast telomerase that binds the RNA subunit, Tlc1 (181). Antibodies raised against Est1 immunoprecipitate Tlc1 as well as telomerase activity (344-345). In addition, *est1* mutants also show the ever shorter phenotype observed in the telomerase reverse transcriptase, *est2* mutants (17).

There is evidence for diverse proteins from different organisms playing conserved roles in telomere biology. For example, our lab showed that *Arabidopsis* CTC1 is a functional homolog of a component of the yeast CST complex, Cdc13, a G-overhang DNA binding protein that recruits the telomerase RNP via a direct interaction with Est1 (98, 247). In line with this, preliminary results suggest that POT1a interacts with the C-terminal region of CTC1. Moreover active telomerase RNPs, containing at TERT, TER1 and POT1a, co-purify with tagged CTC1 (Song X and Shippen D. unpublished data). Together, our results suggest that *Arabidopsis* POT1a may be a functional homolog of yeast EST1 and further suggest that POT1a may function as a telomerase recruitment factor through its interaction with CTC1.

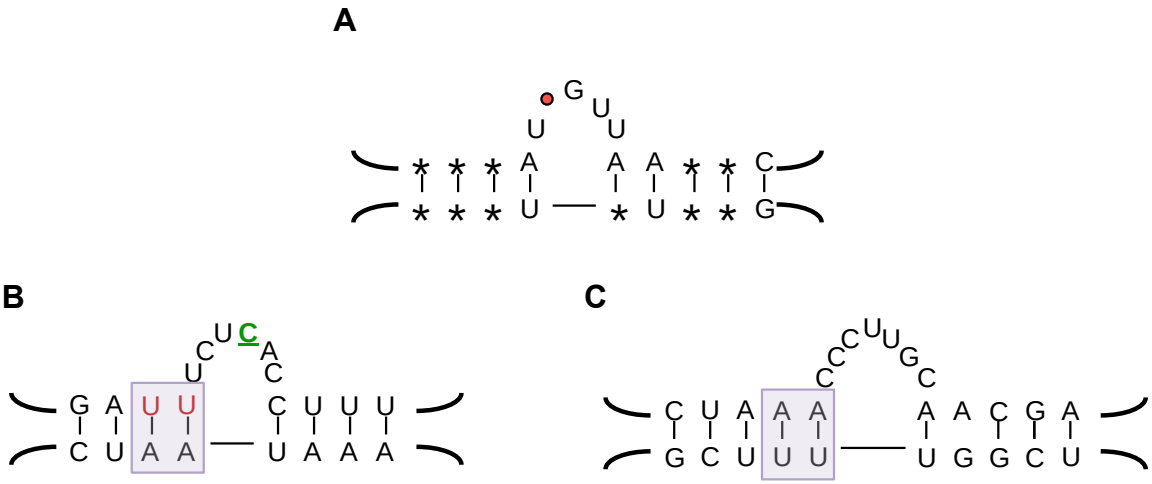


Fig. 42. Structural similarities between the Est1-Tlc1 and POT1a-TER1 binding sites. (A) Est1 binding site in Tlc1 (Seto, et al 2002). Phylogenetically conserved sequence in the bulged stem of Tlc1. Indicated nucleotides are conserved between yeast species; * indicates nucleotides that show co-variation; and red circle indicates a position in the bulge where the sequence is not conserved. (B) Predicted structure of the POT1a-TER1 binding site in *Arabidopsis thaliana*. Red font, mutations in these nucleotides abolishes binding. Underlined green C, Cytidine 167. Purple square, indicates co-variation between *Arabidopsis thaliana* and *Brassica*. (C) Predicted Brassica POT1a-TER1 binding site.

The OB-folds in POT1a mediate its interaction with TER1

The OB-fold is a conserved protein domain consisting of a five-stranded beta-sheet coiled to form a closed beta-barrel that is capped by an alpha-helix between the third and fourth strands (70). It is found in many nucleic acid-binding proteins and can also mediate protein-protein interactions (70). Proteins of the POT1 family contain two N-terminal OB-folds that in most organisms serve as major points of contact ss DNA. *Arabidopsis* POT1a shows a different nucleic acid-binding specificity, binding TER. We found that this interaction is also mediated through OB-folds (Chapter II, Fig 18A). Mutation of F65, a conserved Phenylalanine and a major determinant in the recognition of ss telomeric DNA by POT1 from yeast and vertebrates, leads to reduced TER1 binding and only a partial functional complementation *in vivo* (Song X and Shippen D. unpublished data). This observation suggests that POT1a-TER1 and POT1-DNA complexes may assemble in a similar manner. Interestingly, it was recently reported that the corresponding Phe in human POT1 (Phe 62) plays a major role in preventing RNA binding (272). Thus, the structural characterization of the POT1a-TER1 complex may reveal how *Arabidopsis* POT1a evolved its ability to bind RNA.

The POT1a binding site is conserved in other plant TER1s

Phylogenetic analysis revealed the existence of putative TER1 homologs in various *Arabidopsis* close relatives. Secondary structural prediction showed

conserved bulged-stem structures upstream of the putative templating domain that are the hallmarks of a POT1a binding site. Although additional experiments are required to establish the functional significance of these conserved structural motifs, our results suggest that the interaction between POT1a and TER1 may not be restricted to *Arabidopsis*, but may represent a fundamental function of POT1a from eudicots. Intriguingly, the Est1 binding site of Tlc1 is also conserved among various fungal species (309, 346), which supports our hypothesis of the functional homology between yeast Est1 and POT1a as well as the co-evolution of plant POT1a and TER1.

Materials and Methods

In vitro *binding assays*

Electrophoretic mobility shift assays were performed with RNA transcribed *in vitro* with T7 RNA polymerase and [α-^{32}P]-CTP labeled TER1. Binding reactions contained 3µl of RRL expressed protein, 0.1pmol of 32P labeled TER and 1X binding buffer (25mM Tris-HCl pH 8.0, 10mM $Mg(OAc)_2$, 25mM KCl, 10mM DTT and 5% Glycerol) in a 30µl final volume. 1µM yeast tRNA and 0.5µM RNA $(U_3AG_3)_4$ were used as nonspecific competitors. After 20 min at 30°C the reaction was loaded onto a 0.8% agarose 0.5X TBE gel and run for 2 h at 70 volts at 4°C. Gels were dried and exposed to phosphorimager screens.

For double-filter binding assays, RRL expressed protein was incubated with decreasing concentrations of pre-folded RNA transcribed *in vitro* using 5-end ^{32}P-labeled RNA as a tracer. Binding reactions contained 0.5µl of recombinant protein, pre-folded TER in binding buffer (50mM Tris-HCl pH 7.5, 200mM potassium glutamate, 0.5mg/ml BSA, 0.5mg/ml tRNA, 1mM MgCl2, 1mM DTT and 0.01% NP-40) in a 25µl final volume. After 30 min at 30°C, the reactions were filtered through nitrocellulose and nylon filters using a dot-blot apparatus (BioRad). The membranes were washed with 600µl washing buffer (50mM Tris-HCl pH 7.5, 200mM potassium glutamate, 1mM $MgCl_2$, 1mM DTT and 10% glycerol), dried, exposed to a phosphor storage screen and scanned after 2h. Equilibrium dissociation constants, Kd, were obtained by non-linear regression of the binding data fitted to a one-site binding model using Graphpad Prism software.

Structure prediction of plant POT1 proteins

The secondary structure of each plant POT1 protein sequences was determined using the secondary structure prediction server PSIPRED v2.6 (347). Generation of theoretical structural models for the OB1 domains present in each plant POT1 protein was accomplished using threading techniques. Best structural templates for each plant POT1 protein was selected based on the best alignment score generated from a CLUSTAL alignment (348), which was found to be the N-terminal OB1 domain present in *Oxytricha nova*. Optimum sequence

alignments were done using PROMALS (349), which couples primary sequence homology with secondary structure prediction to align two sequences that are poorly conserved in primary sequence. PROMALS aligned sequences were submitted to the SWISS-MODEL threading server (www.swissmodel.expasy.org) (350). Generated models were then subjected to a GROMOS96 energy minimization to adjust bond lengths, angles and geometries using Swiss-PDB Viewer (351).

CHAPTER V

INITIAL INVESTIGATIONS INTO THE BIOGENESIS OF $TER2_S$

Summary

Telomeres are nucleoprotein complexes that protect the ends of linear chromosomes from being recognized as double-strand breaks. Telomeres are replicated and maintained by a ribonucleoprotein reverse transcriptase called telomerase, which uses its essential RNA subunit (TER) to direct synthesis of telomeric repeats. *Arabidopsis* encodes three TER isoforms, TER1, TER2 and $TER2_S$, a processed from of TER2. Here we present the initial characterization of TER2s production. We show that $TER2_S$ is a bona fide TER isoform, and the product of TER2 self-splicing. Notably, the processing of TER2 is not mediated by a group I or II self-splicing intron mechanism, since $TER2_S$ production does not require a guanosine or Mg^{2+} cofactor. TER2 processing results in lower levels of active telomerase from reconstituted RNPs. These findings suggest that TER2 self-splicing may be a novel mechanism for regulation of telomerase RNP *in vivo*.

Introduction

Organisms with linear chromosomes rely on telomeres for chromosome end protection. Due to the end replication problem, telomeres cannot be fully replicated by the conventional DNA replication machinery. Instead, the ribonucleoprotein telomerase is charged with replenishing the terminal sequences that are lost with each cell division (352). In most cell types, the action of telomerase is not sufficient to compensate for telomeric DNA loss, ultimately resulting in cellular senescence after a determined number of cell divisions (200, 353). Increased levels of telomerase are not the solution for cell aging either, since high levels of telomerase can fuel cancer (190, 270, 354).

Telomerase, at its core, is composed of a reverse transcriptase subunit (TERT) and an RNA component (TER) that serves as template for nucleotide addition to the single-stranded G-rich telomeric 3' overhang (134, 149, 355-356). TERT, the catalytic protein component of telomerase, is highly conserved among most eukaryotes and resembles viral reverse transcriptases in its catalytic core (357). In contrast, TER molecules are extremely divergent in sequence and size: ~150 nt in ciliates (14), ~450 nt in humans (231), ~750nt in *Arabidopsis* (313) and Chapter II) and ~1300 nt in budding yeast (170). Mutations in human TER result in the genetic disorder dyskeratosis congenita, a devastating disorder characterized by insufficient renewal capacity of stem cells in skin and bone marrow (184, 358-360).

Phylogenetic studies and mutational analysis reveal that TER moieties from different organisms share functionally conserved secondary structure elements (169-170, 340-341, 361), including a single-stranded templating domain typically corresponding to one and a half telomeric repeats flanked by a 5' boundary element and a 3' pseudoknot domain (172). Vertebrate TER subunits harbor an additional CR7 domain and a box H/ACA snoRNA motif, which binds dyskerin and is required for RNP maturation and TER nuclear localization (147). In contrast, the budding yeast TER assembles with Sm-proteins and is processed as a snRNA (246).

Recent studies *in Saccharomyces pombe* indicate that processing of the 3' end of TER1 is required for telomerase function (324). A poly(A) site is found in the TER1 locus and TER1 transcripts bearing a poly(A) tail are detected (171, 324), However, the TER1 transcripts that associate with telomerase do not contain a poly(A) tail. Further analysis showed that processing of the SpTER1 3' end is mediated by the spliceosome. Strikingly, the mechanism of processing is not via conventional splicing: only the first transesterification reaction occurs, releasing the active form of the RNA without exon ligation. This process has been termed slicing (324). Failure to process SpTER1 results in telomerase inhibition. Thus, RNA processing may play a key role in telomerase biogenesis and function.

We have shown that *Arabidopsis* telomerase associates with a processed form of TER2, called $TER2_S$ (Chapter III). Cloning and sequencing experiments

indicate that $TER2_S$ is composed of the two conserved regions of TER2, joined together (Chapter III). The exact 5' and 3' ends of this transcript are still unknown. $TER2_S$ associates with TERT *in vitro* and *in vivo* (Chapter III) and telomerase activity can be reconstituted with a transcript corresponding to the conserved regions of TER2 (Chapter II). However, it is TER1 and not TER2 or its processed form, $TER2_S$ that functions as the primary template of telomerase for telomere maintenance *in vivo* (Chapter III). Thus, the role of TER2 processing is unknown. Despite this lack of understanding regarding $TER2_S$ function, the production of this TER isoform is developmentally regulated. Furthermore, $TER2_S$ associates with POT1b, a negative regulator of telomerase activity (Chapter III). These findings suggest that TER2 processing is likely to play a role in telomere biology *in vivo*.

In this chapter, we present an initial characterization of TER2 processing. We show that $TER2_S$ is a bona fide TER isoform product of TER2 self-splicing. It is produced *in vitro* independently of the addition of TERT or any other protein. Also, we show that the production of $TER2_S$ does not require guanosine or Mg^{2+} as cofactors, suggesting that the mechanism of TER2 processing is not related to that of group I or II introns. This finding suggests that $TER2_S$ production may occur via a Hepatitis Delta Virus or hairpin ribozymes self-splicing mechanism. Finally, production of $TER2_S$ negatively affects telomerase activity of *in vitro* reconstituted in RNPs, arguing that TER2 processing can provide a mechanism for regulation of telomerase activity.

Results

$TER2_S$ is a bona fide TER isoform

Alignment of the $TER2_S$ sequence with TER1 and TER2 showed 100% identity to the conserved region of TER2. As expected, about 90% similarity was observed with the conserved regions of TER1 (Fig. 43). Computational analysis of plant splice sites using the NetGene2 software (362-363) did not predict the characteristic landmarks of consensus mRNA splicing machinery (branch point, donor and acceptor splice sites) in TER2 (data not shown), indicating an alternative mechanism of processing. It should be noted, however, that plant mRNAs show a great variation of consensus plant splice sites and branch points regions, despite the high conservation of the splicing machinery throughout eukaryotes (322-323, 364).

It is possible that $TER2_S$ is an artifact of reverse transcription *in vitro* due to the presence of stable RNA structures that are not disrupted during the experimental protocol or to misannealing of the primer used for reverse transcription. To examine this possibility, employed a combination of northern blotting (which is a direct detection method that gives information on molecular size) and secondary structural analyses of TER2.

A

```
Score =  233 bits (126),  Expect = 6e-66
Identities = 195/226 (87%), Gaps = 13/226 (5%)
Strand=Plus/Plus

TER1     234  GACGACAACTAAACCCTACACGCTTACATAAAAACGCGACCTTTGAGAATTTTGTCCGGC  293
              ||||||||||||||||||  |||||||||||||||||||||||||| ||||||||||  ||||
TER2S    1    GACGACAACTAAACCCT--ACGCTTACATAAAAACGCGACCTTTAAGAATTTTGTACGGC  58

TER1     294  GACAGAAATGGCGGGGTTAGGGTT-ACGGCTAAAGGCGGCAAAATGGACGTTGAGGAGCG  352
               ||||  ||    | | |   |  ||| ||||||||||||||| |||||||||||||  |||||
TER2S    59   AACAG-AA--CCAGAG--AACGTTGACGGCTAAAGGCGGCGAAATGGACGTTGATGAGCG  113

TER1     353  GATCAGGTGCTGTATCGAGAGAAT-G--G--TCGTCGGAGATGGGGAAAGGTGTTCGTCG  407
              || ||| | ||||||||||  | || |   |  ||||||||||||||||| ||||||||||||
TER2S    114  GAGCAGCTACTGTATCGAGGGTATGGTCGTCTCGTCGGAGATGGGGAAGGGTGTTCGTCG  173

TER1     408  ATTCTCTACTGAAACAGAGAACGATGTTCCAACTTCAGGAATCAGT  453
               ||||||||||||  |||||||||||||||||||||||||||||||
TER2S    174  GTTCTCTACTGAACCAGAGAACGATGTTCCAACTTCAGGAATCAGT  219
```

B

```
Score =  268 bits (145),  Expect = 2e-76
Identities = 145/145 (100%), Gaps = 0/145 (0%)
Strand=Plus/Plus

TER2     1    GACGACAACTAAACCCTACGCTTACATAAAAACGCGACCTTTAAGAATTTTGTACGGCAA  60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TER2S    1    GACGACAACTAAACCCTACGCTTACATAAAAACGCGACCTTTAAGAATTTTGTACGGCAA  60

TER2     61   CAGAACCAGAGAACGTTGACGGCTAAAGGCGGCGAAATGGACGTTGATGAGCGGAGCAGC  120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TER2S    61   CAGAACCAGAGAACGTTGACGGCTAAAGGCGGCGAAATGGACGTTGATGAGCGGAGCAGC  120

TER2     121  TACTGTATCGAGGGTATGGTCGTCT  145
              |||||||||||||||||||||||||
TER2S    121  TACTGTATCGAGGGTATGGTCGTCT  145


Score =  143 bits (77),  Expect = 1e-38
Identities = 77/77 (100%), Gaps = 0/77 (0%)
Strand=Plus/Plus

TER2     674  TCGTCGGAGATGGGGAAGGGTGTTCGTCGGTTCTCTACTGAACCAGAGAACGATGTTCCA  733
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TER2S    145  TCGTCGGAGATGGGGAAGGGTGTTCGTCGGTTCTCTACTGAACCAGAGAACGATGTTCCA  204

TER2     734  ACTTCAGGAATCAGTTT  750
              |||||||||||||||||
TER2S    205  ACTTCAGGAATCAGTTT  221
```

Figure 43. $TER2_S$ sequence is 100% identical to TER2. (A) Alignment of TER1 with $TER2_S$. (A) Alignment of TER2 with $TER2_S$.

We performed northern blotting using a probe targeting the conserved region in TER2. A product of ~200 nt that corresponds to $TER2_S$ was detected (Fig. 44). RT-PCR indicates that $TER2_S$ is more abundant in flowers than in cell culture (Chapter III), this is observed in the Northern blot. Although precise mapping of the 5' and 3' ends of the $TER2_S$ transcript is still needed, this result indicates that $TER2_S$ is composed mainly of the TER2 conserved region. We also detected a product of ~750 nt, which could correspond to TER1. Alternatively, this band may correspond to both TER1 and TER2, but agarose electrophoresis may not have resolved the two products due to their similar length.

Finally, RNA folding predictions using the Mfold software show that the TER2 regions comprising the $TER2_S$ splice junction are not in close proximity or part of a common structural element in any of the TER2 predicted structures that could lead to a stable secondary structure that is resistant to unfolding during reverse transcription (Fig. 45). Altogether, these data argue that $TER2_S$ is not an aberrant reverse transcription or PCR product, but rather a bona fide transcript. This conclusion is further supported by the developmental regulation of TER2 processing (Chapter III).

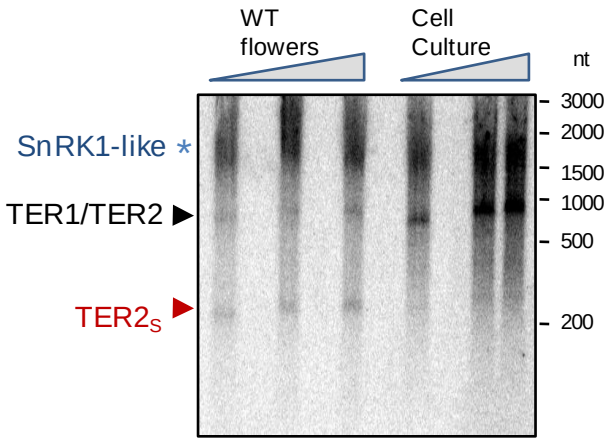


Figure 44. TER2s is a bona fide transcript. Northern blot from wild type flowers and cell culture total RNA (20µg, 40µg and 60µg) using a radiolabeled probe complementary to the conserved region in TER2. Red arrow head, $TER2_S$. Black arrow head, TER1/TER2. Blue asterisk, SnRK1-like mRNA is also detected because the TER conserved regions are embedded in this gene.

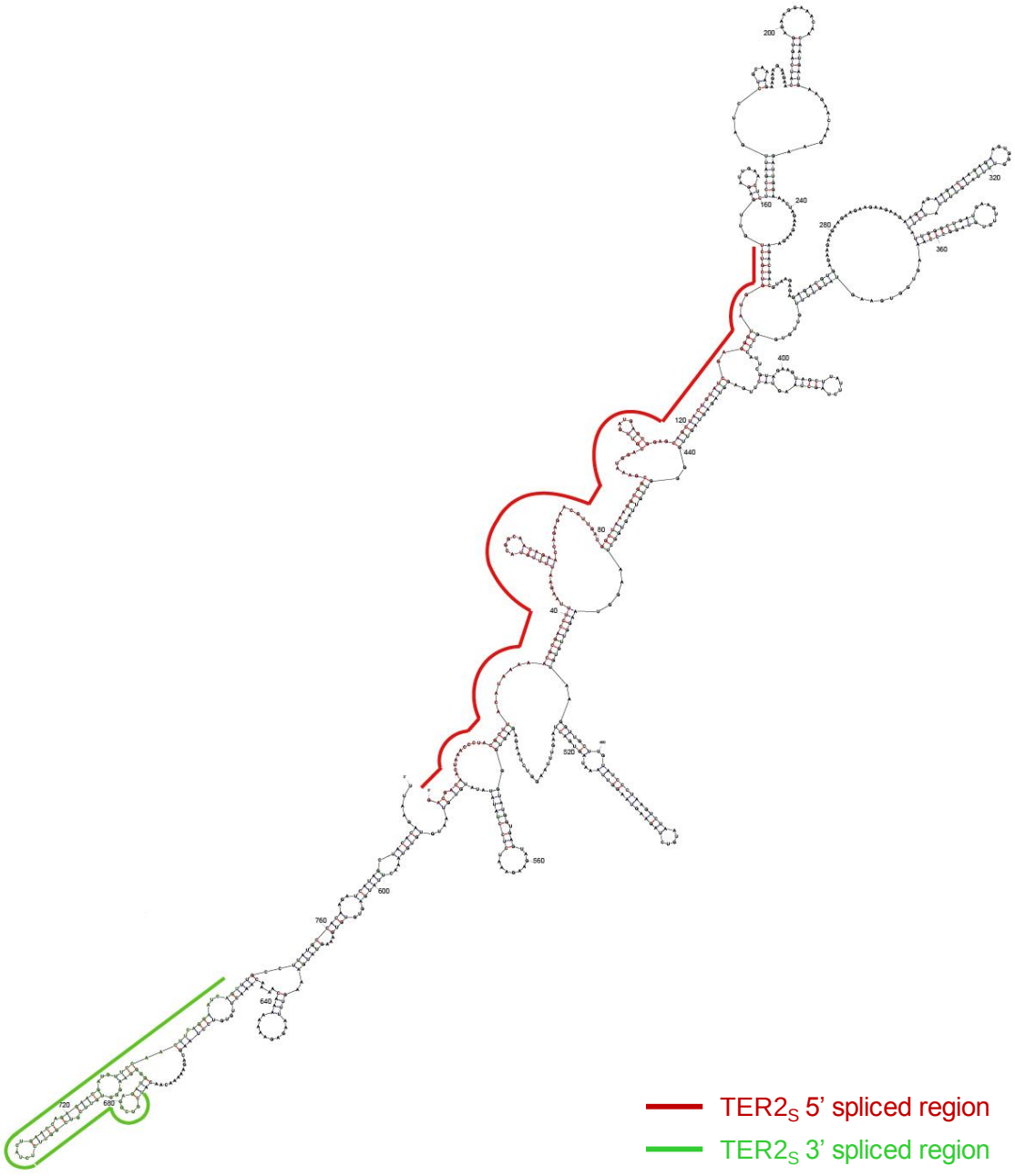


Figure 45. Representative structure prediction of the TER2 transcript. The secondary structure of TER2 was predicted using Mfold software. Red, $TER2_S$ 5' splice region. Green, $TER2_S$ 3' spliced region.

$TER2_S$ is produced in vitro

Unexpectedly, we detected an RNA corresponding to $TER2_S$ in a telomerase reconstitution reaction programmed with TERT and TER2. *In vitro* reconstitution reactions were assembled with T7-tagged TERT and varying concentrations of either TER1 or TER2 (0.1–50 nM). After IP, telomerase activity was monitored by TRAP and RT-PCR was performed to detect RNA using primers specific to conserved region in TER (Fig. 46).

Telomerase activity was retrieved from all reactions, except those containing either TER or TERT alone. A lower level of telomerase activity was observed in reactions assembled with higher concentrations of TER2 (>10 nM) (Fig. 46). Notably, as the amount of TER2 increased, a product of ~200 nt was observed in addition to the ~750 nt band corresponding to TER2. Cloning and sequencing of the smaller product confirmed that it corresponds to $TER2_S$. This finding indicates an inverse relationship between *in vitro* assembled active telomerase RNPs and the processing of TER2 into $TER2_S$ (Fig. 45) and suggests that the TER2 RNP extends the substrate with more efficiency than the TER2 RNP. Altogether, our results indicate that TER2 processing can be reconstituted *in vitro*. They also indicate that the elements required for TER2 processing are contained in our *in vitro* telomerase activity reconstitution system.

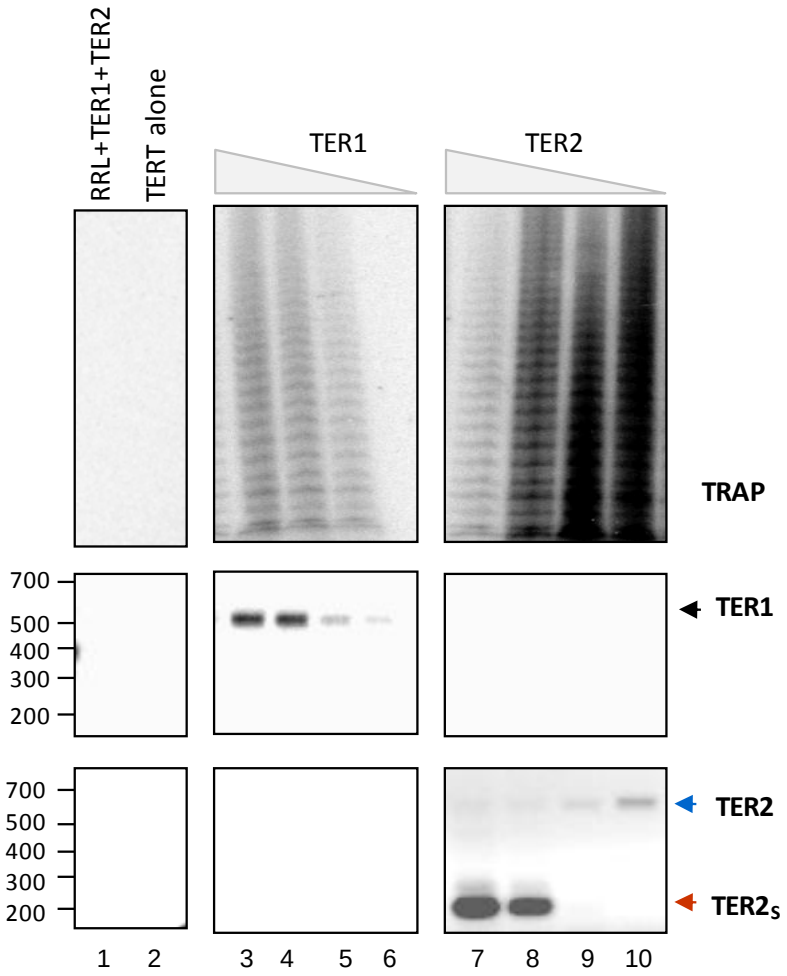


Fig 46. $TER2_S$ is produced *in vitro*. T7-immunoprecipitates of telomerase *in vitro* reconstitution reactions assembled with either TER1 or TER2. Top, Telomerase activity assay. Middle, RT-PCR with primers specific for TER1. RT-PCR with primers specific for the conserved region in TER2. Lane 1, 50nM TER1 and TER2 were added to Rabbit Reticulocyte Lysate (RRL). Lane 2, RRL programmed with TERT only. Lanes 3-6, *In vitro* reconstitution reactions assembled with decreasing concentrations of TER1 (50nM, 10nM, 1nM, 0.1nM). Lanes 7-10, *In vitro* reconstitution reactions assembled with decreasing concentrations of TER2 (50nM, 10nM, 1nM, 0.1nM). $TER2_S$ is detected in reconstitution reactions assembled with >10nM TER2 (Lanes 7-8). Lower telomerase activity levels were detected in these reactions (Lanes 7-8).

The mechanism of TER2 processing into $TER2_S$ does not involve proteins

In vitro reconstitution reactions contain Rabbit Reticulocyte Lysate, buffer, a plasmid encoding T7-tagged TERT and TER transcripts. To further investigate the mechanism of TER2 processing, we modified the *in vitro* reconstitution protocol. First, since $TER2_S$ was found to associate with TERT in reconstitution reactions assembled with TERT and TER2, we asked whether the mechanism of $TER2_S$ production was dependent on TERT. We followed the *in vitro* reconstitution protocol, omitting the IP step, in order to detect TERs independently of their TERT association status. We assembled *in vitro* reconstitution reactions in the presence or absence of TERT and monitored the RNAs by northern blotting. TERT expression was monitored by the incorporation of [^{35}S]-methionine (data not shown). Notably, $TER2_S$ was detected in reactions with or without TERT (Fig. 47A), indicating that $TER2_S$ production occurs independently of TERT.

We next examined the contribution of the RRL in TER2 processing. RRL contains the machinery required for mammalian translation, including RNA-binding proteins as well as RNA helicase activities that may serve a chaperone function and aid in folding TER2 into a processing-competent state. Strikingly, we detected $TER2_S$ in reactions lacking RRL (Fig. 47B). Since the transcribed RNA is purified by denaturing PAGE prior to its addition to the reaction, there is no other protein component present. These findings suggest that TER2 processing may occur autocatalytically.

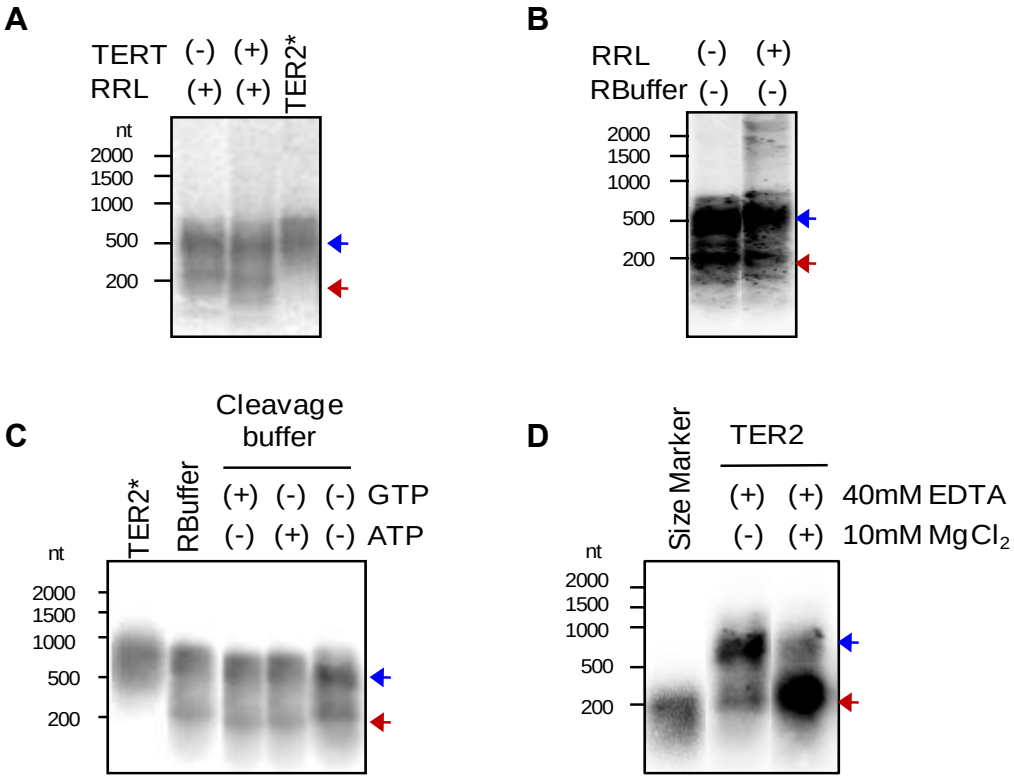


Figure 47. TER2 processing is autocatalytic. Northern blots with *in vitro* transcribed TER2 and a probe specific for the TER2 conserved region. All the incubations were performed at 30ºC. TER2* was defrosted and heated right before loading and was not preincubated. (A) TER2 was incubated with RRL expressed TERT or with RRL alone. (B) TER2 was incubated with RRL or with RRL commercial buffer in the absence of lysate (RBuffer). (C) TER2 was incubated with RRL commercial buffer alone or with cleavage buffer containing (10mM $MgCl_2$ and 50mM NaCl) in the presence or absence of 200µM GTP or ATP. (D) TER2 was incubated in cleavage buffer containing 40mM EDTA in the presence or absence of $MgCl_2$.

The *in vitro* transcription/translation contains nucleoside triphosphates required for transcription. We tested whether TER2 processing required a guanosine nucleophile like group I introns. TER2 was incubated with the RRL buffer, or a cleavage buffer including or lacking GTP (200μM). We observed the appearance of $TER2_S$ under all conditions, except when TER2 was heated right before loading (Fig. 47C). This observation indicates that TER2 processing does not require a guanosine or adenosine cofactor, arguing that TER2 is not processed by a group I intron mechanism.

Finally, since the cleavage buffer contained 10 mM $MgCl_2$, we tested if TER2 processing involved $MgCl_2$ as a co-factor (Fig. 47D). Accordingly, we incubated TER2 with cleavage buffer in the presence or absence of 10 mM $MgCl_2$. Also, 40 mM EDTA was added to the reactions to ensure complete Mg^{2+} depletion. Only a slight decrease in $TER2_S$ production was observed in reactions lacking $MgCl_2$ thus TER2 self-splicing appears to proceed without exogenous Mg^{2+}, in contrast to group I and group II introns.

Discussion

TER2 processing is autocatalytic

Here we present an initial investigation into the biogenesis of $TER2_S$, an RNA that is derived from the two TER2 conserved regions. Because TER2 processing sites do not conform to consensus plant mRNA splice sites, we considered the possibility that TER2 splicing was autocatalytic. In support of this hypothesis, we find that processing of TER2 into $TER2_S$ occurs *in vitro* in reactions devoid of protein.

Splicing of TER2 does not require magnesium, as a vast excess of EDTA does not completely inhibit processing *in vitro*, in contrast to the magnesium-dependent self-splicing group I and group II introns. Additionally, TER2 splicing does not require a guanosine cofactor and preliminary sequence analysis has not revealed features reminiscent of group II introns (data not shown). Further *in vitro* experiments aimed at revealing details of TER2 splicing are needed, as well as detailed sequence and secondary structural analyses. However, the fact that $TER2_S$ formation does not depend on Mg^{2+} suggests that the self-splicing motif of TER2 may be more closely related to the Hepatitis Delta Virus or hairpin ribozymes, metal-independent self-splicing RNAs derived from plant viruses that function in the processing of viral RNA replication intermediates. Furthermore, our data raises the interesting possibility that the functional diversification of *Arabidopsis* TER genes may have been shaped by the invasion of viral sequences into the TER2 locus.

A role for TER2$_S$ in telomerase regulation

The biological function of TER2$_S$ in remains to be elucidated, but our data suggests that processing of TER2 may counteract its negative regulatory role. Since TER2 has higher affinity for TERT than TER1, increasing the concentration of TER2$_S$ would release of TER2-bound TERT, freeing the RT subunit to associate with the more abundant TER1. TER2$_S$ can assemble into a functional telomerase RNP *in vitro*, but the affinity of TERT for TER2$_S$ is low (Kd ~1 μM), indicating that processing of TER2 decreases the strength of its interaction with TERT by 50-fold (compare with 20 nM), this allows the exchange of TERT into a more stable complex with TER1 (Kd 200 nM).

TER2$_S$ associates with the TER2-binding protein POT1b, a negative regulator of telomerase (Chapter III). *pot1b* null plants show increased telomerase activity, at levels similar to those observed in *ter2* plants. This result is consistent with a role for POT1b in the negative regulation of telomerase. Interestingly, neither TER2 nor TER2$_S$ are detected in *pot1b* null mutants, which suggest a role for POT1b in TER2 stability. Thus, TER2s may function as a "sink" for POT1b (or other negative regulators), sequestering the factors needed to stabilize TER2. Increased TER2 splicing, (or increased association of TER2$_S$ with POT1b), would lead to reduced association of POT1b with TER2, causing the destabilization of TER2 RNP and release of TERT.

Further experiments are needed to understand the role(s) of TER2$_S$ *in vivo*. These studies should be combined with an examination of the differences

in the catalytic activities of the TER1-, TER2-, and $TER2_S$-TERT RNPs, and their differential association with telomerase accessory factors. These experiments will shed light into this novel form of telomerase regulation based on alternative RNP subunit composition.

A more detailed understanding of TER2 self-splicing will clarify if this RNA is functionally related to either the Hepatitis Delta Virus or hairpin ribozymes or if it constitutes a brand-new class of self-splicing RNA. Either way the findings will have a profound impact in our understanding of telomerase RNA function and evolution.

Materials and Methods

Northern blotting

Northern blotting was performed with total RNA extracted from *Arabidopsis* suspension culture. Briefly, RNA was denatured for 2 min at 95ºC in formamide loading buffer containing 5M urea and resolved in a 4% acrylamide gel under denaturing conditions. T7 *in vitro* transcribed TER1 was used as control. Then, RNA was transferred to a Hybond-N+® membrane (Amersham) for 10h. After 1h pre-hybridization the membrane was hybridized for 12h at 40ºC with a pool 5'-end 32P-ATP labeled oligonucleotides complementary to TER1. After washes the membrane was exposed to a phosphor screen.

In vitro *telomerase reconstitution*

A TERT-pET28a plasmid with an N-terminal T7 tag was used for telomerase reconstitution experiments. Reactions were assembled with 100ng of TERT-pET28a plasmid and 0.5pmol or 0.1pmol of gel purified DNA template encoding TER1 or TER2 respectively, driven by a T7 promoter, in a mix containing Rabbit Reticulocyte Lysate (RRL) (Promega), amino acids, RNase inhibitors, and T7 RNA polymerase. Reactions were incubated for 90 min at 30°C. T7 agarose beads (Novagen) were blocked with buffer W-100 (20mM TrisOAc [pH 7.5], 10% glycerol, 1mM EDTA, 5mM $MgCl_2$, 0.2M NaCl, 1% NP-40, 0.5mM sodium deoxycholate, and 100mM potassium glutamate) containing 0.5mg/ml BSA, 0.5mg/ml lysozyme, 0.05mg/ml glycogen, 1mM DTT and 1µg/ml yeast tRNA. The reconstitution reaction was mixed with the beads to a 200µl final volume and incubated for 2 h at 4°C with rotation. Beads were washed 6X with 800 µl of W-400 buffer (W-100 containing 400 mM potassium glutamate) and 3X with 800µl of TMG buffer (10mM TrisOAc [pH 7.5], 1mM $MgCl_2$, and 10% glycerol). After the final wash, beads were resuspended in 30µl of TMG. 2µl of beads were used for TRAP assays as previously described (130, 242).

CHAPTER VI

CONCLUSIONS AND FUTURE DIRECTIONS

An essential task for an organism is to maintain genome stability and integrity and thus guarantee its successful function and propagation. In eukaryotes, telomeres and telomerase are key elements required in this enterprise. The 2009 Nobel Prize in Physiology or Medicine recognized the pioneering work of Elizabeth Blackburn, Carol Greider, and Jack Szostak in defining the DNA composition, de novo synthesis and maintenance of telomeres, as well as the detrimental consequences of their loss.

The importance of telomeres in humans is evidenced by the involvement of telomere shortening in aspects of aging and in the chromosomal instability associated with human cancer. In essence, if the telomeres shorten, cells age. In contrast, if telomere length is maintained by telomerase, cellular senescence is delayed. Telomerase activity must to be carefully regulated. Inappropriate telomerase expression results in the loss of control of cell proliferation and is one of the rate-limiting factors in carcinogenesis (365-369). On the other hand, not enough telomerase limits cell proliferation, resulting in loss of stem cell viability by telomere dysfunction and in devastating genetic disorders such as dyskeratosis congenita and aplastic anemia (370). Consequently, a major effort in the telomere field has focused on understanding telomerase composition, biogenesis and regulation at the chromosome terminus.

Although interesting differences exist between organisms, the protection of chromosome ends by telomeres and telomerase is highly conserved in eukaryotes. The flowering plant *Arabidopsis thaliana*, in particular, is an excellent model organism to study telomere metabolism, since it is a multicellular organism with complex development and excellent genetic tools. In addition it is highly tolerant to genome instability, allowing the study of telomere-related genes whose mutations are lethal in mammals. My dissertation focuses in the characterization of the *Arabidopsis thaliana* telomerase RNP. This chapter contains conclusions from this work, and proposed future directions.

The Duplication of Telomerase Components in *Arabidopsis* Results in Functional Diversification

In most organisms studied thus far, the core components of telomerase are encoded in the genome as single copy genes (14, 147, 149, 171, 231, 244, 320, 371). The only two exceptions are found in the ciliated protozoa *Euplotes crasus* and in *Arabidopsis thaliana. Euplotes* encodes a single TER and three TERT isoforms that assemble into different active sub-complexes *in vivo*. These different telomerase RNPs are required in different stages of telomere metabolism and promote a shift in enzyme specificity. Specifically, the enzymatic transitions from de novo synthesis of telomere tracts onto broken chromosome ends to the maintenance of the length of previously synthesized telomere tracts (227).

In *Arabidopsis*, TERT is a single copy gene (143), and three TER isoforms are present: TER1 and TER2 which are distinct genes encoded in different chromosomes and $TER2_S$, a product of TER2 processing (Chapters III and V). Although all three TER isoforms can assemble with TERT to direct telomere synthesis *in vitro*, these RNA subunits play different roles *in vivo*.

TER1 has a conserved function in telomere metabolism

AtTERT uses TER1 as the template for telomere maintenance *in vivo*. Thus, TER1 is a canonical telomerase template with a function similar to TER subunits from other organisms (14, 231, 319-320, 372). Depletion of TER1 results in progressive loss of high molecular weight telomere tracts and decreased telomerase activity. A similar phenotype is observed in plants depleted of both TER1 and TER2. Additionally, mutant telomere repeats are incorporated into the chromosome ends of plants over-expressing $TER1_{CC}$, a TER1 transcript bearing a mutation in the template domain (Chapter II). These data are in agreement with previous observations made in similar experiments with budding yeast (158), *Tetrahymena* (16) and human (373) TER subunits.

A conserved domain in TER1 and TER2 is sufficient to reconstitute telomerase activity in vitro

The *Arabidopsis* TERs share a highly conserved region of 220nt that contains the telomere templating domain and all of the elements required for TERT association. A 150nt “mini-T” within the 5’ (R1) conserved region of TER1 and TER2 and the processed form of TER2, $TER2_S$, is sufficient to reconstitute telomerase activity *in vitro*. This size correlates with *Tetrahymena* TER (14) and is slightly smaller than the minimal TERs from humans (210nt) (293) and from *S. cerevisiae* (170nt) (294). Outside the conserved domain TER1 and TER2 share virtually no sequence identity and cannot be accurately aligned, implying that unique sequences within these RNAs specify their distinct functions and interaction partners *in vivo*.

TER2 functions as a novel negative regulator of telomerase activity

In contrast to TER1, TER2 has an unconventional function that has not been reported for any TER. It works as a negative regulator of telomerase activity. Notably, TER2 depletion results in a dosage-dependent increase in enzyme activity (Chapter III), although no change in telomere length regulation is observed in these plants. Additionally, TER2 over-expression results in decreased telomerase activity and in loss of telomere length heterogeneity as well as in shortened telomeres, similar to those from AtTERT mutants (143). These observations are consistent with a negative regulatory role for TER2.

Moreover, since TER1 levels in the TER2 over-expression mutants are similar to those in wild-type plants, TER2 regulates TER1 function.

TER2 can assemble with TERT into a functional RNP *in vitro*. Strikingly, we find that the affinity of TERT for TER2 is ~9-fold higher than for TER1. Although TER2 levels are ~20-fold lower than TER1 in cell culture, the amount of TER1 associated to TERT is only two-fold higher than TER2 (Chapter II). Thus, TER2 is over-represented in TERT RNP complexes *in vivo*. TER2 is not able to support telomere repeat maintenance *in vivo*. We find a very low number of modified telomere repeats in chromosome ends in plants over-expressing $TER2_{RSA}$, a TER2 transcript bearing a mutation in the templating domain (Chapter III). Thus, the cellular levels of TER2 and the differences in affinity for TER1 and TER2 may be important for the mechanism of TER2 regulation. We conclude that TER2 plays a complex role in telomerase regulation.

Duplication and diversification of a core telomerase RNP subunit is not without precedent. There are reports of variant RNA subunits within the spliceosome and the ribosome. Humans and *Drosophila* encode multiple isoforms of U5 snRNA, which assemble into different RNP particles, and display different expression profiles *in vivo* (374-375). Similarly, *Plasmodium* harbors two different small subunit rRNAs (376), whose differential expression and processing lead to stage-specific ribosome populations thought to regulate the translation of blood-stage or mosquito mRNAs (377). These findings support

our conclusions that duplication of TER can result in a novel regulatory mechanism for telomerase.

Arabidopsis Assembles Different Telomerase RNPs That Have Opposing Roles in Telomere Metabolism

Our data demonstrate that *Arabidopsis* harbors at least two variant telomerase RNPs with distinct RNA and protein subunit composition (Chapter III). These complexes share the catalytic subunit TERT and the RNP maturation factor dyskerin, but assemble with either TER1 or TER2 and the accompanying TER1- or TER2-specific accessory proteins. Both TER1 and TER2 provide a template for telomerase reverse transcription *in vivo* and both can direct the incorporation of telomere repeats onto chromosome ends. Notably, however, the incorporation mediated by TER2 RNP is very inefficient compared to the TER1 RNP (Chapters II and III). Thus TER2 may associate with chromosome termini, but it is not able to efficiently incorporate telomeric repeats (Fig. 48).

We find that POT1a associates with the TER1 RNP responsible for maintaining telomere tracts *in vivo*. Telomerase activity is reduced when either TER1 or POT1a is depleted (Chapter II and (130, 313)). Thus, the POT1a-TER1 RNP association is central for telomerase function.

In contrast, POT1b and KU associate with the TER2 RNP, a negative regulator of telomerase. POT1b and KU appear to reinforce TER2-mediated

negative regulation. *pot1b-1* mutants exhibit a three-fold increase in telomerase activity, similar to the increase observed in *ter2-1* mutants (Chapter III).

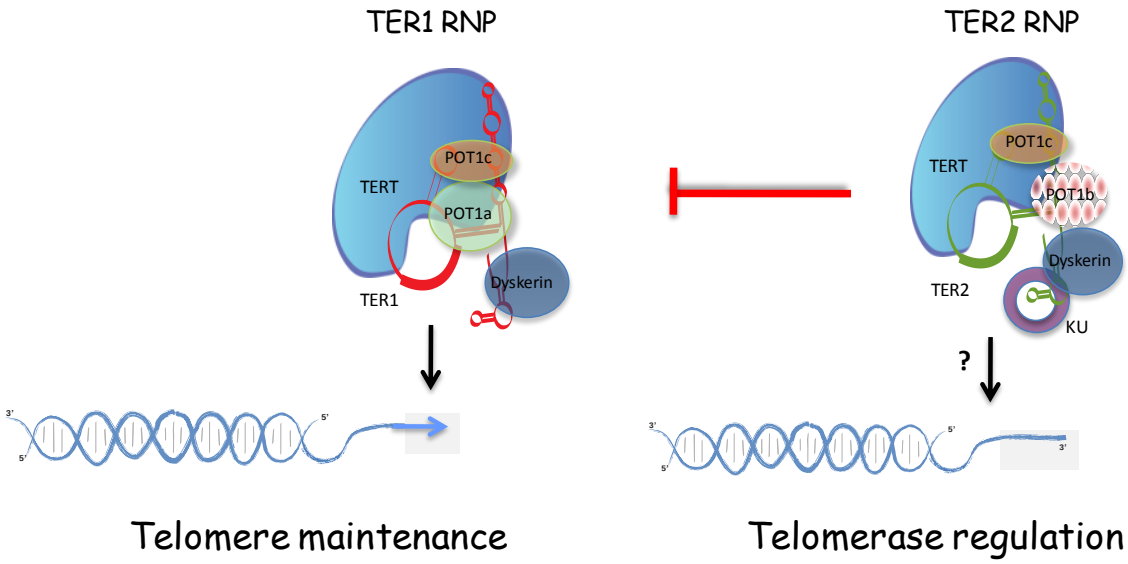


Figure 48. *Arabidopsis* encodes two telomerase RNPs with distinct RNA and protein composition and different roles *in vivo.*

In addition, the TER2-interacting factor KU is a strong negative regulator of telomerase-mediated telomere elongation in *Arabidopsis*. *ku70* plants show an increase in telomere length (296) but telomerase enzyme activity levels are not increased (Nelson A. and D. Shippen unpublished data). These findings suggest that KU regulates telomere elongation at the chromosome end rather than enzyme activity. Previous studies reveal that KU physically associates with the TERs from *S. cerevisiae* (182) and humans (312), but not from *S. pombe* (319). Our data support the conclusion that the KU-TER interaction is dispensable for telomerase function.

How does KU negatively regulate telomere function *in vivo*? One intriguing possibility is that KU is involved in recruitment of telomerase to the telomere in a cell cycle dependent manner. In budding yeast KU brings telomerase to the telomere in G1, but the enzyme does not promote telomere elongation until late S/G2 (252). Because KU is associated with the TER2 RNP in *Arabidopsis*, recruitment of the negative regulatory complex in G1, could interfere with TER1 RNP recruitment and telomere extension (Fig. 49). It is striking that all of the TER2-specific RNP components (POT1b, TER2 and KU) identified thus far serve as a negative regulators for telomerase function.

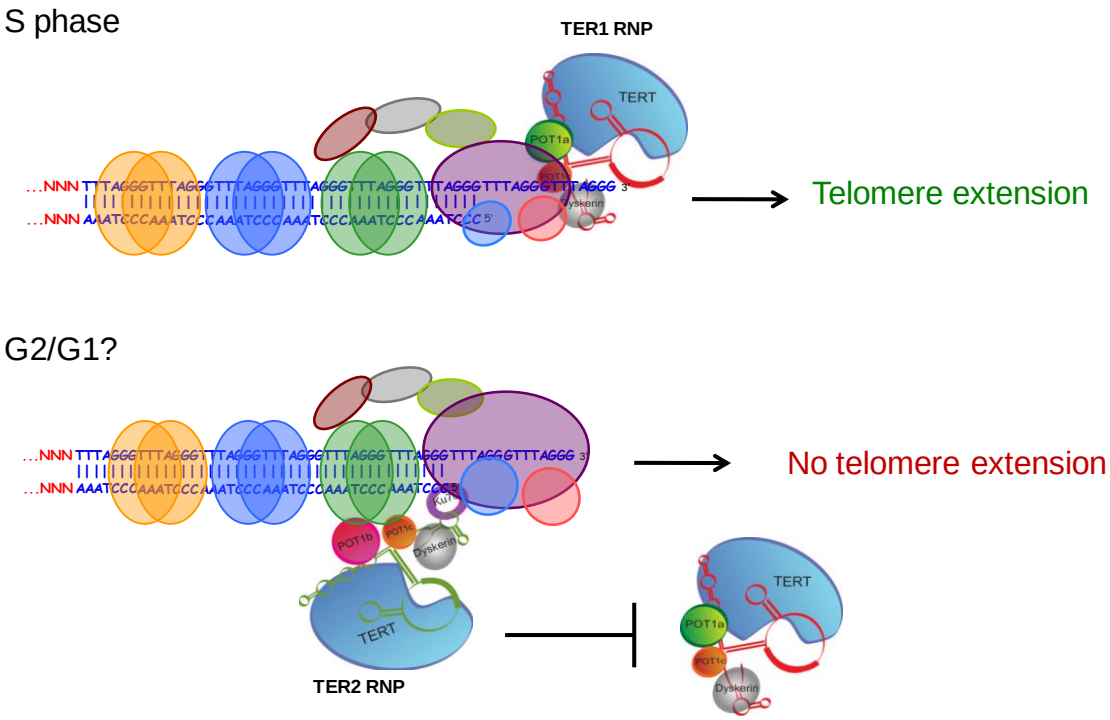


Figure 49. Proposed model for the regulation of telomerase function at the chromosome end by the TER1 and TER2 RNPs. See text.

AtPOT1 Paralogs Encode TER Binding Proteins That Play Distinct Functions in Telomere Biology

Fission yeast and humans encode a single copy POT1 protein (69), while other organisms, like mice (118-119), *Tetrahymena* (326), *Euplotes* (325) and *C. elegans* (40) contain more than one POT1. POT1 proteins from fission yeast (28, 334), mammals (28, 118) and some plants (132), share a ss telomeric DNA binding activity and play crucial roles in distinguishing the chromosome terminus from ds breaks.

Arabidopsis thaliana is unusual because it encodes three POT1 proteins: POT1a, POT1b and POT1c (129) and Nelson A. and Shippen DE unpublished data). Furthermore, unlike POT1 orthologs in other eukaryotes, binding experiments have failed to demonstrate DNA binding activity for AtPOT1 family members (131). Instead, AtPOT1 proteins have evolved to bind TER (Chapters II, III and IV), and distribute to the two TER isoforms. POT1a associates with TER1, while POT1b associates with TER2. POT1c, on the other hand, associates with both TER1 and TER2.

Interestingly, the dissociation constant for the POT1a-TER1 interaction (Kd=~200 nM) is similar to that for the hPOT1-ssDNA (Kd=~100 nM) (120-121). In addition, deletional mutagenesis experiments showed that, like hPOT1 (29, 122), nucleic acid recognition by POT1a is mediated through the OB-folds, with the first N-terminal OB-fold being sufficient for TER1 recognition (Chapters II and IV).

We find that mutations in a highly conserved phenylalanine residue (Phe65) in the first OB-fold results in decreased TER1 specific binding. This finding correlates with observations for hPOT1-ssDNA interaction, in which the corresponding phenylalanine (Phe62) is a major contact with its DNA substrate (29, 122) and discriminantes between RNA and DNA recognition (272). Structural modeling analysis showed that AtPOT1a OB-fold 1 models onto the crystal structure of hPOT1 (Croy, J. and Wuttke, D., University of Colorado and Chapter IV). Thus, the switch from ss telomeric DNA to TER recognition by AtPOT1a does not involve appear to involve significant structural changes of its nucleic acid binding pocket.

AtPOT1a is a positive regulator of telomerase. It associates with active telomerase RNP *in vivo* through a direct interaction with TER1 (Chapter II) and it acts in the same genetic pathway as telomerase (130). Our findings indicate that like POT1a, POT1b and POT1c associate with telomerase, but in this case with the negative regulatory RNP. We find that POT1b associates with both TER2, a negative regulator of telomerase activity and $TER2_S$ *in vivo* (Chapter III).

POT1c, which contains a single OB-fold, associates with both TER1 and TER2, suggesting that TER specificity may be conferred by interactions through the second OB-fold, since this domain is absent in POT1c. It is possible that POT1c competes with POT1a and POT1b binding sites and alters enzyme activity or recruitment of the RNP complex to the telomeres. Although additional experiments are required to determine POT1c function *in vivo*, knockdown of

POT1c by RNAi results in increased telomerase activity levels, supporting a role for this protein in telomerase regulation (Nelson A. and Shippen D. unpublished data).

Altogether, our results indicate that POT1 proteins are novel telomerase accessory factors that play different functions in telomerase regulation. They also indicate that indicate that duplication and diversification of *Arabidopsis* telomerase may be the end result of the co-evolution of the TER and POT1 RNP components.

Future Directions

POT1a, an Est1-like recruitment factor?

Several lines of evidence support the hypothesis that AtPOT1a functions in a manner analogous to the yeast protein Est1, a telomerase-associated protein that bind the telomerase RNA subunit and positively regulates telomere extension (181), Chapters II and IV). First, an Ever-Shorter-Telomere phenotype, characterized by the progressive shortening of telomeres from one generation to the next due to the lack of telomerase, is observed in both *pot1a* and *est1* null mutants (18, 129). Second, telomerase activity co-immunoprecipitates with POT1a and Est1, indicating a physical interaction among these components and active telomerase (130, 164). Third, the stability of both proteins appears to be regulated during the cell cycle (378) and Armstrong, S., personal communication). Fourth, POT1a interacts with CTC1,

the functional homolog of the G-overhang-binding protein Cdc13 (Song. X and Shippen D, unpublished data), which mediates Est1-dependent telomerase recruitment to the chromosome end in yeast (379). Finally, the architecture of the POT1a binding site in TER1 (Chapter IV) resembles the bulged stem characteristic of the Est1 binding site in yeast TLC1 (181).

Based on these finding I hypothesize that, like yeast EST1, POT1a functions as a recruitment factor that brings telomerase to the chromosome end through its interaction with CTC1. This hypothesis can be tested by: 1) mapping the regions of POT1a and CTC1 that mediate their interaction *in vitro* to identify mutations that disrupt interaction but not their nucleic acid binding properties. These experiments would be followed by expressing POT1a mutant alleles in *pot1a* null plants to identify those residues that affect telomere length but not telomerase activity. Results of these experiments would test the hypothesis that interaction between POT1a and CTC1 is required for the recruitment of telomerase to the chromosome end. A second line of investigation is to establish the pattern of telomerase-CST interactions during the cell cycle by performing Chromatin IP and co-localization experiments for TERT, POT1a and CTC1 as well as telomere and TER FISH during different stages of the cell cycle. A third set of experiments would involve disruption of the POT1a-TER1 interaction and testing by ChIP or co-localization experiments if telomerase is recruited to telomeres or if there are changes in the pattern of association of telomerase with telomeres during the cell cycle.

POT1a as a regulator of telomerase enzymatic properties

Notably, Est1 is not required for telomerase activity (380). Additionally, EST1 binds ss telomeric DNA, although its affinity for DNA is lower than for Tlc1 (381). Similarly, *pot1a* mutants show a 10-fold decrease in telomerase activity (130), indicating that POT1a is required for telomerase activity. This raises the possibility that POT1a may play additional roles in the regulation of telomerase.

The human sheltering components, POT1-TTP1, alternate between two main functions in telomere biology: protecting the chromosome end and serving as a processivity factor for telomerase during telomere extension (120), possibly through a direct interaction of TPP1 with TERT. A plausible hypothesis is that, like human TPP1, POT1a alone or in complex with an unidentified protein may affect telomerase enzymatic properties like processivity or substrate recognition.

This hypothesis can be tested by identifying proteins that interact with POT1a using yeast-two hybrid and by mass spectrometry of affinity-purified POT1a complexes. Potential ss telomeric DNA binding proteins or CTC1-interacting proteins could be found that influence POT1a function. Another important goal is to develop a direct activity assay for *Arabidopsis* telomerase that does not rely on PCR. A direct assay would allow us to examine enzymatic properties like processivity, since the PCR-based TRAP assay is not suitable for these goals. With a direct assay, one could examine the effect of POT1a on telomerase processivity using the *in vitro* reconstitution system. Additionally, reactions could be carried out to examine the effect of POT1a on CTC1 binding.

In *Tetrahymena*, the p65 protein mutually stabilizes TER and TERT in enzymatically-active telomerase conformations (382). Thus, another possibility is that POT1a is required for enzyme activity because it enhances active an active RNP conformation. This possibility can be tested by: determining the effects of POT1a on the affinity of the TERT-TER1 interaction *in vitro*. The dissociation rate constant of the TERT-TER1 interaction could also be measured by FRET. If my hypothesis is correct, I expect to see an increase in affinity of the TERT-TER1 interaction or the formation of a more stable TERT-TER1 complex in the presence of POT1a.

Dissecting the mechanism of negative regulation by the TER2 RNP

Human core telomerase is active as a dimer *in vivo* (356). Preliminary results from glycerol gradient fractionation suggest *Arabidopsis* telomerase forms oligomers *in vitro*. To study the role of oligomerization in telomerase regulation, one could determine the oligomerization state of active recombinant *Arabidopsis* telomerase complex by gel filtration chromatography as has been done for *Tetrahymena* (155) and humans (356). I expect to observe distinct telomerase complexes, with a fraction of telomerase in TER1- or TER2-monomers, in addition to homo- and hetero-dimeric telomerase RNP complexes that differ in activity in response to the oligomeric state. Alternatively one could use a combination of RNA-tagging and template mutations to determine if

Arabidopsis telomerase forms hetero-oligomers and whether both TER1 and TER2 templates are used for extension.

TER2 processing into TER2s

TER2 is processed into $TER2_S$ (Chapter III). I hypothesize that processing of TER2 counteracts its negative regulatory role. TER2 splicing into $TER2_S$ may result in disassembly of the TER2 RNP, boosting the assembly of free TERT into TER1 RNPs, since the interaction of $TER2_S$ with TERT is weak. Thus, processing of TER2 would lead to dissociation of TER2 RNPs, allowing the exchange of TERT into a more stable complex with TER1 (Fig. 50). This model can be tested *in vitro* by evaluating the exchange of TERT from assembled TER2 RNP in the presence of TER1 and $TER2_S$.

$TER2_S$ associates with the TER2-binding protein POT1b, a negative regulator of telomerase (Chapter III). The fact that neither TER2 nor $TER2_S$ are detected in *pot1b* null mutants, suggests a role for POT1b in TER2 stability. It is possible that TER2s functions by destabilizing the TER2 transcript through sequestration of POT1b. One could test the role of POT1b in TER2 stability *in vivo*, by ectopically expressing POT1b under the control of an inducible promoter in *pot1b* mutants. I would expect to find TER2 and $TER2_S$ when POT1b is expressed.

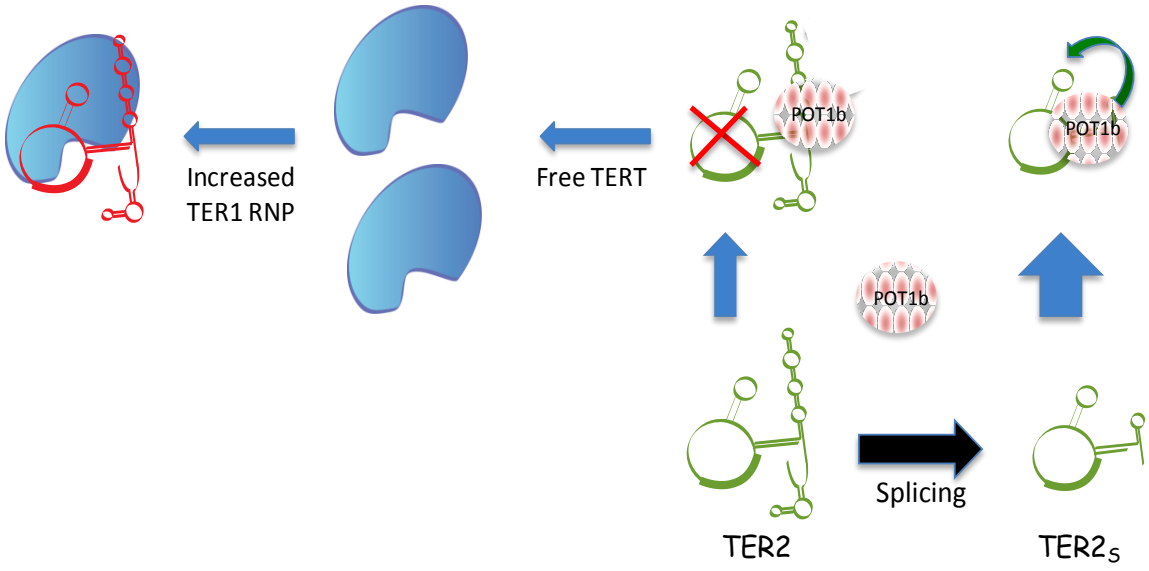


Figure 50. Proposed model for how TER2 may influence TER1 function *in vivo*. See text.

Finally, TER2 processing occurs in the absence of proteins, guanosine and Mg^{2+} cofactors (Chapter V), which suggest that the mechanism of TER2 self-splicing is not similar to those employed by group I and II introns. An alternative is that TER2s is produced by a mechanism similar to the Hepatitis Delta Virus or hairpin ribozymes, which does not depend on any metal cofactor for catalysis. This can be tested by secondary structure and FRET analysis aimed to identify domains in close proximity that can result in cleavage.

In conclusion, this study indicates that duplication and diversification of *Arabidopsis* telomerase RNP may be the end result of the co-evolution of the TER and POT1 RNP components. Additionally, the discovery of a novel negative regulatory mechanism for plant telomerase argues that additional modes of restraining telomerase remain to be elucidated in mammals where misregulation of this enzyme can lead to carcinogenesis.

REFERENCES

1. Weissmann A (1889) *Essays upon heredity and kindred biological problems.* (Clarendon Press., Oxford).

2. Carrel A & Ebeling AH (1921) The multiplication of fibroblasts *in vitro*. *J Exp Med* 34(4):317-337.

3. Hayflick L (1961) The establishment of a line (WISH) of human amnion cells in continuous cultivation. *Exp Cell Res* 23:14-20.

4. Hayflick L (1965) The limited *in vitro* lifetime of human diploid cell strains. *Exp Cell Res* 37:614-636.

5. Muller HJ (1938) The remaking of chromosomes. *Collect. Net, Woods Hole* 13:181-195.

6. McClintock B (1932) A Correlation of ring-shaped chromosomes with variegation in *Zea Mays*. *Proc Natl Acad Sci USA* 18(12):677-681.

7. McClintock B (1939) The behavior in successive nuclear divisions of a chromosome broken at meiosis. *Proc. Natl. Acad. Sci USA.* 25:405-416.

8. Watson JD (1972) Origin of concatemeric T7 DNA. *Nat New Biol* 239(94):197-201.

9. Olovnikov AM (1973) A theory of marginotomy. *J. Theor. Biol.* 41:181-190.

10. Blackburn EH & Gall JG (1978) A tandemly repeated sequence at the termini of the extrachromosomal ribosomal RNA genes in *Tetrahymena*. *J. Mol. Biol.* 120:33-53.

11. Szostak JW (1982) Cloning yeast telomeres on linear plasmid vectors. *Cell* 29:245-255.

12. Greider CW & Blackburn EH (1985) Identification of a specific telomere terminal transferase activity in *Tetrahymena* extracts. *Cell* 43:405-413.

13. Greider EW & Blackburn EH (1987) The telomere terminal transferase of *Tetrahymena* is a ribonucleoprotein enzyme with two kinds of primer specificity. *Cell* 51:887-898.

14. Greider CW & Blackburn EH (1989) A telomeric sequence in the RNA of *Tetrahymena* telomerase required for telomere repeat syntheses. *Nature* 337:331-337.

15. Shippen-Lentz D & Blackburn EH (1990) Functional evidence for an RNA template in telomerase. *Science* 247:546-552.

16. Yu G-L, Bradley JD, Attardi LD, & Blackburn EH (1990) *In vivo* alteration of telomere sequences and senescence caused by mutated *Tetrahymena* telomerase RNAs. *Nature* 344:126-132.

17. Lundblad V & Szostak JW (1989) A mutant with a defect in telomere elongation leads to senescence in yeast. *Cell* 57:633-643.

18. Lendvay TS, Morris DK, Sah J, Balasubramanian B, & Lundblad V (1996) Senescence mutants of *Saccharomyces cerevisiae* with a defect in telomere replication identify three additional EST genes. *Genetics* 144:1399-1412.

19. Burr B, Burr FA, Matz EC, & Romero-Severson J (1992) Pinning down loose ends: mapping telomeres and factors affecting their length. *Plant Cell* 4:953-960.

20. Higashiyama T, Maki S, & Yamada T (1995) Molecular organization of Chlorella vulgaris chromosome 1:presence of telomeric repeats that are conserved in higher plants. *Mol. Gen. Genet.* 246:29-36.

21. Kilian A, Stiff C, & Kleinhofs A (1995) Barley telomeres shorten during differentiation but grow in callus culture. *Proc. Natl. Acad. Sci. USA.* 92:9555-9559.

22. Richards ER & Ausubel FM (1988) Isolation of a higher eukaryotic telomere from *Arabidopsis thaliana*. *Cell* 53:127-136.

23. Klobutcher LA, Swanton MT, Donini O, & Prescott DM (1981) All gene-sized DNA molecules in four species of hypotrichs have the same terminal sequence and an unusual 3' terminus. *Proc. Natl. Acad. Sci. USA.* 78:3015-3019.

24. Wellinger RJ, Wolf AJ, & Zakian VA (1993) *Saccharomyces* telomeres acquire single-stranded TG1-3 tails late in S phase. *Cell* 72:51-60.

25. Makarov VL, Hirose Y, & Langmore JP (1997) Long G tails at both ends of human chromosomes suggest a C strand degradation mechanism for telomere shortening. *Cell* 88:657-666.

26. Venkatesan RN & Price C (1998) Telomerase expression in chickens: constitutive activity in somatic tissues and down-regulation in culture. *Proc Natl Acad Sci USA* 95(25):14763-14768.

27. Riha K, Fajkus J, McKnight TD, Vyskot B, & Shippen DE (2000) Analysis of the G-overhang structures on plant telomeres: evidence for two distinct telomere architectures. *Plant J.* 23:1-11.

28. Baumann P & Cech TR (2001) Pot1, the putative telomere end-binding protein in fission yeast. *Science*. 11;292(5519):1171-5.

29. Lei M, Podell ER, Baumann P, & Cech TR (2003) DNA self-recognition in the structure of Pot1 bound to telomeric single-stranded DNA. *Nature* 426(6963):198-203.

30. Loayza D, Parsons H, Donigian J, Hoke K, & de Lange T (2004) DNA binding features of human POT1: a nonamer 5'-TAGGGTTAG-3' minimal binding site, sequence specificity, and internal binding to multimeric sites. *J Biol Chem* 279(13):13241-13248.

31. Dionne I & Wellinger RJ (1998) Processing of telomeric DNA ends requires the passage of a replication fork. *Nucl. Acids Res.* 26:5365-5371.

32. Bonetti D, Martina M, Clerici M, Lucchini G, & Longhese MP (2009) Multiple pathways regulate 3' overhang generation at *S. cerevisiae* telomeres. *Mol Cell* 35(1):70-81.

33. Larrivee M & Wellinger RJ (2006) Telomerase- and capping-independent yeast survivors with alternate telomere states. *Nat Cell Biol* 8(7):741-747.

34. Vodenicharov MD & Wellinger RJ (2006) DNA degradation at unprotected telomeres in yeast is regulated by the CDK1 (Cdc28/Clb) cell-cycle kinase. *Mol Cell* 24(1):127-137.

35. Frank CJ, Hyde M, & Greider CW (2006) Regulation of telomere elongation by the cyclin-dependent kinase CDK1. *Mol Cell* 24(3):423-432.

36. Huffman KE, Levene SD, Tesmer VM, Shay JW, & Wright WE (2000) Telomere shortening is proportional to the size of the G-rich telomeric 3'-overhang. *J Biol Chem* 275(26):19719-19722.

37. Wright WE, Tesmer VM, Huffman KE, Levene SD, & Shay JW (1997) Normal human chromosomes have long G-rich telomeric overhangs at one end. *Genes and Dev.* 11:2801-2809.

38. Wright WE, Tesmer VM, Liao ML, & Shay JW (1999) Normal human telomeres are not late replicating. *Exp. Cell Res.* 251:492-499.

39. Cimino-Reale G, Pascale E, Alvino E, Starace G, & D'Ambrosio E (2003) Long telomeric C-rich 5'-tails in human replicating cells. *J Biol Chem* 278(4):2136-2140.

40. Raices M*,* Verdun RE, Compton SA, Haggblom CI, Griffith JD, *et al.* (2008) *C. elegans* telmomeres contain G-strand and C-strand overhangs that are bound by distinct proteins. *Cell* 132:745-757.

41. Griffith JD*, et al.* (1999) Mammalian telomeres end in a large duplex loop. *Cell* 97:503-514.

42. Stansel RM, de Lange T, & Griffith JD (2001) T-loop assembly *in vitro* involves binding of TRF2 near the 3' overhang. *EMBO J.* 20:5532-5540.

43. Grunstein M (1997) Molecular model for telomeric heterochromatin in yeast. *Curr Opin Cell Biol* 9(3):383-387.

44. Bystricky K, Laroche T, van Houwe G, Blaszczyk M, & Gasser SM (2005) Chromosome looping in yeast: telomere pairing and coordinated movement reflect anchoring efficiency and territorial organization. *J Cell Biol* 168(3):375-387.

45. Munoz-Jordan JL, Cross GAM, de Lange T, & Griffith JD (2001) T-loops at trypanosome telomeres. *EMBO J.* 20:579-588.

46. Murti KG & Prescott DM (2002) Topological organization of DNA molecules in the macronucleus of hypotrichous ciliated protozoa. *Chromosome Res* 10(2):165-173.

47. LeBel C & Wellinger RJ (2005) Telomeres: what's new at your end? *J. Cell Sci.* 118:2785-2788.

48. Fajkus J, Korarik A, Kralovics R, & Bezdek M (1995) Organization of telomeric and subtelomeric chromatin in the higher plant Nicotiana tabacum. *Mol. Gen. Genet.* 247:633-638.

49. Riha K, Heacock ML, & Shippen DE (2006) The role of the nonhomologous end-joining DNA double-strand break repair pathway in telomere biology. *Annu Rev Genet* 40:237-277.

50. McEachern MJ & Blackburn EH (1995) Runaway telomere elongation caused by telomerase RNA gene mutations. *Nature* 376:403-409.

51. Nakamura TM, Morin GB, Chapman KB, Weinrich SL, Andrews WH, *et al.* (1997) Telomerase catalytic subunit homologs from fission yeast and human. *Science* 277:955-959.

52. Oguchi K, Hongtu L, Tamura K, & Takahashi H (1999) Molecular cloning and characterization of AtTERT, a telomerase reverse transcriptase homolog in *Arabidopsis* thaliana. *FEBS Lett.* 457:465-469.

53. Kobryn K & Chaconas G (2001) The circle is broken: telomere resolution in linear replicons. *Curr Opin Microbiol* 4(5):558-564.

54. Biessmann H, *et al.* (1992) Frequent transpositions of Drosophila melanogaster HeT-A transposable elements to receding chromosome ends. *EMBO J* 11(12):4459-4469.

55. Levis RW, Ganesan R, Houtchens K, Tolar LA, & Sheen F-m (1993) Transposons inplace of telomeric repeats at a Drosophila telomere. *Cell* 75:1083-1093.

56. Abad JP, *et al.* (2004) Genomic analysis of Drosophila melanogaster telomeres: full-length copies of HeT-A and TART elements at telomeres. *Mol Biol Evol* 21(9):1613-1619.

57. Pardue M-L (1995) Drosophila telomeres: another way to end it all. *Telomeres*, ed Greider CW. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp 339-370.

58. Lundblad V & Blackburn EH (1993) An alternative pathway for yeast telomere maintenance rescues est1- senescence. *Cell* 73:347-360.

59. Teng SC & Zakian VA (1999) Telomere-telomere recombination is an efficient bypass pathway for telomere maintenance in *Saccharomyces cerevisiae*. *Mol Cell Biol* 19(12):8083-8093.

60. Lundblad V (2002) Telomere maintenance without telomerase. *Oncogene* 21(4):522-531.

61. Cesare AJ & Reddel RR (2010) Alternative lengthening of telomeres: models, mechanisms and implications. *Nat Rev Genet* 11(5):319-330.

62. Perrem K, Colgin LM, Neumann AA, Yeager TR, & Reddel RR (2001) Coexistence of alternative lengthening of telomeres and telomerase in hTERT-transfected GM847 cells. *Mol Cell Biol* 21(12):3862-3875.

63. Gottschling DE & Zakian VA (1986) Telomere proteins: specific recognition and protection of the natural termini of *Oxytricha* macronuclear DNA. *Cell* 47:195-205.

64. Price CM & Cech TR (1987) Telomeric DNA-protein interactions of *Oxytricha* macronuclear DNA. *Genes and Dev.* 1:783-793.

65. De Lange T (2005) Telomere-related genome instability in cancer. *Cold Spring Harb Symp Quant Biol* 70:197-204.

66. Blackburn EH (2001) Switching and signaling at the telomere. *Cell* 106(6):661-673.

67. Gottschling DE & Cech TR (1984) Chromatin structure of the molecular ends of *Oxytricha* macronuclear DNA: phased nucleosome and and a telomeric complex. *Cell* 38:501-510.

68. Garvik B, Carson M, & Hartwell L (1995) Single-stranded DNA arising at telomeres in cdc13 mutants may constitute a specific signal for the RAD9 checkpoint. *Mol. Cell. Biol.* 15:6128-6138.

69. Baumann P & Cech TR (2001) Pot1, the putative telomere end-binding protein in fission yeast and humans. *Science* 292(5519):1171-1175.

70. Theobald DL, Mitton-Fry RM, & Wuttke DS (2003) Nucleic acid recognition by OB-fold proteins. *Annu Rev Biophys Biomol Struct* 32:115-133.

71. Cooper JP, Nimmo ER, Allshire RC, & Cech TR (1997) Regulation of telomere length and function by a myb-domain protein in fission yeast. *Nature* 385:744-747.

72. Shore D & Nasmyth K (1987) Purification and cloning of a DNA binding protein from yeast that binds to both silencer and activator elements. *Cell* 51:721-732.

73. Biluad T, Brun C, Ancelin K, Koering CE, Laroche T*, et al.* (1997) Telomeric localization of TRF2, a novel human telobox protein. *Nat. Genet.* 17:236-239.

74. Broccoli DS, A., Chong L, & Delange T (1997) Human telomeres contain two distinct myb-related proteins, TRF1 and TRF2. *Nature Genetics* 17:231-235.

75. Chong L, van Steensel B, Broccoli D, Erdjument-Bromage H, Hanish J, *et al.* (1995) A human telomeric protein. *Science* 270(5242):1663-1667.

76. Biluad T*,* Koering CE, Binet-Brasselet E, Ancelin K, Pollice A, *et al.* (1996) The telobox, a myb-related telomeric DNA binding motif found in proteins from yeast, plants and human. *Nucl. Acids Res.* 24:1294-1303.

77. Broccoli D, Chong L, Oelmann S, Fernald AA, Marziliano N*, et al.* (1997) Comparison of the human and mouse genes endcoding the telomeric protein, TRF1: chromosomal localization, expression and conserved protein domains. *Human Mol. Genet.* 6:69-76.

78. Konig P, Giraldo R, Chapman L, & Rhodes D (1996) The crystal structure of the DNA-Binding domain of yeast RAP1 in complex with telomeric DNA. *Cell* 85:125-136.

79. Nishikawa T, Okamura H, Nagadoi A, Koig P, Rhodes D, *et al.* (2001) Structure of the DNA-binding domain of human telomeric protein, TRF1 and its interaction with telomeric DNA. *Nucleic Acids Res Suppl* (1):273-274.

80. Nishikawa T*, et al.* (2001) Solution structure of a telomeric DNA complex of human TRF1. *Structure* 9(12):1237-1251.

81. Evans SK & Lundblad V (2000) Positive and negative regulation of telomerase access to the telomere. *J. Cell Sci.* 113:3357-3364.

82. Fairall L, Chapman L, Moss H, de Lange T, & Rhodes D (2001) Structure of the the TRFH dimerization domain of the human telomeric proteins TRF1 and TRF2. *Mol. Cell* 8:351-361.

83. Shore D (1997) Telomerase and telomere-binding proteins-controlling the end game. *Trends in Biochemical Sciences* 22:233-235.

84. Hicke BJ, Celander DW, MacDonald GH, Price CM, & Cech TR (1990) Two versions of the gene encoding the 41-kilodalton subunit of the telomere binding protein of *Oxytricha nova*. *Proc. Natl. Acad. Sci. USA.* 87:1481-1485.

85. Fang G & Cech TR (1993) The B subunit of *Oxytricha* telomere binding protein promotes G-quartet formation by telomeric DNA. *Cell* 74:875-885.

86. Gray JT, Celander DW, Price CM, & Cech TR (1991) Cloning and expression of genes for the *Oxytricha* telomere-binding protein: specific subunit interactions in the telomeric complex. *Cell* 67:807-814.

87. Fang G, Gray JT, & Cech TR (1993) *Oxytricha* telomere-binding protein: separatable DNA-binding and dimerization domains of the a-subunit. *Genes and Development* 7:870-882.

88. Horvath MP, Schweiker VL, Bevilacqua JM, Ruggles JA, & Schultz SC (1998) Crystal structure of the *Oxytricha nova* telomere end binding protein complexed with single-strand DNA. *Cell* 95:963-974.

89. Marcand S, Wotton D, Gilson E, & Shore D (1997) Rap1p and telomere length regulation in yeast. *Ciba Found Symp* 211:76-93; discussion 93-103.

90. Hardy CFJ, Sussel L, & Shore D (1992) A RAP1-interacting protein involved in transcriptional silencing and telomere length regulation. *Genes and Development* 6:801-804.

91. Wotton D & Shore D (1997) A novel Rap1p-interacting factor, Rif2p, cooperates with Rif1p to regulate telomere length in *Saccharomyces cerevisiae*. *Genes Dev* 11(6):748-760.

92. Marcand S, Gilson E, & Shore D (1997) A protein-counting mechanism for telomere length regulation in yeast. *Science* 275:986-990.

93. Loayza D & De Lange T (2003) POT1 as a terminal transducer of TRF1 telomere length control. *Nature* 423(6943):1013-1018.

94. Wei C & Price CM (2003) Protecting the terminus: t-loops and telomere end-binding proteins. *Cell Mol Life Sci* 60(11):2283-2294.

95. Nugent CI, Hughes TR, Lue NF, & Lundblad V (1996) Cdc13p: a single-strand telomeric DNA binding protein with a dual role in yeast telomere maintenance. *Science* 274:249-252.

96. Lustig AJ (2001) Cdc13 subcomplexes regulate multiple telomere functions. *Nat Struct Biol* 8(4):297-299.

97. Chandra A, Hughes TR, Nugent CI, & Lundblad V (2001) Cdc13 both positively and negatively regulates telomere replication. *Genes Dev* 15(4):404-414.

98. Pennock E, Buckley K, & Lundblad V (2001) Cdc13 delivers separate complexes to the telomere for end protection and replication. *Cell* 104(3):387-396.

99. Grandin N, Damon C, & Charbonneau M (2000) Cdc13 cooperates with the yeast Ku proteins and Stn1 to regulate telomerase recruitment. *Mol. Cell. Biol.* 20:8397-8408.

100. Qi HaZ, V.A. (2000) The *Saccharomyces* telomere-binding protein Cdc13p interacts with both the catalytic subunit of DNA polymerase alpha and the telomerase associated Est1 proten. *Genes Dev* 14:1777-1788.

101. Grandin N, Reed SI, & Charbonneau M (1997) Stn1, a new *Saccharomyces cerevisiae* protein, is implicated in telomere size regulation in association with Cdc13. *Genes Dev* 11(4):512-527.

102. Martin V, Du LL, Rozenzhak S, & Russell P (2007) Protection of telomeres by a conserved Stn1-Ten1 complex. *Proc Natl Acad Sci USA* 104(35):14038-14043.

103. Ferreira MG & Cooper JP (2001) The fission yeast Taz1 protein protects chromosomes from Ku-dependent end -to-end fusions. *Molecular Cell* 7:55-63.

104. Miller KM & Cooper JP (2003) The telomere protein Taz1 is required to prevent and repair genomic DNA breaks. *Mol Cell* 11(2):303-313.

105. Kanoh J & Ishikawa F (2001) spRap1 and spRif1, recruited to telomeres by Taz1, are essential for telomere function in fission yeast. *Curr Biol* 11(20):1624-1630.

106. Miller KM, Ferreira MG, & Cooper JP (2005) Taz1, Rap1 and Rif1 act both interdependently and independently to maintain telomeres. *EMBO J* 24(17):3128-3135.

107. Bunch JT, Bae NS, Leonardi J, & Baumann P (2005) Distinct requirements for Pot1 in limiting telomere length and maintaining chromosome stability. *Mol Cell Biol* 25(13):5567-5578.

108. Palm W & de Lange T (2008) How shelterin protects mammalian telomeres. *Annu Rev Genet* 42:301-334.

109. de Lange T (2005) Shelterin: the protein complex that shapes and safeguards human telomeres. *Genes Dev* 19(18):2100-2110.

110. Liu D, O'Connor MS, Qin J, & Songyang Z (2004) Telosome, a mammalian telomere-associated complex formed by multiple telomeric proteins. *J Biol Chem* 279:51338-51342.

111. Ye JZ, Hockemeyer D, Krutchinsky AN, Loayza D, Hooper SM, *et al.* (2004) POT1-interacting protein PIP1: a telomere length regulator that recruits POT1 to the TIN2/TRF1 complex. *Genes Dev* 18(14):1649-1654.

112. Ye JZ, Donigian JR, van Overbeek M, Loayza D, Luo Y, Krutchinsky AN, *et al.* (2004) TIN2 binds TRF1 and TRF2 simultaneously and stabilizes the TRF2 complex on telomeres. *J Biol Chem* 279(45):47264-47271.

113. van Steensel B & de Lange T (1997) Control of telomere length by the human telomeric protein TRF1. *Nature* 385(6618):740-743.

114. van Steensel B, Smogorzewska A, & de Lange T (1998) TRF2 protects human telomeres from end-to-end fusions. *Cell* 92(3):401-413.

115. Kim SH, Kaminker P, & Campisi J (1999) TIN2, a new regulator of telomere length in human cells. *Nat. Genet.* 23:405-412.

116. Kim SH, Beausejour C, Davalos AR, Kaminker P, Heo SJ, *et al.* (2004) TIN2 mediates functions of TRF2 at human telomeres. *J Biol Chem* 279(42):43799-43804.

117. Li B, Oestreich S, & de Lange T (2000) Identification of human Rap1: implications for telomere evolution. *Cell* 101:471-483.

118. Hockemeyer D, Daniels JP, Takai H, & de Lange T (2006) Recent expansion of the telomere complex in rodents: two distinct POT1 proteins protect mouse telomeres. *Cell* 126:63-77.

119. Wu L, Multani AS, He H, Cosme-Blanco W, Deng Y, Deng JM, *et al.* (2006) Pot1 deficiency initiates DNA damage checkpoint activation and aberrant homologous recombination at telomeres. *Cell* 126(1):49-62.

120. Wang R, Podell, E.R., Zaug, A.J., Yang, Y., Baciu, P., Cech, T.R. and Lei, M. (2007) The POT1-TPP1 telomere complex is a telomerase processivity factor. *Nature* 445:506-510.

121. Xin H, *et al.* (2007) TPP1 is a homologue of ciliate TEBP-beta and interacts with POT1 to recruit telomerase. *Nature* 445(7127):559-562.

122. Lei M, Podell ER, & Cech TR (2004) Structure of human POT1 bound to telomeric single-stranded DNA provides a model for chromosome end protection. *Nat Struc Biol* 11:1223-1229.

123. Kelleher C, Kurth I, & Lingner J (2005) Human protetion of telomeres 1 (POT1) is a negative regulator of telomerase activity *in vitro*. *Mol Cell Biol* 25:808-818.

124. Cristofari G, Sikora K, & Lingner J (2007) Telomerase unplugged. *ACS Chem Biol* 2(3):155-158.

125. Song X, Leehy K, Warrington RT, Lamb JC, Surovtseva YV, *et al.* (2008) STN1 protects chromosome ends in *Arabidopsis thaliana*. *Proc Natl Acad Sci USA* 105(50):19815-19820.

126. Surovtseva YV*, et al.* (2009) Conserved telomere maintenance component 1 interacts with STN1 and maintains chromosome ends in higher eukaryotes. *Mol Cell* 36(2):207-218.

127. Karamysheva Z, Surovstova Y, Vespa L, Shakirov EV, & Shippen DE (2004) A C-terminal Myb-extension domain defines a novel family of double-strand telomeric DNA binding proteins in *Arabidopsis*. *J. Biol. Chem.* 279:47799-47807.

128. Hwang MG & Cho MH (2007) *Arabidopsis* thaliana telomeric DNA-binding protein 1 is required for telomere length homeostasis and its Myb-extension domain stabilizes plant telomeric DNA binding. *Nucleic Acids Res* 35(4):1333-1342.

129. Shakirov EV, Surovtseva Y, Osbun N, & Shippen DE (2005) The *Arabidopsis* Pot1 and Pot2 proteins function in telomre length homeostasis and chromosome end protection. *Mol Biol Cell* 2005 Sep;25(17):7725-33.

130. Surovtseva YV, Shakirov EV, Vespa L, Osbun N, Song X, *et al.* (2007) *Arabidopsis* POT1 associates with the telomerase RNP and is required for telomere maintenance. *EMBO J* 26(15):3653-3661.

131. Shakirov EV, McKnight TD, & Shippen DE (2009) POT1-independent single-strand telomeric DNA binding activities in Brassicaceae. *Plant J* 58(6):1004-1015.

132. Shakirov EV, Song X, Joseph JA, & Shippen DE (2009) POT1 proteins in green algae and land plants: DNA-binding properties and evidence of co-evolution with telomeric DNA. *Nucleic Acids Res* 37(22):7455-7467.

133. Autexier C & Greider CW (1994) Functional reconstitution of wild-type and mutant *Tetrahymena* telomerase. *Genes and Devel.* 8:563-575.

134. Beattie TL, Zhou W, Robinson MO, & Harrington L (1998) Reconstitution of human telomerase activity *in vitro*. *Curr. Biol.* 8:177-180.

135. Kim NW, Piatyszek MA, Prowse KR, Harley CB, West MD, *et al.* (1994) Specific association of human telomerase activity with immortal cells and cancer. *Science* 266:2011-2015.

136. Lingner J & Cech TR (1996) Purification of telomerase from *Euplotes aediculatus*-requirement of a primer 3' overhang. *Proc. Natl. Acad. Sci USA.* 93:10712-10171.

137. Collins K & Greider CW (1995) Utilization of ribonucleotides and RNA primers by *Tetrahymena* telomerase. *EMBO J.* 14(21):5422-32.

138. Cohn M & Blackburn EH (1995) Telomerase in yeast. *Science* 269:396-400.

139. Greenberg RA, Allsopp RC, Chin L, Morin GB, & DePinho RA (1998) Expression of mouse telomerase reverse transcriptase during development, differentiation and proliferation. *Oncogene* 16:1723-1730.

140. Guo W, Okamoto M, Park NH, & Lee YM (2001) Cloning and expression of hamster telomerase catalytic subunit cDNA. *Int J Mol Med* 8(1):73-78.

141. Kuramoto M, Ohsumi K, Kishimoto T, & Ishikawa F (2001) Identification and analyses of the *Xenopus* TERT gene that encodes the catalytic subunit of telomerase. *Gene* 277(1-2):101-110.

142. Meyerson M*,* Counter CM, Eaton EN, Ellisen LW, Steiner P, *et al.* (1997) hEST2, the putative human telomerase catalytic subunit gene, is up-regulated in tumor cells and during immortalization. *Cell* 90:785-795.

143. Fitzgerald MS*,* Riha K, Gao F, Ren S, McKnight TD, *et al.* (1999) Disruption of the telomerase catalytic subunit gene from *Arabidopsis* inactivates telomerase and leads to a slow loss of telomeric DNA. *Proc. Natl. Acad. Sci. USA* 96:14813-14818.

144. Eickbush TH (1997) Telomerase and retrotransposons: Which came first? *Science* 277:911-912.

145. O'Reilly M, Teichmann SA, & Rhodes D (1999) Telomerases. *Curr Opin Struct Biol* 9(1):56-65.

146. Beattie TL, Zhou W, Robinson MO, & Harrington L (2001) Functional multimerization of the human telomerase reverse transcriptase. *Mol Cell Biol* 21:6151-6160.

147. Collins K (2006) The biogenesis and regulation of telomerase holoenzymes. *Nat Rev Mol Cell Biol* 7:484-494.

148. Harrington L (2003) Biochemical aspects of telomerase function. *Cancer Lett* 194(2):139-154.

149. Autexier CaL, N.F. (2006) The structure and function of telomerase reverse transcriptase. *Annu Rev Biochem* 75:493-517.

150. Collins K (1999) Ciliate telomerase biochemistry. *Ann. Rev. Biochem.* 68:187-218.

151. Hammond PW, Lively TN, & Cech TR (1997) The anchor site of telomerase from *Euplotes aediculatus* revealed by photo-cross-linking to single- and double-stranded DNA primers. *Mol. Cell. Biol.* 17:296-308.

152. Greider CW (1991) Telomerase is processive. *Mol. Cell. Biol.* 11:4572-4580.

153. Prowse KR, Avilion AA, & Greider CW (1993) Identification of a nonprocessive telomerase activity from mouse cells. *Proc. Natl. Acad. Sci. USA.* 90:1493-1497.

154. Harrington L & Greider C (1991) Telomerase primer specificity and chromsome healing. *Nature* 353:451-453.

155. Bryan TM, Goodrich KJ, & Cech TR (2003) Tetrahymena telomerase is active as a monomer. *Mol Biol Cell* 14:4794-4804.

156. Ly H, Xu L, Rivera MA, Parslow TG, & Blackburn EH (2003) A role for a novel 'trans-pseudoknot' RNA-RNA interaction in the functional dimerization of human telomerase. *Genes Dev* 17(9):1078-1083.

157. Teixeira MT & Gilson E (2007) La sets the tone for telomerase assembly. *Nat Struct Mol Biol* 14(4):261-262.

158. Prescott J & Blackburn EH (1997) Telomerase RNA mutations in *Saccharomyces cerevisiae* alter telomerase action and reveal nonprocessivity *in vivo* and *in vitro*. *Genes and Dev.* 11:528-540.

159. Prescott J & Blackburn EH (1997) Functionally interacting telomerase RNAs in the yeast telomerase complex. *Genes and Dev.* 11:2790-2800.

160. Rivera MA & Blackburn EH (2004) Processive utilization of the human telomerase template: Lack of a requirement for template switching. *J Biol Chem*.

161. Chen ZZ*, ,* Xu RX, Jiang XD, Teng XH, Li GT, *et al.* (2006) Lack of telomerase activity in rabbit bone marrow stromal cells during differentiation along neural pathway. *Chin J Traumatol* 9(4):201-205.

162. Witkin KL & Collins K (2004) Holoenzyme proteins required for the physiological assembly and activity of telomerase. *Genes Dev* 18(10):1107-1118.

163. Bianchi A & deLange T (1999) Ku binds telomeric DNA *in vitro*. *J. Biol. Chem.* 274:21223-21227.

164. Hughes TR, Evans SK, Weilbaecher RG, & Lundblad V (2000) The Est3 protein is a subunit of yeast telomerase. *Curr Biol* 10(13):809-812.

165. Friedman KL, Heit JJ, Long DM, & Cech TR (2003) N-terminal domain of yeast telomerase reverse transcriptase: recruitment of Est3p to the telomerase complex. *Mol Biol Cell* 14(1):1-13.

166. Lundblad V ed (2006) *Budding yeast telomeres*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp 345-386.

167. Forsythe HL, Jarvis JL, Turner JW, Elmore LW, & Holt SE (2001) Stable association of hsp90 and p23, but Not hsp70, with active human telomerase. *J Biol Chem* 276(19):15571-15574.

168. Downs JA & Jackson SP (2004) A means to a DNA end: the many roles of Ku. *Nat Rev Mol Cell Biol* 5(5):367-378.

169. Chen JL, Blasco MA, & Greider C (2000) Secondary structure of vertebrate telomerase RNA. *Cell* 100:503-514.

170. Dandijinou AT, Lévesque N, Larose S, Lucier JF, Abou Elela S, *et al.* (2004) A phylogenetically based secondary structure for the yeast telomerase RNA. *Curr Biol* 14:1148-1158.

171. Leonardi J, Box JA, Bunch JT, & Baumann P (2008) TER1, the RNA subunit of fission yeast telomerase. *Nat Struc & Mol Biol* 15:26-33.

172. Chen JL & Greider CW (2004) An emerging consensus for telomerase RNA structure. *Proc Natl Acad Sci U S A* 101(41):14683-14684.

173. Autexier C & Greider CW (1995) Boundary elements of the *Tetrahymena* telomerase RNA template and alignment domains. *Genes and Dev.* 9:2227-2239.

174. Seto AG, Umansky K, Tzfati Y, Zaug AJ, Blackburn EH, *et al.* (2003) A template-proximal RNA paired element contributes to *Saccharomyces cerevisiae* telomerase activity. *RNA* 9(11):1323-1332.

175. Tzfati Y, Knight Z, Roy J, & Blackburn EH (2003) A novel pseudoknot element is essential for the action of a yeast telomerase. *Genes Dev* 17(14):1779-1788.

176. Gilley D & Blackburn EH (1996) Specific RNA residue interactions required for enzymatic functions of Tetrahymena telomerase. *Molecular and Cellular Biology* 16:66-75.

177. Antal M, Boros E, Solymosy F, & Kiss T (2002) Analysis of the structure of human telomerase RNA *in vivo*. *Nucleic Acids Res* 30(4):912-920.

178. Lai CK, Miller MC, & Collins K (2003) Roles for RNA in telomerase nucleotide and repeat addition processivity. *Mol Cell* 11(6):1673-1683.

179. Aigner S, Postberg J, Lipps HJ, & Cech TR (2003) The *Euplotes* La motif protein p43 has properties of a telomerase-specific subunit. *Biochemistry* 42(19):5736-5747.

180. Aigner S, Lingner J, Goodrich KJ, Grosshans CA, Shevchenko A, Mann M, *et al.* (2000) *Euplotes* telomerase contains an La motif protein produced by apparent translational frameshifting. *EMBO J* 19(22):6230-6239.

181. Seto AG, Livengood AJ, Tzfati Y, Blackburn EH, & Cech TR (2002) A bulged stem tethers Est1p to telomerase RNA in budding yeast. *Genes Dev* 16(21):2800-2812.

182. Stellwagen AE, Haimberger ZW, Veatch JR, & Gottschling DE (2003) Ku interacts with telomerase RNA to promote telomere addition at native and broken chromosome ends. *Genes Dev* 14:2384-2395.

183. Tesmer VM, Ford LP, Holt SE, Frank BC, Yi X, *et al.* (1999) Two inactive fragments of the integral RNA cooperate to assemble active telomerase with the human protein catalytic subunit (hTERT) *in vitro*. *Mol Cell Biol* 19:6207-6216.

184. Mitchell JR, Wood E, & Collins K (1999) A telomerase component is defective in the human disease dyskeratosis congenita. *Nature* 402:551-555.

185. Jady BE, Bertrand E, & Kiss T (2004) Human telomerase RNA and box H/ACA scaRNAs share a common Cajal body-specific localization signal. *J Cell Biol* 164(5):647-652.

186. Legassie JD & Jarstfer MB (2006) The unmasking of telomerase. *Structure (Camb)* 14:1603-1609.

187. Moriarty TJ, Marie-Egyptienne DT, & Autexier C (2004) Functional organization of repeat addition processivity and DNA synthesis determinants in the human telomerase multimer. *Mol Cell Biol* 24(9):3720-3733.

188. Harley CB, Futcher AB, & Greider CW (1990) Telomeres shorten during ageing of human fibroblasts. *Nature* 345:458-460.

189. Bodnar AG, Ouellette M, Frolkis M, Holt SE, Chiu CP*, et al.* (1998) Extension of life-span by introduction of telomerase into normal human cells. *Science* 279:349-352.

190. Shay JW & Bacchetti S (1997) A survey of telomerase activity in human cancer. *Eur. J. Cancer* 33:787-791.

191. Vaziri H & Benchimol S (1998) Reconstitution of telomerase activity in normal human cells leads to elongation of telomeres and extended replicative life span. *Curr Biol* 8(5):279-282.

192. Harley CB, Homayoun V, Counter CM, & Allsopp RC (1992) The telomere hypothesis of cellular aging. *Exp. Gerontol.* 27:375-382.

193. Kyo S, Takakura M, Fujiwara T, & Inoue M (2008) Understanding and exploiting hTERT promoter regulation for diagnosis and treatment of human cancers. *Cancer Sci* 99(8):1528-1538.

194. Cairney CJ, Hoare SF, Daidone MG, Zaffaroni N, & Keith WN (2008) High level of telomerase RNA gene expression is associated with chromatin modification, the ALT phenotype and poor prognosis in liposarcoma. *Br J Cancer* 98(8):1467-1474.

195. Chang M, Arneric M, & Lingner J (2007) Telomerase repeat addition processivity is increased at critically short telomeres in a Tel1-dependent manner in *Saccharomyces cerevisiae*. *Genes Dev* 21(19):2485-2494.

196. Zaug AJ, Podell ER, Nandakumar J, & Cech TR (2010) Functional interaction between telomere protein TPP1 and telomerase. *Genes Dev* 24(6):613-622.

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

198. Robart AR, O'Connor CM, & Collins K (2010) Ciliate telomerase RNA loop IV nucleotides promote hierarchical RNP assembly and holoenzyme stability. *RNA* 16(3):563-571.

199. Robart AR & Collins K (2010) Investigation of human telomerase holoenzyme assembly, activity, and processivity using disease-linked subunit variants. *J Biol Chem* 285(7):4375-4386.

200. Shay JW & Wright WE (2005) Senescence and immortalization: role of telomeres and telomerase. *Carcinogenesis* 26:867-874.

201. Shay JW & Wright WE (2006) Telomerase therapeutics for cancer: challenges and new directions. *Nat Rev Drug Discov* 5(7):577-584.

202. Cong YS, Wright WE, & Shay JW (2002) Human telomerase and its regulation. *Microbiol Mol Biol Rev* 66:407-425.

203. Grandori C, Cowley SM, James LP, & Eisenman RN (2000) The Myc/Max/Mad network and the transcriptional control of cell behavior. *Annu Rev Cell Dev Biol* 16:653-699.

204. Wu K-J*, et al.* (1999) Direct activation of TERT transcription by c-MYC. *Nature Genet.* 21:220-224.

205. Greenberg RA*,* O'Hagan RC, Deng H, Xiao Q, Hann SR, *et al.* (1999) Telomerase reverse transcriptase gene is a direct target of c-Myc but is not functionally equivalent in cellular transformation. *Oncogene* 18(5):1219-1226.

206. Oh S, Song Y, Yim J, & Kim TK (1999) The Wilms' tumor 1 tumor suppressor gene represses transcription of the human telomerase reverse transcriptase gene. *J Biol Chem* 274(52):37473-37478.

207. Oh S, Song YH, Kim UJ, Yim J, & Kim TK (1999) *In vivo* and *in vitro* analyses of Myc for differential promoter activities of the human telomerase (hTERT) gene in normal and tumor cells. *Biochem Biophys Res Commun* 263(2):361-365.

208. Englert C (1998) WT1--more than a transcription factor? *Trends Biochem Sci* 23(10):389-393.

209. Bodnar AG, Kim NW, Effros RB, & Chiu CP (1996) Mechanism of telomerase induction during T cell activation. *Exp Cell Res* 228(1):58-64.

210. Xu D*,* Popov N, Hou M, Wang Q, Björkholm M, *et al.* (2001) Switch from Myc/Max to Mad1/Max binding and decrease in histone acetylation at the telomerase reverse transcriptase promoter during differentiation of HL60 cells. *Proc Natl Acad Sci USA* 98(7):3826-3831.

211. Fujimoto K*,* Kyo S, Takakura M, Kanaya T, Kitagawa Y, *et al.* (2000) Identification and characterization of negative regulatory elements of the human telomerase catalytic subunit (hTERT) gene promoter: possible role of MZF-2 in transcriptional repression of hTERT. *Nucleic Acids Res* 28(13):2557-2562.

212. Crowe DL, Nguyen DC, Tsang KJ, & Kyo S (2001) E2F-1 represses transcription of the human telomerase reverse transcriptase gene. *Nucleic Acids Res* 29(13):2789-2794.

213. Yang H, Kyo S, Takatura M, & Sun L (2001) Autocrine transforming growth factor beta suppresses telomerase activity and transcription of human telomerase reverse transcriptase in human cancer cells. *Cell Growth Differ* 12(2):119-127.

214. Venkitaraman AR (2002) Cancer susceptibility and the functions of BRCA1 and BRCA2. *Cell* 108(2):171-182.

215. Xu D*,* Wang Q, Gruber A, Björkholm M, Chen Z, *et al.* (2000) Downregulation of telomerase reverse transcriptase mRNA expression by wild type p53 in human tumor cells. *Oncogene* 19(45):5123-5133.

216. Liu K*,* Schoonmaker MM, Levine BL, June CH, Hodes RJ, *et al.* (1999) Constitutive and regulated expression of telomerase reverse transcriptase (hTERT) in human lymphocytes. *Proc Natl Acad Sci USA* 96(9):5147-5152.

217. Lawlor MA & Alessi DR (2001) PKB/Akt: a key mediator of cell proliferation, survival and insulin responses? *J Cell Sci* 114(Pt 16):2903-2910.

218. Kang SS, Kwon T, Kwon DY, & Do SI (1999) Akt protein kinase enhances human telomerase activity through phosphorylation of telomerase reverse transcriptase subunit. *J Biol Chem* 274(19):13085-13090.

219. Liu WJ & Ding J (2001) Molecular mechanisms involved in the regulation of telomerase activity. *Sheng Li Ke Xue Jin Zhan* 32(3):220-224.

220. Minamino T & Kourembanas S (2001) Mechanisms of telomerase induction during vascular smooth muscle cell proliferation. *Circ Res* 89(3):237-243.

221. Kim JH, Park SM, Kang MR, Oh SY, Lee TH, *et al.* (2005) Ubiquitin ligase MKRN1 modulates telomere length homeostasis through a proteolysis of hTERT. *Genes Dev* 19(7):776-781.

222. Tomlinson RL, Ziegler TD, Supakorndej T, Terns RM, & Terns MP (2006) Cell cycle-regulated trafficking of human telomerase to telomeres. *Mol Biol Cell* 17(2):955-965.

223. Wong JM, Kusdra L, & Collins K (2002) Subnuclear shuttling of human telomerase induced by transformation and DNA damage. *Nat Cell Biol* 4(9):731-736.

224. Kanaya T, Kyo S, Hamada K, Takakura M, Kitagawa Y, *et al.* (2000) Adenoviral expression of p53 represses telomerase activity through down-regulation of human telomerase reverse transcriptase transcription. *Clin Cancer Res* 6(4):1239-1247.

225. Stone MD*,* Mihalusova M, O'connor CM, Prathapam R, Collins K, *et al.* (2007) Stepwise protein-mediated RNA folding direct assembly of telomerase ribonucleoprotein. *Nature* 446(458-461).

226. Mason DX, Autexier C, & Greider CW (2001) Tetrahymena proteins p80 and p95 are not core telomerase components. *Proc Natl Acad Sci USA* 98(22):12368-12373.

227. Karamysheva Z*,* Wang L, Shrode T, Bednenko J, Hurley LA, *et al.* (2003) Developmentally programmed gene elimination in *Euplotes crassus* facilitates a switch in the telomerase catalytic subunit. *Cell* 113(5):565-576.

228. Chapon C, Cech TR, & Zaug AJ (1997) Polyadenylation of telomerase RNA in budding yeast. *RNA* 3(11):1337-1351.

229. Seto AG, Zaug AJ, Sobel SG, Wolin SL, & Cech TR (1999) *Saccharomyces cerevisiae* telomerase is an Sm small nuclear ribonucleoprotein particle. *Nature.* 402(6764):898.

230. Baumann P (2005) Taking control of G-quadruplexes. *Nat Struct Mol Biol* 12(10):832-833.

231. Feng J, Funk WD, Wang SS, Weinrich SL, Avilion AA, *et al.* (1995) The RNA component of human telomerase. *Science* 269:1236-1241.

232. Meier UT (2005) The many facets of H/ACA ribonucleoproteins. *Chromosoma* 114(1):1-14.

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

234. Gallant P (2007) Control of transcription by Pontin and Reptin. *Trends Cell Biol* 17(4):187-192.

235. Venteicher AS, Meng Z, Mason PJ, Veenstra TD, & Artandi SE (2008) Identification of ATPases pontin and reptin as telomerase components essential for holoenzyme assembly. *Cell* 132(6):945-957.

236. Zhu Y, Tomlinson RL, Lukowiak AA, Terns RM, & Terns MP (2004) Telomerase RNA accumulates in Cajal bodies in human cancer cells. *Mol Biol Cell* 15(1):81-90.

237. Fu D & Collins K (2006) Human telomerase and Cajal body ribonucleoproteins share a unique specificity of Sm protein association. *Genes Dev* 20(5):531-536.

238. Holt SE*,* Dara L. Aisner, Joseph Baur, Valerie M. Tesmer, Marife Dy, *et al.* (1999) Functional requirement of p23 and Hsp90 in telomerase complexes. *Genes and Dev.* 13:817-826.

239. Seimiya H*,* Sawada H, Muramatsu Y, Shimizu M, Ohko K, *et al.* (2000) Involvement of 14-3-3 proteins in nuclear localization of telomerase. *EMBO J* 19(11):2652-2661.

240. Shippen DE (2005) Plant telomeres. *Telomeres*, ed Blackburn EH, de Lange, T. and Lunblad, V. (Cold Spring Habor Press, Cold Spring Harbor, NY), 2nd Ed, pp 525-550.

241. Riha K, McKnight TD, Griffing L, & Shippen DE (2001) Living with genome instability: plant responses to telomere dysfunction. *Science* 291:1797-1800.

242. Kannan K, Nelson ADL, & Shippen DE (2008) Dyskerin is a component of the *Arabidopsis* telomerase RNP required for telomere maintenance. *Mol Cell Biol* 28:2332-2341.

243. Chan A, Boule JB, & Zakian VA (2008) Two pathways recruit telomerase to *Saccharomyces cerevisiae* telomeres. *PLoS Genet* 4(10):e1000236.

244. Lingner J*, et al.* (1997) Reverse transcriptase motifs in the catalytic subunit of telomerase. *Science* 276:561-567.

245. Hughes TR, Weilbaecher RG, Walterscheid M, & Lundblad V (2000) Identification of the single-strand telomeric DNA binding domain of the *Saccharomyces cerevisiae* Cdc13 protein. *Proc Natl Acad Sci USA* 97(12):6457-6462.

246. Zappulla DC & Cech TR (2004) Yeast telomerase RNA: a flexible scaffold for protein subunits. *Proc Natl Acad Sci USA* 101:10024-10029.

247. Evans SK & Lundblad V (1999) Est1 and Cdc13 as comediators of telomerase access. *Science* 286(5437):117-120.

248. Peterson SE*, et al.* (2001) The function of a stem-loop in telomerase RNA is linked to the DNA repair protein Ku. *Nature Genetics* 27:64-67.

249. Gravel S, Larrivee M, Labrecque P, & Wellinger RJ (1998) Yeast Ku as a regulator of chromsomal DNA structure. *Science* 280:741-744.

250. Gallego ME, Bleuyard JY, Daoudal-Cotterell S, Jallut N, & White CI (2003) Ku80 plays a role in non-homologous recombination but is not required for T-DNA integration in *Arabidopsis*. *Plant J.* 35:557-565.

251. Gallego ME, Jalut N, & White CI (2003) Telomerase dependence of telomere lengthening in Ku80 mutant *Arabidopsis*. *Plant Cell* 15(3):782-789.

252. Fisher TS, Taggart AK, & Zakian VA (2004) Cell cycle-dependent regulation of yeast telomerase by Ku. *Nat Struct Mol Biol* 11(12):1198-1205.

253. Leonardi J, Box JA, Bunch JT, & Baumann P (2008) TER1, the RNA subunit of fission yeast telomerase. *Nat Struct Mol Biol* 15(1):26-33.

254. Reichenbach P*,* Höss M, Azzalin CM, Nabholz M, Bucher Pet, *et al.* (2003) A human homolog of yeast Est1 associates with telomerase and uncaps chromosome ends when overexpressed. *Curr Biol* 13(7):568-574.

255. Zhou XZ & Lu KP (2001) The Pin2/TRF1-interacting protein PinX1 is an potent telomerase inhibitor. *Cell* 107:347-359.

256. Lin J & Blackburn EH (2004) Nucleolar protein PinX1p regulates telomerase by sequestering its protein catalytic subunit in an inactive complex lacking telomerase RNA. *Genes Dev* 18(4):387-396.

257. Boule JB, Vega LR, & Zakian VA (2005) The yeast Pif1p helicase removes telomerase from telomeric DNA. *Nature* 438(7064):57-61.

258. Azzalin CM, Reichenbach P, Khoriauli L, Giulotto E, & Lingner J (2007) Telomeric repeat containing RNA and RNA surveillance factors at mammalian chromosome ends. *Science* 318(5851):798-801.

259. Luke B*, et al.* (2008) The Rat1p 5' to 3' exonuclease degrades telomeric repeat-containing RNA and promotes telomere elongation in *Saccharomyces cerevisiae*. *Mol Cell* 32(4):465-477.

260. Schoeftner S & Blasco MA (2008) Developmentally regulated transcription of mammalian telomeres by DNA-dependent RNA polymerase II. *Nat Cell Biol* 10(2):228-236.

261. Feuerhahn S, Iglesias N, Panza A, Porro A, & Lingner J (2010) TERRA biogenesis, turnover and implications for function. *FEBS Lett* 584(17):3812-3818.

262. Nergadze SG*, et al.* (2009) CpG-island promoters drive transcription of human telomeres. *RNA* 15(12):2186-2194.

263. Chan CS & Tye BK (1983) Organization of DNA sequences and replication origins at yeast telomeres. *Cell* 33(2):563-573.

264. Zhu Y, Zou S, Wright DA, & Voytas DF (1999) Tagging chromatin with retrotransposons: target specificity of the *Saccharomyces* Ty5 retrotransposon changes with the chromosomal localization of Sir3p and Sir4p. *Genes Dev* 13(20):2738-2749.

265. Vrbsky J*,* Akimcheva S, Watson JM, Turner TL, Daxinger L*, et al.* (2010) siRNA-mediated methylation of *Arabidopsis* telomeres. *PLoS Genet* 6(6):e1000986.

266. Redon S, Reichenbach P, & Lingner J (2010) The non-coding RNA TERRA is a natural ligand and direct inhibitor of human telomerase. *Nucleic Acids Res* 38(17):5797-5806.

267. Shakirov EV & Shippen DE (2004) Length regulation and dynamics of individual telomere tracts in wild-type *Arabidopsis*. *Plant Cell* 16(8):1959-1967.

268. Heacock M, Spangler E, Riha K, Puizina J, & Shippen DE (2004) Molecular analysis of telomere fusions in *Arabidopsis*: multiple pathways for chromosome end-joining. *EMBO J* 23(11):2304-2313.

269. Osterhage JL & Friedman KL (2009) Chromosome end maintenance by telomerase. *J Biol Chem* 284(24):16061-16065.

270. Artandi SE & DePinho RA (2010) Telomeres and telomerase in cancer. *Carcinogenesis* 31(1):9-18.

271. Theobald DL & Wuttke DS (2004) Prediction of multiple tandem OB-fold domains in telomere end-biding proteins Pot1 and Cdc13. *Structure (Camb)* 12:1877-1879.

272. Nandakumar J, Podell ER, & Cech TR (2010) How telomeric protein POT1 avoids RNA to achieve specificity for single-stranded DNA. *Proc Natl Acad Sci USA* 107(2):651-656.

273. Xin H, Liu D, & Songyang Z (2008) The telosome/shelterin complex and its functions. *Genome Biol* 9(9):232.

274. Shakirov EV, Perroud PF, Nelson AD, Cannell ME, Quatrano RS, *et al.* (2010) Protection of Telomeres 1 is required for telomere integrity in the moss Physcomitrella patens. *Plant Cell* 22(6):1838-1848.

275. Gilley D & Blackburn EH (1999) The telomerase RNA pseudoknot is critical for the stable assembly of a catalytically active ribonucleoprotein. *Proc Natl Acad Sci USA* 96(12):6621-6625.

276. Tzfati Y, Fulton TB, Roy J, & Blackburn EH (2000) Template boundary in a yeast telomerase specified by RNA structure. *Science* 288(5467):863-867.

277. Lukowiak AA, Narayanan A, Li ZH, Terns RM, & Terns MP (2001) The snoRNA domain of vertebrate telomerase RNA functions to localize the RNA within the nucleus. *RNA* 7(12):1833-1844.

278. Autexier C & Greider CW (1994) Functional reconstitution of wild-type and mutant Tetrahymena telomerase. *Genes Dev* 8(5):563-575.

279. Beattie TL, Zhou W, Robinson MO, & Harrington L (1998) Reconstitution of human telomerase activity *in vitro*. *Curr Biol* 8(3):177-180.

280. Zappulla DC, Goodrich K, & Cech TR (2005) A miniature yeast telomerase RNA functions *in vivo* and reconstitutes activity *in vitro*. *Nat Struct Mol Biol* 12(12):1072-1077.

281. Taggart AK, Teng SC, & Zakian VA (2002) Est1p as a cell cycle-regulated activator of telomere-bound telomerase. *Science* 297(5583):1023-1026.

282. Zappulla DC & Cech TR (2004) Yeast telomerase RNA: a flexible scaffold for protein subunits. *Proc Natl Acad Sci USA* 101(27):10024-10029.

283. Stellwagen AE, Haimberger ZW, Veatch JR, & Gottschling DE (2003) Ku interacts with telomerase RNA to promote telomere addition at native and broken chromosome ends. *Genes Dev* 17(19):2384-2395.

284. Ting NS, Yu Y, Pohorelic B, Lees-Miller SP, & Beattie TL (2005) Human Ku70/80 interacts directly with hTR, the RNA component of human telomerase. *Nucleic Acids Res* 33(7):2090-2098.

285. Webb CJ & Zakian VA (2008) Identification and characterization of the *Schizosaccharomyces pombe* TER1 telomerase RNA. *Nat Struct Mol Biol* 15(1):34-42.

286. Fisher TS & Zakian VA (2005) Ku: a multifunctional protein involved in telomere maintenance. *DNA Repair (Amst)* 4(11):1215-1226.

287. Cohen SB*,* Graham ME, Lovrecz GO, Bache N, Robinson PJ, *et al.* (2007) Protein composition of catalytically active human telomerase from immortal cells. *Science* 315(5820):1850-1853.

288. Feng J*,* Funk WD, Wang SS, Weinrich SL, Avilion AA, *et al.* (1995) The RNA component of human telomerase. *Science* 269(5228):1236-1241.

289. Liu Y, Kha H, Ungrin M, Robinson MO, & Harrington L (2002) Preferential maintenance of critically short telomeres in mammalian cells heterozygous for mTert. *Proc Natl Acad Sci USA* 99(6):3597-3602.

290. Heacock ML, Idol RA, Friesner JD, Britt AB, & Shippen DE (2007) Telomere dynamics and fusion of critically shortened telomeres in plants lacking DNA ligase IV. *Nucleic Acids Res* 35(19):6490-6500.

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

292. Karamysheva ZN, Surovtseva YV, Vespa L, Shakirov EV, & Shippen DE (2004) A C-terminal Myb extension domain defines a novel family of double-strand telomeric DNA-binding proteins in *Arabidopsis*. *J Biol Chem* 279(46):47799-47807.

293. Chen JL & Greider CW (2005) Functional analysis of the pseudoknot structure in human telomerase RNA. *Proc Natl Acad Sci USA* 102(23):8080-8085; discussion 8077-8089.

294. Qiao F & Cech TR (2008) Triple-helix structure in telomerase RNA contributes to catalysis. *Nat Struct Mol Biol* 15(6):634-640.

295. Miller MC & Collins K (2002) Telomerase recognizes its template by using an adjacent RNA motif. *Proc Natl Acad Sci USA* 99(10):6585-6590.

296. Riha K, Watson JM, Parkey J, & Shippen DE (2002) Telomere length deregulation and enhanced sensitivity to genotoxic stress in *Arabidopsis* mutants deficient in Ku70. *EMBO J* 21(11):2819-2826.

297. Menges M & Murray JAH (2002) Synchronous *Arabidopsis* suspension cultures for anlaysis of cell-cycle gene activity. *Plant J.* 30:203-212.

298. Herbert BS, Hochreiter AE, Wright WE, & Shay JW (2006) Nonradioactive detection of telomerase activity using the telomeric repeat amplification protocol. *Nat Protoc* 1(3):1583-1590.

299. Wege H, Chui MS, Le HT, Tran JM, & Zern MA (2003) SYBR Green real-time telomeric repeat amplification protocol for the rapid quantification of telomerase activity. *Nucleic Acids Res* 31(2):E3-3.

300. Liu XD*, et al.* (2010) Down-regulation of telomerase activity and activation of caspase-3 are responsible for tanshinone I-induced apoptosis in monocyte leukemia cells *in vitro*. *Int J Mol Sci* 11(6):2267-2280.

301. Liu JP*, et al.* (2010) Regulation of telomerase activity by apparently opposing elements. *Ageing Res Rev* 9(3):245-256.

302. Cairney CJ & Keith WN (2008) Telomerase redifined: integrated regulation of hTR and hTERT for telomere maintenance and telomerase activity. *Biochemie* 90:13-23.

303. Hsu M, Yu EY, Singh SM, & Lue NF (2007) Mutual dependence of Candida albicans Est1p and Est3p in telomerase assembly and activation. *Eukaryot Cell* 6(8):1330-1338.

304. Madireddi MT, Davis MC, & Allis CD (1994) Identification of a novel polypeptide involved in the formation of DNA-containing vesicles during macronuclear development in *Tetrahymena*. *Dev Biol* 165(2):418-431.

305. Pogacic V, Dragon F, & Filipowicz W (2000) Human H/ACA small nucleolar RNPs and telomerase share evolutionarily conserved proteins NHP2 and NOP10. *Mol Cell Biol* 20(23):9028-9040.

306. He J*, et al.* (2002) Targeted disruption of Dkc1, the gene mutated in X-linked dyskeratosis congenita, causes embryonic lethality in mice. *Oncogene* 21(50):7740-7744.

307. Fu D & Collins K (2003) Distinct biogenesis pathways for human telomerase RNA and H/ACA small nucleolar RNAs. *Mol Cell* 11(5):1361-1372.

308. Fu D & Collins K (2007) Purification of human telomerase complexes identifies factors involved in telomerase biogenesis and telomere length regulation. *Mol Cell* 28:773-785.

309. Singh SM & Lue NF (2003) Ever shorter telomere 1 (EST1)-dependent reverse transcription by Candida telomerase *in vitro*: evidence in support of an activating function. *Proc Natl Acad Sci USA* 100:5718-5723.

310. Taggart AK, Teng SC, & Zakian VA (2002) Est1p as a cell cycle-regulated activator of telomere-bound telomerase. *Science* 297:1023-1026.

311. Lee J, Mandell, E.K., Tucey, T.M, Morris, D.K. and Lundblad, V. (2008) The Est3 protein associates with yeast telomerase through and OB-fold domain. *Nat. Struct. & Mol Bio* 15:990-997.

312. Ting NS, Yu Y, Pohorelic B, Lees-Miller SP, & Beattie TL (2005) Human Ku70/80 interacts directly with hTR, the RNA component of human telomerase. *Nucleic Acids Res* 33:2090-2098.

313. Cifuentes-Rojas C. KK, Tseng L. and Shippen DE (2010) Dissecting the *Arabidopsis* telomerase RNP. *Proc. Natl. Acad. Sci. USA.* In press.

314. Baumann P, Podell ER, & Cech TR (2002) Human Pot1 (Protection of telomere) protein: cytolocalization, gene structure, and alternative splicing. *Mol Cell Biol* 22:8079-8087.

315. He H*,* Multani AS, Cosme-Blanco W, Tahara H, Ma J, *et al.* (2006) POT1b protects telomeres from end-to-end chromosomal fusions and aberrant homologous recombination. *EMBO J* 25(21):5180-5190.

316. Hockemeyer D, Palm W, Else T, Daniels JP, Takai KK, *et al.* (2007) Telomere protection by mammalian Pot1 requires interaction with Tpp1. *Nat Struct Mol Biol* 14(8):754-761.

317. Cifuentes-Rojas C, Kannan K, Tseng L & Shippen DE. (2010) Two RNA subunits and POT1a are components of *Arabidopsis* telomerase. *Proc. Natl. Acad. Sci. USA*. In press.

318. Fitzgerald MS, McKnight TD, & Shippen DE (1996) Characterization and developmental patterns of telomerase expression in plants. *Proc. Natl. Acad. Sci.* 93:14422-14427.

319. Webb CJ & Zakian VA (2008) Identification and characterization of the *Schizosaccharomyces pombe* TER1 telomerase RNA. *Nat Struc & Mol Biol* 15:34-42.

320. Blasco MA, Funk W, Villeponteau B, & Greider CW (1995) Functional characterization and developmental regulation of mouse telomerase RNA. *Science* 269:1267-1270.

321. Hinkley CS, Blasco MA, Funk WD, Feng J, Villeponteau B, *et al.* (1998) The mouse telomerase RNA 5' end lies just upstream of the telomerase RNA template sequence. *Nucl. Acids Res.* 26:532-536.

322. Schuler MA (2008) Splice site requirements and switches in plants. *Curr Top Microbiol Immunol* 326:39-59.

323. Wiebauer K, Herrero JJ, & Filipowicz W (1988) Nuclear pre-mRNA processing in plants: distinct modes of 3'-splice-site selection in plants and animals. *Mol Cell Biol* 8(5):2042-2051.

324. Box JA, Bunch JT, Tang W, & Baumann P (2008) Spliceosomal cleavage generates the 3' end of telomerase RNA. *Nature* 456(7224):910-914.

325. Wang W, Skopp R, Scofield M, & Price C (1992) *Euplotes crassus* has genes encoding telomere-binding proteins and telomere-binding protein homologs. *Nucleic Acids Res* 20(24):6621-6629.

326. Jacob NK, Lescasse R, Linger BR, & Price CM (2007) Tetrahymena POT1a regulates telomere length and prevents activation of a cell cycle checkpoint. *Mol Cell Biol* 27:1592-1601.

327. Linger BR & Price CM (2009) Conservation of telomere protein complexes: shuffling through evolution. *Crit Rev Biochem Mol Biol* 44(6):434-446.

328. Prathapam R, Witkin KL, O'Connor CM, & Collins K (2005) A telomerase holoenzyme protein enhances telomerase RNA assembly with telomerase reverse transcriptase. *Nat Struct Mol Biol* 12(3):252-257.

329. Pfaffl MW (2001) A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res* 29(9):e45.

330. Lingner J & Cech TR (1998) Telomerase and chromosome end maintenance. *Curr Opin Genet Dev* 8(2):226-232.

331. Singer MS & Gottschling DE (1994) TLC1: template RNA component of *Saccharomyces cerevisiae* telomerase. *Science* 266:404-409.

332. Autexier C, Pruzan R, Funk WD, & Greider CW (1996) Reconstitution of human telomerase activity and identification of a minimal functional region of the human telomerase RNA. *EMBO J.* 15(21):5928-35

333. Witkin KL, Prathapam R, & Collins K (2007) Positive and negative regulation of *Tetrahymena* telomerase holoenzyme. *Mol Cell Biol* 27:2074-2083.

334. Trujillo KM, Bunch JT, & Baumann P (2005) Extended DNA binding site in Pot1 broadens sequence specificity to allow recognition of heterogeneous fission yeast telomeres. *J Biol Chem* 280:9119-9128.

335. Pitt CW, Moreau E, Lunness PA, & Doonan JH (2004) The pot1+ homologue in Aspergillus nidulans is required for ordering mitotic events. *J. Cell Sci.* 117:199-209.

336. Churikov D, Wei C, & Price CM (2006) Vertebrate POT1 restricts G-overhang length and prevents activation of a telomeric DNA damage checkpoint but is dispensable for overhang protection. *Mol Cell Biol* 26(18):6971-6982.

337. Price CM (1994) Telomere binding proteins of ciliated protozoa. *Nucleic acids in molecular biology* 9:299–307.

338. Gonzalo S*, et al.* (2006) DNA methyltransferases control telomere length and telomere recombination in mammalian cells. *Nat Cell Biol* 8(4):416-424.

339. Hockemeyer D, Palm W, Wang RC, Couto SS, & de Lange T (2008) Engineered telomere degradation models dyskeratosis congenita. *Genes Dev* 22(13):1773-1785.

340. Bhattacharyya A & Blackburn EH (1994) Architecture of telomerase RNAs. *EMBO J* 13(23):5721-5730.

341. Romero DP & Blackburn EH (1991) A conserved secondary structure for telomerase RNA. *Cell* 67:343-353.

342. Zuker M (2003) Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res* 31(13):3406-3415.

343. Parisien M & Major F (2008) The MC-Fold and MC-Sym pipeline infers RNA structure from sequence data. *Nature* 452(7183):51-55.

344. Lin J-J & Zakian VA (1995) An *in vitro* assay for *Saccharomyces* telomerase requires EST1. *Cell* 81:1127-1135.

345. Steiner BR, Hidaka K, & Futcher B (1996) Association of the Est1 protein with telomerase activity in yeast. *Proc Natl Acad Sci USA* 93(7):2817-2821.

346. DeZwaan DC & Freeman BC (2009) The conserved Est1 protein stimulates telomerase DNA extension activity. *Proc Natl Acad Sci USA* 106(41):17337-17342.

347. McGuffin LJ, Bryson K, & Jones DT (2000) The PSIPRED protein structure prediction server. *Bioinformatics* 16(4):404-405.

348. Larkin MA*, et al.* (2007) Clustal W and Clustal X version 2.0. *Bioinformatics* 23(21):2947-2948.

349. Pei J & Grishin NV (2007) PROMALS: towards accurate multiple sequence alignments of distantly related proteins. *Bioinformatics* 23(7):802-808.

350. Arnold K, Bordoli L, Kopp J, & Schwede T (2006) The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. *Bioinformatics* 22(2):195-201.

351. Guex N & Peitsch M (1997) SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis* 18(15):2714-2723.

352. Greider CW (1995) Telomerase biochemistry and regulation. *Telomeres*, ed Greider CW. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp 35-68.

353. Greider CW (1990) Telomeres, telomerase and senescence. *BioEssays* 12:363-369.

354. Greider CW (1998) Telomerase activity, cell proliferation, and cancer. *Proc. Natl. Acad. Sci. USA.* 95:90-92.

355. Collins K (1996) Structure and function of telomerase. *Curr. Opin. Cell Biol.* 8:374-380.

356. Cohen SB, Graham ME, Lovrecz GO, Bache N, Robinson PJ, *et al.* (2007) Protein composition of catalytically active human telomerase from immortal cells. *Science* 315:1850-1853.

357. Cech TR (2004) Beginning to understand the end of the chromosome. *Cell* 116(2):273-279.

358. Chen JL & Greider CW (2004) Telomerase RNA structure and function: implications for dyskeratosis congenita. *Trends Biochem Sci* 29(4):183-192.

359. Mason PJ, Wilson DB, & Bessler M (2005) Dyskeratosis congenita-a disease of dysfunctional telomere maintenance. *Curr Mol Med* 5:159-170.

360. Vulliamy T, Marrone A, Goldman F, Dearlove A, Bessler M, *et al.* (2001) The RNA component of telomerase is mutated in autosomal dominant dyskeratosis congenita. *Nature* 413(6854):432-435.

361. McCormick-Graham M & Romero DP (1995) Ciliate telomerase RNA structural features. *Nucl. Acids Res.* 23:1091-1097.

362. Hebsgaard SM, Korning PG, Tolstrup N, Engelbrecht J, Rouzé P, *et al.* (1996) Splice site prediction in *Arabidopsis thaliana* pre-mRNA by combining local and global sequence information. *Nucleic Acids Res* 24(17):3439-3452.

363. Brunak S, Engelbrecht J, & Knudsen S (1991) Prediction of human mRNA donor and acceptor sites from the DNA sequence. *J Mol Biol* 220(1):49-65.

364. Tolstrup N, Rouze P, & Brunak S (1997) A branch point consensus from *Arabidopsis* found by non-circular analysis allows for better prediction of acceptor sites. *Nucleic Acids Res* 25(15):3159-3163.

365. Bacchetti S & Counter CM (1995) Telomeres and telomerase in human cancer. *Int. J. Oncol.* 7(423-432).

366. Harley CB (1994) Telomerase, cell immortality, and cancer. *Cold Spring Harb. Symp. Quant. Biol.* 59:307-315.

367. Harley CB & Sherwood SW (1997) Telomerase, checkpoints and cancer. *Cancer Surveys* 29:263-284.

368. Harley CB & Villeponteau B (1995) Telomeres and telomerase in aging and cancer. *Curr. Opin. Genet. Devel.* 5:249-255.

369. Oulton R & Harrington L (2000) Telomeres, telomerase, and cancer: life on the edge of genomic stability. *Curr. Opin. Oncol.* 12:74-81.

370. Mason PJ (2003) Stem cells, telomerase and dyskeratosis congenita. *BioEssays* 25:126-133.

371. Cristofari G & Lingner J (2006) The telomerase ribonucleoprotein particle. *Telomeres*, eds de Lange T, Lundblad V, & Blackburn EH. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY. Pp, 21-47.

372. Blackburn EH (1992) Telomerases. *Annu. Rev. Biochem.* 61:113-129.

373. Kim MM, Rivera MA, Botchkina IL, Shalaby R, Thor AD, *et al.* (2001) A low threshold level of expression of mutant-template telomerase RNA inhibits human tumor cell proliferation. *Proc Natl Acad Sci USA* 98(14):7982-7987.

374. Sontheimer EJ & Steitz JA (1992) Three novel functional variants of human U5 small nuclear RNA. *Mol Cell Biol* 12(2):734-746.

375. Chen L, Lullo DJ, Ma E, Celniker SE, Rio DC, *et al.* (2005) Identification and analysis of U5 snRNA variants in *Drosophila*. *RNA* 11(10):1473-1477.

376. Gunderson JH, Sogin ML, Wollett G, Hollingdale M, de la Cruz VF, *et al.* (1987) Structurally distinct, stage-specific ribosomes occur in *Plasmodium*. *Science* 238(4829):933-937.

377. Waters AP, Syin C, & McCutchan TF (1989) Developmental regulation of stage-specific ribosome populations in *Plasmodium*. *Nature* 342(6248):438-440.

378. Osterhage JL, Talley, J.M. and Friedman, K.L. (2006) Proteasome-dependent degradation of Est1p regulates the cell cycle-restricted assembly of telomerase in *Saccharomyces cerevisiae*. *Nat. Struct. & Mol Bio* 13:720-728.

379. Bianchi A & Shore D (2008) How telomerase reaches its end: mechanism of telomerase regulation by the telomeric complex. *Mol Cell* 31(2):153-165.

380. Lingner J, Cech TR, Hughes TR, & Lundblad V (1997) Three Ever Shorter Telomere (EST) genes are dispensable for *in vitro* yeast telomerase activity. *Proc Natl Acad Sci USA* 94(21):11190-11195.

381. Virta-Pearlman V, Morris DK, & Lundblad V (1996) Est1 has the properties of a single-stranded telomere end-binding protein. *Genes Dev* 10(24):3094-3104.

382. Berman AJ, Gooding AR, & Cech TR (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(20):4965-4976.

APPENDICES

APPENDIX I

MINIMAL PRE-mRNA SUBSTRATES WITH NATURAL AND CONVERTED SITES FOR FULL-ROUND U INSERTION AND U DELETION RNA EDITING IN TRYPANOSOMES*

Summary

Trypanosome RNA editing by uridylate insertion or deletion cycles is a mitochondrial mRNA maturation process catalyzed by multi-subunit complexes. A full-round of editing entails three consecutive steps directed by partially complementary guide RNAs: pre-mRNA cleavage, U addition or removal, and ligation. The structural and functional composition of editing complexes is intensively studied, but their molecular interactions in and around editing sites are not completely understood. In this study, we performed a systematic analysis of distal RNA requirements for full-round insertion and deletion by purified editosomes. We define minimal substrates for efficient editing of A6 and CYb model transcripts, and established a new substrate, RPS12.

*Reprinted with permission from: "Minimal pre-mRNA substrates with natural and converted sites for full-round U insertion and U deletion RNA editing in trypanosomes" by Cifuentes-Rojas C, Halbig K, Sacharidou A, De Nova-Ocampo M, Cruz-Reyes J. 2005 *Nucleic Acids Res.* Nov 23; Vol 33, No. 20:6610-20. Copyright 2005 Oxford University Press.

Important differences were observed in the composition of substrates for insertion and deletion. Furthermore, we also showed for the first time that natural sites can be artificially converted in both directions: from deletion to insertion or from insertion to deletion. Our site conversions enabled a direct comparison of the two editing kinds at common sites during substrate minimization and demonstrate that all basic determinants directing the editosome to carry out full-round insertion or deletion reside within each editing site. Surprisingly, we were able to engineer a deletion site into CYb, which exclusively undergoes insertion in nature.

Introduction

Maturation of mitochondrial pre-mRNAs in kinetoplastid species including *Trypanosoma* and the evolutionarily distant *Leishmania* involves three-step cycles of uridylate insertion or deletion at many editing sites (ESs). Multisubunit editing complexes switch between insertion and deletion modes as editing progresses with a general 3'-to-5' polarity. The multisubunit composition and function of editing complexes is currently under intense study [for reviews see (383-385)], and their interactions with editing substrates have not been identified. Each editing cycle, directed by partially complementary mitochondrial guide RNAs (gRNAs), entails three basic consecutive catalytic steps: first, pre-mRNA cleavage; second, 30 terminal processing of the upstream piece by either nucleotide addition mediated by terminal U transferase 'TUTase' (in insertion), or

nucleotide removal by a U-specific exonuclease (in deletion); finally, mRNA resealing by ligation (386-389). Apart from these basic catalytic steps, annealing and unwinding steps are also likely involved (390-391). The most evident landmarks of deletion or insertion ESs are unpaired mRNA uridylates or unpaired gRNA purines, respectively. Site-specific mutagenesis affecting the ES or adjoining residues, particularly their potential for mRNA/gRNA pairing can significantly impact the specificity and efficiency of full-round and partial ('pre-cleaved') editing (387-388, 392).

In trypanosomes, it is generally accepted that natural ESs lie between two flanking duplexes: a proximal upstream duplex formed by a pre-mRNA 5' purine-rich and gRNA 3' poly(U) sequences (386,387,423), and an adjacent pre-mRNA/gRNA downstream 'anchor' duplex that directs cleavage (386,387,388,389). Site-specific mutagenesis of the gRNA 30 region, to artificially stabilize the upstream duplex, can enhance full-round editing *in vitro* (392.444). In *Leishmania*, the need for such upstream duplex remains somewhat controversial (448,449).

Trypanosome full-round deletion and insertion are currently studied *in vitro* using substrates representing fragments of the pre-edited domain in two mRNAs encoding ATPase subunit 6 (A6) and apo-cytochrome b (CYb) (387,444,447). The natural A6 pre-mRNA contains numerous insertion and deletion ESs, but only ES1 and ES2 are currently used to analyze full-round deletion and insertion *in vitro*, respectively. A6 substrates between 65 and 72 nt

long have been used (387,444). In contrast, CYb pre-mRNA exclusively contains insertion ESs, of which only ES1 has been analyzed *in vitro* using a CYb 55 nt substrate (444). The influence of pre-mRNA structural properties, such as substrate composition, folding and potential cis-elements on full-round insertion and deletion, has not been systematically characterized in trypanosomes. This is of interest because the two kinds of editing in a common complex may be differentially regulated, as they involve separate activities and enzymes, and there is evidence for their physical separation in the complex (426,427,428,429). Also, different gRNA features appear to be required for deletion and insertion (392,444,446). Thus, it is feasible that editing complex interactions with a processing site and its environs differs between the two editing types.

Here, we have analyzed the effect of substrate minimization on full-round insertion and deletion at natural and converted sites in different substrates. Our studies comparing catalysis by moderately and significantly enriched editing complexes show that, at least for some substrates, the latter are more sensitive to substrate minimization. This suggests that one or more associated RNA-binding activities are sensitive to the overall size of RNAs, and also to stringent purification of the complexes. Substrate minimization had significantly distinct effects on insertion and deletion, and the extent and direction of such effects varied among substrates. This implies that editing complexes are quite sensitive to RNA context. We also show for first time that natural sites can be converted in either direction, from deletion to insertion and vice versa. Such site conversion

not only allowed comparison of both types of editing at specific sites during substrate minimization, but also underscores the basic nature of the RNA determinants that direct complexes to carry out full-round insertion or deletion.

Results

A6 insertion and deletion editing

The influence that pre-mRNAs features may have on editing efficiency, including potential cis-elements, substrate composition and RNA folding, has not been systematically studied. We began analyzing the contribution of upstream and downstream pre-mRNA features, by performing gradual terminal truncations on an A6 72 nt substrate (RNA1, which includes the 30-most 26 ESs) paired to a gRNA that directs 3U insertion at the natural ES2 (Fig. A-1A). In these studies, we used catalytic complexes that were enriched by Q-Sepharose ion-exchange chromatography ('Q fraction'), or complexes that were significantly purified by consecutive Q-Sepharose and DNA-cellulose chromatography ('D fraction'), as previously reported (28). Deletions of 10 nt from the 50 end (62 nt RNA2), 17 nt from the 30 end (55 nt RNA3) or both combined (45 nt RNA4) supported accurate 3U insertion, with both Q and D fractions (Fig. A-1B, lanes 1–4 and 5–8, respectively).

Quantification of accurate 3U insertion products showed that editing with the 45 nt RNA4 was 2-fold more efficient than with the parental 72 nt RNA1 (Fig. A-1C and E). The abundance of accurately edited product for each construct

tested, was initially calculated as the percentage of total input RNA and then normalized to the abundance of the corresponding product with the most active substrate identified. Mean and error bars from two experiments were plotted on a linear scale. The 45 nt RNA4 has 30 nt (including 27 purines) upstream and 15 nt downstream of ES2. Interestingly, progressive 50 truncations of this substrate (43 nt RNA5, 40 nt RNA6, 37 nt RNA7 and 34 nt RNA8) showed quantitatively different inhibitory effects with the Q and D fractions. 3U insertion was moderately reduced with decreased template length in the Q fraction (Fig. A-1D, lanes 1–5; and Fig. A-1C). However, with the D fraction, increased truncation of the template had a more dramatic effect (Fig. A-1D, lanes 6–10; and Fig. A-1E), and no 3U insertion was detected with the shortest 34 nt template (Fig. A-1D, lane 10; and Fig. A-1E). Notably, the 45 nt RNA4 and shorter constructs are capable of forming the same upstream and downstream duplexes with gRNA and exhibit the same DG (Fig. A-1A), and the observed gradual drop in insertion parallels the extent of upstream sequence truncation.

Thus, an ~45 nt A6 substrate is optimal for insertion at ES2 in both Q and D fractions, and the significant inhibition of insertion with shorter constructs may reflect, at least in part, loss of stimulatory cis-features (e.g. purinerichness) or a more general property, such as overall length or folding. To distinguish between these possibilities we extended the 34 nt RNA8, that is inactive in the D fraction, either with an artificial upstream pyrimidine run (47 nt RNA9) or its natural downstream pyrimidine-rich sequence (51 nt RNA10). Note that the size of these

two constructs is similar to the efficient 45 nt RNA4, although their predicted structure is 3- to 4-fold more stable (see DGs in Fig. A-1A). Both pyrimidine-rich extensions reestablished editing to at least 50%of the maximal level with RNA4 (Fig. A-1F and G). This suggests that reduced purine-richness was not the cause of substrate inactivation. Consistent with this notion, a 3' polypurine extension reestablished editing to a level comparable with the pyrimidine extensions (data not shown). Interestingly, shortening of the A6 substrate caused significant reduction in the ratio of 3U to 1U insertion (Fig. A-1D). In the case of the smallest 34 nt RNA8, 1U insertion was the major product in the Q fraction (lane 5), and the only detectable product in the D fraction (lane 10). This effect is more evident in the Q fraction. We are currently addressing the mechanistic reasons for this apparent reduction in insertion fidelity. Combined, the data above indicate that for *in vitro* A6 insertion at ES2, an ~45 nt long substrate is optimal, and neither specific composition (upstream and downstream of the 34 nt RNA8 sequence) nor potential pre-mRNA folding prior to gRNA pairing appears critical. The particularly dramatic inhibitory effect of template shortening in the D fraction may be due to loss of relevant protein factors in the purified editing complexes.

We next decided to compare the effect of A6 pre-mRNA minimization between insertion and deletion. In nature, insertion or deletion occur at separate ESs. Instead of analyzing different ESs, we considered the possibility of using a single pre-mRNA site to directly compare full-round insertion and deletion, i.e. a

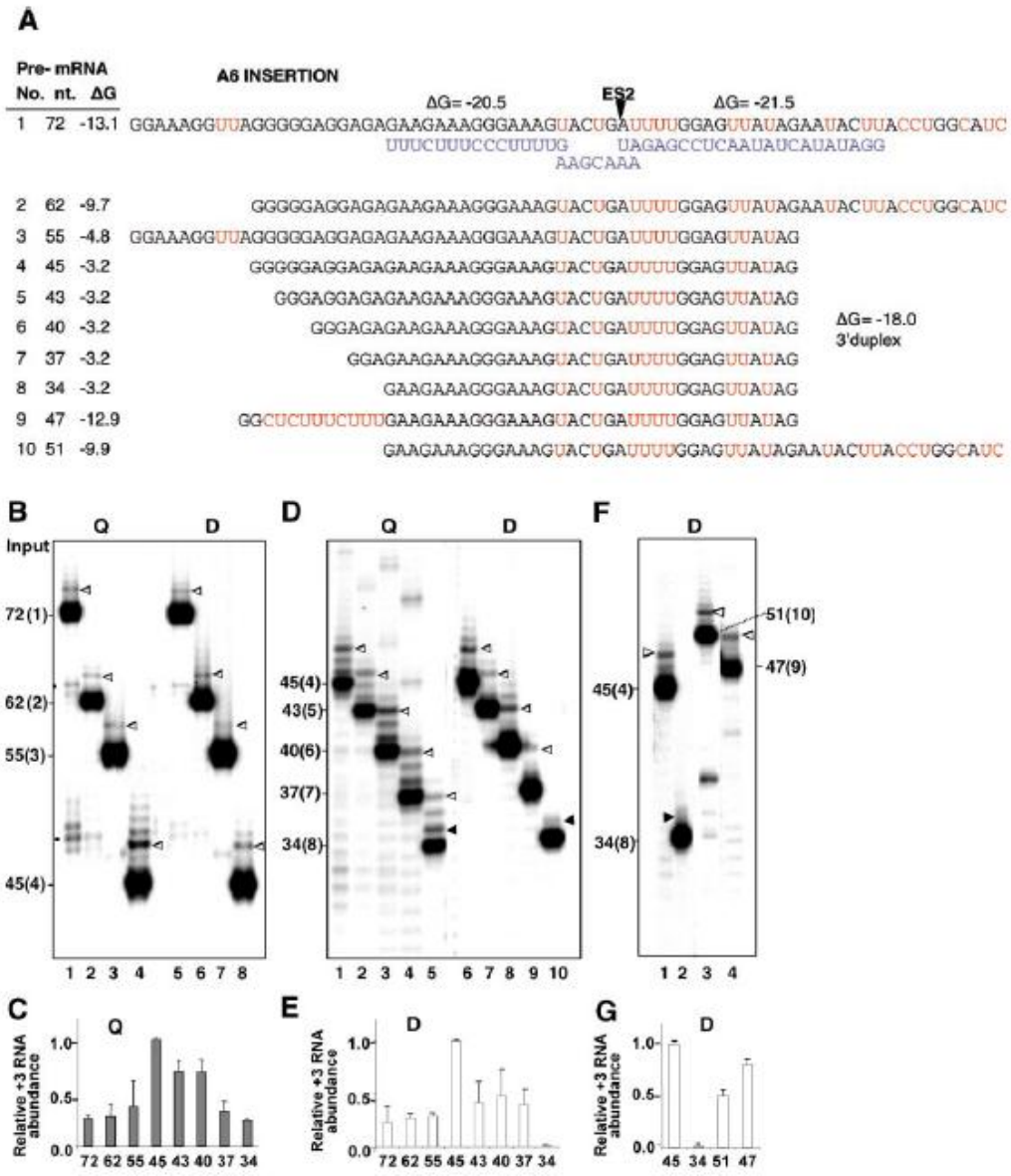


Figure A-1. Mutagenesis of A6 pre-mRNA for full-round insertion at ES2. (A) Purines are in black and pyrimidines in red. The gRNA gA6[14]USD-3A [(15); lower strand in blue] was paired with all constructs. ES2 (arrowhead) and predicted stability (kcal/mol; 3.0 Mfol, M. Zuker) of upstream and downstream duplexes are indicated. ΔG° of the downstream duplex for RNAs1, 2 and 10 is -21.5, and for RNAs 3–9 is -18.0. pre-mRNA number, size and predicted stability of the pre-mRNA alone are indicated. (B,Dand F) Full-round insertion assayed with the peak editing fractions from Q-Sepharose ('Q') or consecutive Q-Sepharose and DNA-cellulose columns ('D'). The pre-mRNA size and number (in parentheses) are indicated, edited +3U RNA (open arrowhead). +1U RNA's also indicated when it is the major product (filled arrowhead). Dots indicate sites of spontaneous RNA breakage (B and D).

site in a natural or artificially converted form (i.e. from deletion to insertion or vice versa). Previous reports have shown that A6 and CYb ESs can be converted at least at the cleavage step (414,427), and substrates for partial 'precleaved' editing efficiently support the second and third enzymatic steps (447,451), suggesting that relatively simple ES determinants may govern a complete editing cycle. Indeed, A6 (insertion to deletion) conversions have been used to establish multi-site full-round editing substrates (J. Cruz-Reyes, A. Zhelonkina and B. Sollner-Webb, unpublished data), but opposite conversions (deletion to insertion) had not been possible.

Conversion of natural ESs in both directions would enable direct comparisons of full-round insertion and deletion 50 of a common pre-mRNA/gRNA downstream 'anchor' duplex. To this end, we introduced three uridylates at ES2 in the 72 nt RNA1 (generating 75 nt RNA11, Fig. A-2A), as well as appropriate gRNA changes to preserve single-strandedness of the targeted uridylates and other residues between the duplexes, as well as weak paring of proximal pairs in the upstream duplex (392). Both RNA11 and parental RNA1 form the same upstream and downstream duplexes with gRNA, so that the two substrates differ only in the structure of the internal loop (Fig. A-2B shows the sequence changes). Fig. A-2C shows that the converted RNA11 supported accurate 3U deletion, thus enabling analysis of the same truncations in RNAs 2, 3 and 4 (compare with RNAs 12, 13 and 14; Fig. A-1A and A-2A).

Notably, deletion at both converted and natural sites was optimal in presence of mM concentrations of an adenine ribonucleotide (428) (data not shown).

We found that the above mutations similarly affected deletion in both Q and D fractions (Fig. A-2C–E). Interestingly, the 75 nt RNA11 was quite efficient for ES2 deletion, and an upstream 10 nt resection (65 nt RNA12) had a limited effect. In contrast, a downstream 17 nt resection (58 nt RNA13) inhibited deletion by 50%, and both truncations combined (48 nt RNA14) were strongly inhibitory. Notably, the predicted folding stability of the latter construct is nearly half of the 65 nt RNA12, and about one-third of the 75 nt RNA11.

Furthermore, terminal extensions of the inactive 48 nt RNA14 with (either 50 or 30) artificial pyrimidine-rich sequences (65 nt RNA15 or 65 nt RNA16) re-established efficient deletion (Fig. A-2F). Control reactions devoid of gRNA confirmed the position of the deletion products (data not shown). Dots indicate sites of spontaneous breakage typically at U positions. Together, the data in Fig. A-1 and A-2 indicate that in both Q and D fractions, the optimal A6 substrate for ES2 insertion was about 20 nt shorter than for ES2 deletion (45 nt RNA4 and 65 nt RNA12, respectively, in Fig. A-1B and A-2B). Strikingly, while the 45 nt RNA4 is the most efficient insertion substrate, the comparable 48 nt RNA14 failed to support deletion. Furthermore, our analysis suggests that distal A6 pre-mRNA composition (beyond the sequence of the shortest RNA analyzed) and folding stability may influence, but are not critical for, efficient ES2 insertion and deletion.

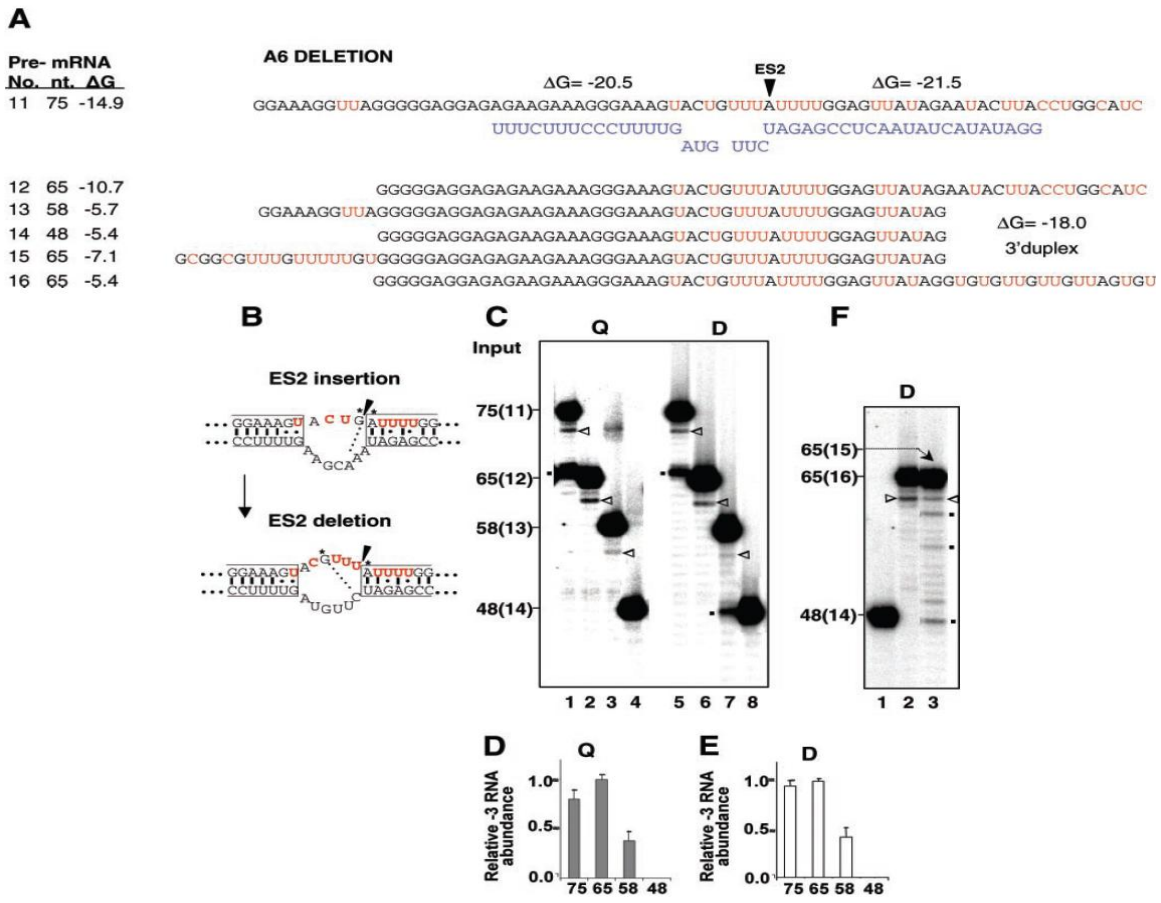


Figure A-2. Mutagenesis of A6 pre-mRNA for full-round deletion at converted ES2. (A) RNA11 is similar to RNA1 in Fig. A-1A but with three extra Us at ES2 to create a 3U deletion site. The gRNA (lower strand in blue) was modified to mimic a deletion site. All labeling is as in Fig. A-1A. (B) Diagram of parental and converted ES2 for full-round editing (the arrow indicates the direction of conversion). Watson–Crick (bar) or non-standard (dot) pairs..A potential base pair in the internal loop is depicted with a dotted line (10–12). (C and F). Full-round deletion assayed with Q or D fractions, and the indicated 3' end-labeled pre-mRNA. deletion is indicated by an arrowhead C) RNA11 to 48 nt RNA14, and (F) inactive 48 nt RNA14 with reconstituted derivatives bearing heterologous 5' or 3' terminal extensions, (D and E) Plots of relative abundance of -3U RNA.

CYb insertion and deletion editing

To determine whether the observed size differences in minimal substrates for A6 deletion and insertion are specific to this pre-mRNA transcript, or are

shared in other substrates, we analyzed the CYb model system. To this end, we started with a 66 nt pre-mRNA/gRNA pair for insertion at ES1, originally described by Igo et al. [RNA17 in Fig. A-3A; (11)]. This substrate contains the entire editing domain (including 13 ESs), which only supports insertion *in vivo*. In both Q and D fractions, a 14 nt upstream resection (52 nt RNA18) and a 21 nt resection (45 nt RNA19) caused a gradual decrease in insertion (Fig. A-3).

In contrast, a 45 nt template for A6 insertion (RNA4) was optimal in both Q and D fractions (Fig. A-1). In further 50 truncations, the 43 nt RNA20 is similarly efficient as the 45 nt RNA19, whereas the 40 nt RNA21 and 37 nt RNA22 were significantly inhibited in both Q and D fractions (Fig. A-3C and D). The shortest substrate appeared inactive in the D fraction, and only 10% active in the Q fraction. All plotted quantifications are the average of two independent experiments, each one including replica series (see material and methods).

All pre-mRNA constructs tested have similar predicted folding stability (Fig. A-3A). Thus, we observed efficient CYb insertion in the Q and D fractions with 52 nt and 66 nt substrates (RNA18 and RNA17, respectively). In contrast with CYb insertion, A6 insertion was quite efficient with a shorter 45 nt substrate (RNA4) but relatively inhibited with larger substrates (Fig. A-1B and C). This difference between the two substrates suggests that editing complexes are sensitive to transcript-specific features, in addition to absolute substrate size. *In vivo* CYb editing only involves insertion; however, it is unknown whether or not the CYb transcript is intrinsically resistant to deletion.

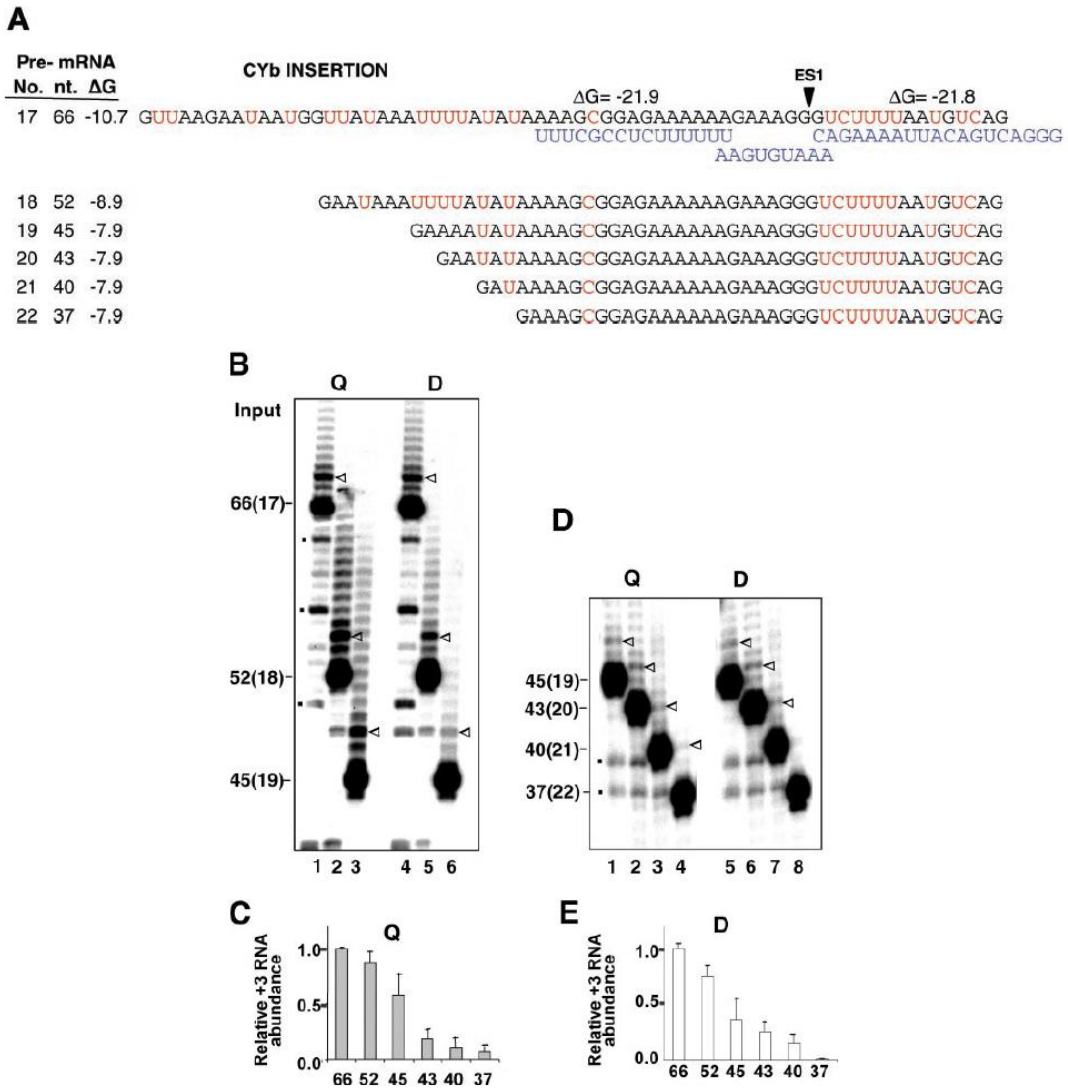


Figure A-3. Mutagenesis of CYb pre-mRNA for full-round insertion at ES1. (A) All labeling is as in Fig. A-1. (B and D) CYb insertion assays, with 3U insertion products indicated by an arrowhead. (C and E) Plots of relative accumulation of +3U RNA.

To address this question, we attempted artificial conversion of the 66 nt RNA17 by introducing uridylates at ES1 (generating 70 nt RNA23, Fig. A-4A) and corresponding gRNA sequence changes to mimic a deletion site (Fig. A-4B

details the changes). Notably, this construct supported accurate gRNA-directed 3U deletion at the converted ES1 (Fig. A-4C, lane 1); therefore, the CYb substrate is not intrinsically resistant to deletion editing. This CYb deletion construct allowed us to test the upstream resections (Fig. A-4A) previously analyzed for CYb insertion. In the Q and D fractions, 14 and 21 nt resections (56 nt RNA24 and 49 nt RNA25, respectively) have relatively small effects.

In further 50 truncations, a 44 nt RNA26, 41 nt RNA27 and 36 nt RNA28 were less efficiently processed in both protein preparations. The shortest construct did not exhibit visible editing in the D fraction (Fig. A-4E, lane 8). Thus, CYb deletion is quite efficient with a 49 nt substrate on average in both Q and D fractions. In contrast, A6 deletion was significantly inhibited with the equivalent 48 nt RNA14 in both Q and D fractions. Interestingly, CYb insertion and deletion exhibited similar profiles of editing activity relative to substrate size (compare plots in Fig. A-3 and A-4).

This could be a reflection of the natural substrate not being exposed to selection pressures associated with deletion editing. Thus, the optimal substrate for deletion differs between A6 and CYb. Such differences in deletion, together with the insertion differences described above, further suggest that pre-mRNA size and nucleotide composition affect editing efficiency in a transcript-specific manner. RPS12 insertion and deletion editing The only model substrates currently used to study trypanosome full-round RNA editing *in vitro* are A6 and CYb.

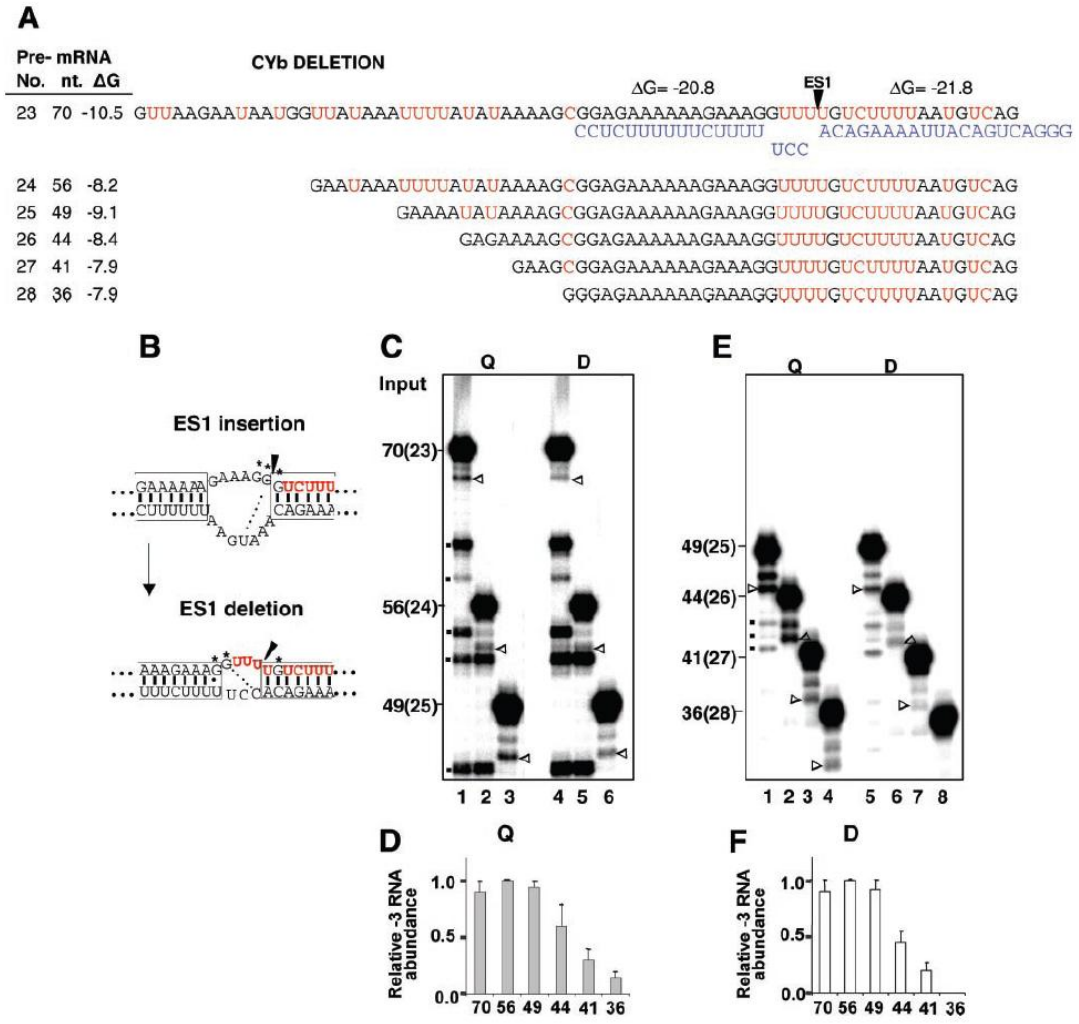


Fig. A-4. Mutagenesis of CYb pre-mRNA for full-round deletion at converted ES1. (A) Starting RNA23 derivatives. All labeling is as in Fig. A-1.(B) Parental and converted ES1 for full-round editing (C and E) CYb deletion assays. (D and F) Plots of relative accumulation of -3U RNA.

We decided to apply the current understanding of these two systems to establish an RPS12 pre-mRNA (which encodes the mitochondrial ribosomal protein, subunit 12) substrate for both insertion and deletion editing. To this end, we tested a 70 nt RNA29 sequence, as well as a longer 112 nt RNA30 and shorter 55 nt substrate (Fig. A-5A) (data not shown). These transcripts were annealed to a complementary gRNA designed to direct 3U deletion at the natural ES2. Interestingly, the largest substrate was most efficient, despite its significant predicted folding stability (Fig.A- 5A), and the shorter 70 nt RNA30 and 55 nt RNA (data not shown) exhibited corresponding decreasing efficiency in both Q and D fractions (Fig. A-5C) (data not shown).

So far we showed conversion from insertion to deletion. To determine if the opposite conversion can be performed (i.e. from deletion to insertion), we altered the natural deletion ES2 in RPS12 to mimic an insertion site, similar to the natural CYb ES1 and A6 ES2 (Fig. A-5B). Based on our A6 and CYb analysis, we tested a 43 nt construct with a complementary gRNA design directing 3U insertion. This RNA pair forms a potential ligation ‘bridge’ with residues adjoining the ES, which stimulates insertion but is dispensable for deletion (392,). Also, it favors single-strand character of residues between the duplexes, and enables weak pairing at the 30 end of the upstream duplex, as in deletion sites (427,447).

As predicted, the resulting 43 nt RNA31 supported accurate 3U insertion in both Q and D fractions (Fig. A-5D) (data not shown). Longer 51 nt and shorter 40 nt RNAs appeared slightly more and less efficient, respectively (data not shown). The absolute percent of correctly edited product with the novel RPS12 substrates reported here varied between experiments and protein preparations, but generally ranged between 5 and 10% of remaining pre-mRNA input. Thus, RPS12 substrates were established for full-round deletion and insertion. Furthermore, the conversion of a natural deletion site into an insertion site confirms that all genetic information that commit editing complexes to carry out full-round deletion or insertion resides within the ES and suggests that such information can be manipulated to reprogram the mode of the editing machinery at virtually any natural site.

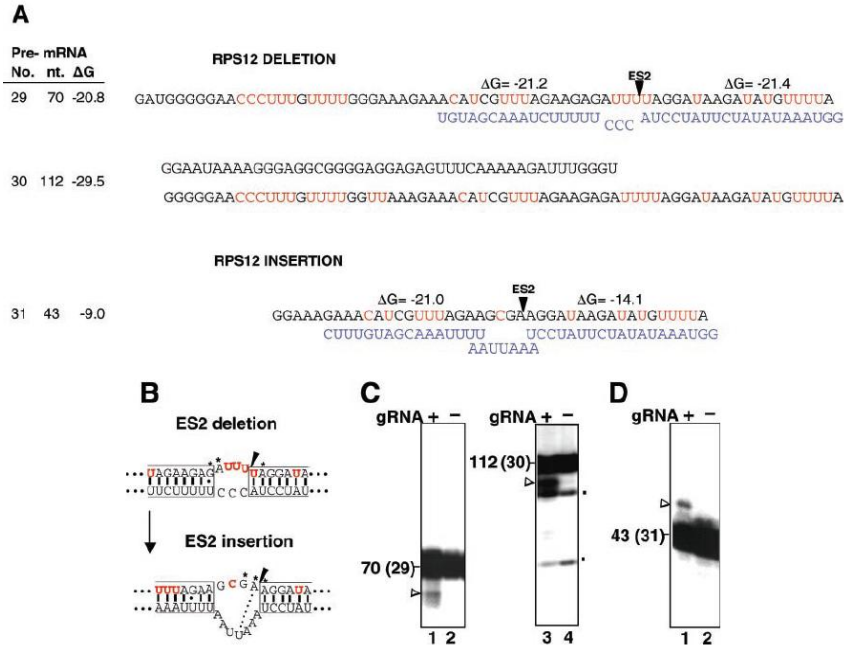


Figure A-5. Establishment of RPS12 as substrate deletion and insertion. (A) RNA29 RNA30 for deletion and 43 nt RNA31 for insertion with partially complementary gRNAs. All labeling is as in Fig. A-1. (B) Parental deletion and converted insertion ES2. (C and D) Deletion and insertion assays.

Discussion

In this study, we defined minimal pre-mRNAs substrates for full-round insertion and deletion at specific ESs, and systematically assessed the contribution of distal regions on editing efficiency. We compared editing with complexes moderately enriched by Q-Sepharose (peak editing fraction, ‘Q fraction’) or significantly enriched by consecutive Q-Sepharose and DNA-cellulose (peak fraction, ‘D fraction’) columns (393). Three different model transcript substrates were analyzed, A6, CYb and RPS12 (the latter established in this study). Interestingly, in some reactions, substrate shortening caused greater inhibition of

insertion activity in the D fraction than in the Q fraction (e.g. see Fig. A-1D and A-4E), suggesting that stringent purification of editing complexes causes partial loss or inactivation of a relevant RNA-binding activity. Furthermore, at least in the case of A6, substrate shortening negatively affected the fidelity of insertion (Fig. A-1D, compare lanes 1 and 5, or 6 and 10).

Our proposed loss of a relevant factor upon extensive purification of editing complexes may explain why some protein preparations by other laboratories support efficient pre-cleaved editing (using short substrates), but not full-round editing reactions. The best substrate size identified for each transcript analyzed was similar in both Q and D peak fractions (e.g. A6 insertion was best with a 45 nt substrate in both protein preparations); however, significant differences between deletion and insertion were observed: 45 and 65 nt A6 substrates were best for insertion and deletion, respectively (Fig. A-1 and A-2), and longer substrates were somewhat less efficient in deletion. Notably, in contrast to our optimized~45 nt insertion substrate, the corresponding 48 nt deletion substrate (with 3Us at ES2 accounting for the size difference) was strongly inhibited (Fig. A-2D and E).

In the case of CYb editing, insertion was most efficient with a 66 nt transcript (Fig. A-3) (447), whereas deletion was best with 49 and 56 nt substrates (Fig. A-4). Significantly larger CYb insertion substrates (80 and 100 nt long) are somewhat less efficient (data not shown). The molecular basis of these differences between deletion and insertion in the same substrate, and between

the two substrates, is presently unclear. They may reflect the complexity of substrate recognition due to the proposed partition of functions associated with insertion and deletion in editing complexes, and the evidently large diversity of nucleotide sequences proximal and distal to sites encountered by editing complexes. Thus, it is feasible that one or more relevant RNA–protein recognitions differ between full-round insertion and deletion.

Furthermore, such interactions appear to be significantly influenced by transcript-specific features in and around editing sites. Purine-rich regions typically found upstream of editing sites have been proposed to anneal with natural gRNA 3' (poly)U tails (386,387), and structural studies in trypanosomes have confirmed this interaction (422). Whether or not purine-rich regions or other pre-mRNA features exhibit additional cis-acting functions has not been determined in trypanosomes. In *Leishmania*, ND7 and CYb pre-mRNA substrates that are devoid of natural purine-rich regions are functional, although the presence of an uncharacterized small 5' determinant near the ND7 editing site was proposed (413,449).

Furthermore, a separate 34 nt A/U element, in CYb pre-mRNA, appears to modulate gRNA-directed, and induces gRNA-independent, insertion (30). In our current study, inactivating resections of the A6 upstream purine-rich region were rescued by pyrimidine-rich extensions. Similarly, inactivating downstream resections can be rescued by heterologous sequences. This indicates that the specific sequence and purine-richness of the truncated A6 regions have little if

any modulatory functions. Whether more proximal 5' and 3' sequences (not dissected in this study) bear specific determinants remains to be analyzed. Our *in vitro* analysis of CYb has not revealed yet the presence of critical cis-elements in substrates up to 100 nt long (this work and data not shown), but further studies may be necessary. Based on our observations with the A6 and CYb systems, we established RPS12 substrates for deletion and converted insertion at ES2.

Interestingly, an ~100 nt long RSP12 substrate supports the best deletion level thus far, and shorter substrates (70 nt and 55 nt long; Fig. A-5C and data not shown) are correspondingly less efficient. Similar to the A6 system, we obtained efficient insertion with shorter substrates (~43–50 nt long, Fig. A-5C) (data not shown) than those needed for optimal deletion. We suspect that efficient model substrates of transcripts naturally undergoing both types of editing and deletion may be shorter for insertion than for deletion. CYb may be an exception as this transcript has not been exposed to natural pressures imposed by deletion editing.

Finally, the effect of pre-mRNA minimization on full-round deletion and insertion was directly compared at natural and artificially converted sites in three different substrates. Such conversions were performed in both directions: from deletion to insertion and from insertion to deletion, thus confirming expectations that the type of editing is exclusively determined by mRNA and gRNA residues within the ES (414,426). Importantly, both natural and converted sites for full-

round deletion studied here exhibited the reported [mM] adenine ribonucleotide requirement for optimal cleavage. In contrast, both natural and converted insertion sites do not require such [mM] adenine ribonucleotide supplement (19). Thus, both the type of editing and associated nucleotide requirements were converted. Sites for partial (pre-cleaved) deletion and insertion editing have also been established on significantly altered pre-mRNA sequences (15,24). Site conversion within the CYb editing domain appears surprising as this substrate is only known to undergo insertion *in vivo*. If the CYb editing domain is not intrinsically refractory to U deletion, inserted or genomically encoded Us may be occasionally removed by 'misguiding' (31), or by proofreading cycles (to eliminate extra U insertions) with cognate gRNAs. It is also conceivable that the natural substrate (1151 bases) uses specialized features, and/or CYb-specific binding factors to downregulate deletion functions of bound complexes.

Materials and methods

Pre-mRNA and gRNA substrates

The ATPase 6 (A6) and CYb (CYb) pre-mRNA substrates used for full-round insertion at ES2 and ES1, respectively, are derived from the A6AC and CYb anchor substrates, and enhanced gRNAs, originally described by Igo et al. (11,15). gCYb[558]USD was modified to direct +3U insertion, and all A6 mRNA constructs used pre-edited ES1 to increase the stability of the downstream duplex. All RNAs were transcribed using the Uhlenbeck single-stranded T7

transcription method (25) and gel purified. Converted ESs in A6, CYb and RPS12 were engineered on natural pre-mRNA sequences paired with enhanced gRNAs (427,444,447).

Preparation of crude mitochondrial extracts and enrichment of editing complexes

Procyclic Trypanosoma brucei strain TREU667 was grown in Cunningham media, and mitochondrial extracts were prepared as described previously (414,434). Mitochondrial crude extract purified by Q-Sepharose ion exchange chromatography, or further purified by DNA-cellulose ion exchange chromatography to seven-major silver-stained components of editing complexes, was described previously (393). The elution fractions with the peak of editing activity were used in each case as indicated in the text (Q fractions and D fractions). The differences in activity by Q and D fractions described in this study were reproduced. Editing and quantification analysis Full-round U deletion and U insertion were performed as described by Cruz-Reyes et al. (19). Briefly, 30 end-labeled pre-mRNA (10 fmol) and gRNA (1.25 pmol) were preannealed and incubated in 10 mM-MRB buffer [10 Mm Mg(OAc)2, 10 mM KCl, 1 mM EDTA, pH 8, 25 mM Tris– HCl, pH 8 and 5% glycerol]. Insertion reactions were supplemented with 150 mM UTP and 3 mM ATP. Deletion reactions were supplemented with 3 mM AMP-CP and 30 mM ATP. Aliquots containing 20 ml editing reactions, including 2 ml of peak editing fraction, were incubated at 26 C

for 60 min. Upon incubation, the RNA was deproteinized and resolved in 9% acrylamide, 8 M urea gels. Data were visualized by phosphorimaging and/or X-ray autoradiography. Quantification of editing products was performed using a STORM PhosphorImager (Image Quant 5.0, Molecular Dynamics). Each panel in these figures corresponds to one of two replica series performed simultaneously (i.e. one experiment). At least two independent experiments were performed for each figure and the data shown is representative. The editing activity varied between protein preparations, but the relative efficiency of the constructs was consistently observed. The abundance of accurately edited product for each construct tested was initially calculated as the percentage of total input RNA and then normalized to the abundance of the corresponding product with the most active substrate identified. Mean and error bars were plotted on a linear scale.

APPENDIX II

RNA EDITING COMPLEX INTERACTIONS WITH A SITE FOR FULL-ROUND U DELETION IN *TRYPANOSOMA BRUCEI**

Summary

Trypanosome U-insertion and U-deletion RNA editing of mitochondrial pre-mRNAs is catalyzed by multisubunit editing complexes as directed by partially complementary guide RNAs. The basic enzymatic activities and protein composition of these high-molecular mass complexes have been under intense study, but their specific protein interactions with functional pre-mRNA/gRNA substrates remains unknown.

We show that editing complexes purified through extensive ion-exchange chromatography and immunoprecipitation make specific cross-linking interactions with A6 pre-mRNA containing a single ^{32}P and photoreactive 4-thioU at the scissile bond of a functional site for full-round U deletion. At least four direct protein–RNA contacts are detected at this site by cross-linking. All four interactions are stimulated by unpaired residues just 5' of the pre-mRNA/gRNA anchor duplex, but strongly inhibited by pairing of the editing site region.

---

*Reprinted with permission from: "RNA editing complex interactions with a site for full-round U deletion in Trypanosoma brucei" by Anastasia Sacharidou, Catherine Cifuentes-Rojas, Kari Halbig, Alfredo Hernandez, Lawrence J. Dangott, Monica De Nova-Ocampo, and Jorge Cruz-Reyes. 2006. RNA. July; Vol. 12. No. 7: 1219–1228.

Furthermore, competition analysis with homologous and heterologous transcripts suggests preferential contacts of the editing complex with the mRNA/gRNA duplex substrate. This apparent structural selectivity suggests that the RNA–protein interactions we observe may be involved in recognition of editing sites and/or catalysis in assembled complexes.

Introduction

Mitochondrial mRNAs in trypanosomatid protozoa including *Trypanosoma, Leishmania*, and *Crithidia* species undergo a unique form of RNA editing by cycles of uridylate insertion or deletion at numerous editing sites (ESs). This post-transcriptional mRNA maturation progresses with a general 3'–5' polarity and is catalyzed by a large multisubunit RNA editing complex (also termed 20S editosome or L-complex) proposed to contain between 8 and 20 polypeptides depending on the purification protocol (393-397). The smaller number presumably reflects high-stringency purification conditions and tight association of the subunits in the resulting complexes.

Partially complementary guide RNA (gRNA) transcripts direct this process, which is believed to initiate with the formation of an ‘‘anchor duplex’’ with pre-mRNA. Catalysis of a single editing cycle involves three basic activities, namely, mRNA endonuclease, 3' terminal uridylyltransferase (Tutase in insertion) or 3' to 5' U-specific exoribonuclease (in deletion), and RNA ligase. So far, the known catalytic subunits in the editing complex include a TUTase

(KRET2, also termed LC-6b; (396, 398), a U-specific exonuclease (KREP6, LC-2; (399), two RNA ligases (KREL1, Band IV, LC-7a and KREL2, band V, LC-9; (400-402), deletion and insertion endonucleases (KREN1 and KREN2; (403-404), respectively), and an endonuclease/exonuclease (KREPA3, band VI, LC-7b; (405)). All these protein subunits have been cloned and characterized *in vitro* and *in vivo*. A significant amount of information has been obtained on the structural and functional composition of editing complexes (for reviews, see (383-385); however, the specific RNA–protein interactions in assembled complexes during recognition of pre-mRNA/gRNA duplex substrates and catalysis of full editing cycles are unknown.

Several reported protein subunits contain conserved motifs for nucleic acid binding, but only a purified recombinant KREPA3 has been shown to exhibit RNA-binding activity (405). In addition to core essential subunits, a few auxiliary components involved in editing are known, including the annealing factors MRP1 (gBP21) and MRP2 (gBP25) (390-391, 406-407), and the gRNA-binding factor RBP16 (408). Other proposed factors are an RNA helicase, REAP1, and TbRGG1 (395, 409-411). All factors mentioned above are either weakly or not associated with editing complexes and dispensable for *in vitro* editing (393, 395-396, 412).

Here, using photocross-linking we report four protein interactions in intimate contact with the first editing site (ES1) for full-round U deletion in an A6 pre-mRNA/gRNA substrate that co-purify and co-immunoprecipitate with editing

complexes. All four RNA–protein cross-links exhibit structural selectivity for the single-stranded character of the editing site region. Together, the data indicate that the cross-linking events described here are mediated by one or more stably bound core subunits. To our knowledge, this is the first report of specific RNA–protein interactions of editing complexes with a functional site for full-round RNA editing.

Results

To search for RNA–protein interactions in assembled RNA editing complexes, we generated a 72-nt A6 pre-mRNA substrate containing a single ^{32}P and 4-thioU at the scissile bond of the first editing site (ES1) for U deletion (Fig. A-6A). Prior to photo-crosslinking, this thiolated pre-mRNA was pre-annealed with gRNA and mixed with editing complex preparations, as in standard *in vitro* reactions (see Materials and Methods). Importantly, the thiolated pre-mRNA supports accurate *in vitro* deletion of three uridylates as directed by the partially complementary gRNA D33 (392), although slightly less efficiently than unmodified pre-mRNA (Fig. A-6A, B). This indicates that the presence of a thio-uridylate immediately 3' of the scissile bond does not significantly interfere with editing activity.

We initially utilized crude mitochondrial lysate that was fractionated by Q-sepharose chromatography (Q1 column; Fig. A-6C) to detect protein cross-links to ES1. This column has been previously used to enrich active editing

complexes (393-394, 413). Several cross-links of various intensities are evident across the fractionated lysate upon irradiation with 365-nm UV light. At least four of them, at about 40, 50, 60, and 100 kDa, appeared to closely co-purify with editing complexes as detected by immunoblots of known core subunits, particularly in the peak fractions 9–11 (‘‘peak Q1 fractions’’; Fig. A-6C,D).

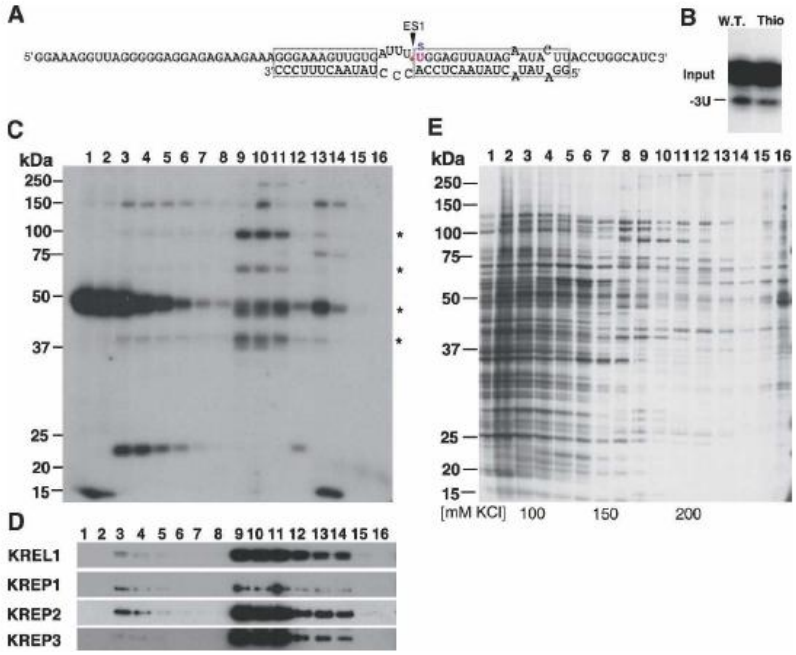


FIGURE A-6. RNA–protein interactions detected by photocross-linking co-purify with RECC in Q-sepharose fractionated mitochondrial extract. (A) Diagram of the 72-nt A6 pre-mRNA substrate annealed with 33-nt gRNA D33. Boxes indicate predicted duplexes flanking ES1. The 4-thioU (sU) and 5’-^{32}P-radiolabed bond (*) are positioned at the double-strand/single-strand junction that defines ES1 (arrowhead). (B) U-deletion *in vitro* assay of unmodified (W.T.) and thiolated A6 pre-mRNA paired with gRNA D33. Input and accurate -3U deletion RNAs are indicated. (C) UV irradiation (365 nm) of the pre-mRNA/gRNA substrate with Q-sepharose fractions. The asterisks indicate the positions of four proteinase K-sensitive cross-links that co-purify with RECCs. RECC was detected in immunoblots (D) of four known subunits KREL1 (also termed TbMP52), KREP1 (MP8), KREPA2 (MP63), and KREPA3 (TbMP42). RECC and co-purifying crosslinks peak in fractions 9–11. The molecular size is indicated in kilodaltons (E) Silver staining of Q1-sepharose fractions. The peak Q1 fractions elute between 150 and 200 mM KCl.

Other prominent cross-links were detected at about 75, 150, and 250 kDa in or near these fractions. Proteinase K inactivation of all cross-links in the peak Q1 fractions showed that they are protein dependent (not shown), so from here onward we will refer to them as p40, p50, p60, and p100. The peak Q1 fractions eluted away, between 150 and 200 mM KCl, from most proteins in the mitochondrial crude extract, and therefore appear significantly enriched (Fig. A-6E). These fractions were pooled and further purified by two subsequent steps of ion-exchange chromatography in DNA-cellulose and Q-sepharose columns, respectively (Fig. A-7; data not shown).

Notably, p40, p50, p60, and p100 co-purify with editing activity in both columns. However, additional bands are detected in the DNA-cellulose column (‘‘D’’) peak fractions, although not reproducibly in our protein preparations (not shown). The peak fractions of the second Q-sepharose column (‘‘Q2’’ fractions 13–15) show primarily p40, p50, p60, and p100 (Fig. A-7A), precisely co-purifying with isolated silver-stained polypeptides and full-round deletion activity (Fig. A-7B, C). Notably, our peak Q2 fractions exhibit a pattern of major stained protein bands, plus a few additional fainter bands (Fig. A-7B), that is remarkably similar to that of editing complexes purified with either the same protocol (414) or another biochemical purification strategy (394).

Both the same protein pattern and relative intensity of individual bands are conserved whether silver or SYPRO Ruby staining is used (data not shown).

Importantly, p40, p50, p60, and p100 co-localize with stained bands in the Q2 peak fractions (Fig. A-7D). Furthermore, these cross-links are only detected if the targeted residue is thiolated and, therefore, upon 365-nm but not 260-nm UV light irradiation (data not shown).

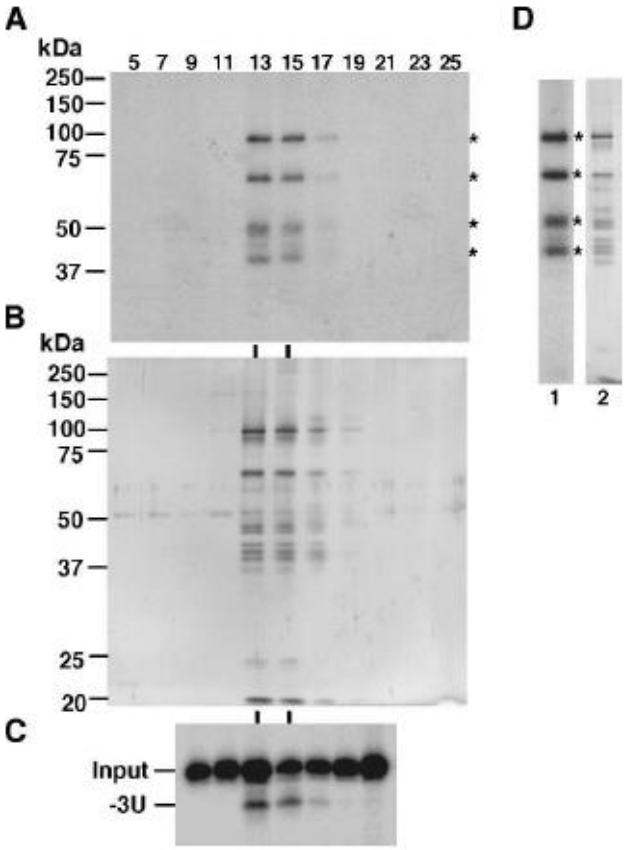


FIGURE A-7. p40, p50, p60, and p100 co-purify with RECC after extensive ion-exchange chromatography. Peak Q1 fractions from Fig. A-6 were fractionated on DNA-cellulose (D) and a second Q-sepharose column (Q2). Shown are the relevant odd fractions of the Q2 elution. The four protein–RNA cross-links (A) precisely co-purify with silver-stained RECCs (B) and U-deletion activity (C). The U-deletion activity of Q2 fractions was assayed at the ES1 of the 3'-end-labeled A6 pre-mRNA. The RNA input and accurate deletion product (-3U) are indicated. (D) The four protein–RNA cross-links (lane 1) co-localize with silver-stained protein components (lane 2) of the peak Q2 fractions (no. 13–15).

A direct comparison of the protein content and cross-linking pattern of Q1, D, and Q2 peak fractions (Fig. A-8A,B) indicates that the four RNA–protein interactions described above are conserved throughout the purification of active editing complexes (Fig. A-8C) and most likely involve the same proteins. Other cross-links previously observed in Q1 and occasionally in D fractions are significantly reduced or lost in Q2 fractions. The peak Q2 fractions (13–15) contain ~1/6,000 of the original crude mitochondrial extract protein and exhibit a simpler protein pattern than the parental D and Q1 fractions. This extent of purification is consistent with others reported using similar protocols (393-394, 415). There is at least a ~10-fold further purification compared to the whole-cell protein content; however, the specific activity of editing complexes could not be calculated since the *in vitro* editing assay is not linear with protein added, particularly in cruder fractions (393-394, 415); data not shown). Together, these data suggest that p40, p50, p60, and p100 are tightly associated with purified active editing complexes and that they make intimate contacts with the targeted editing site.

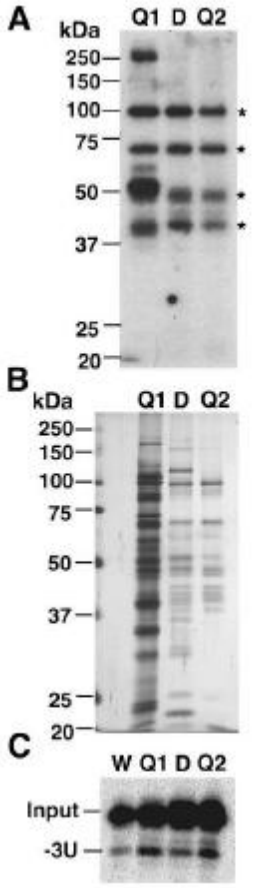


FIGURE A-8. Side-by-side gel analyses of Q1, D, and Q2 peak fractions. (A) protein–RNA cross-linking interactions, (B) silver staining, and (C) full-round U-deletion activity. The latter includes a lane with the original whole mitochondrial extract (W).

To further confirm this association, we performed co-immunoprecipitation assays (co-IP) using monoclonal antibodies that are known to immunoprecipitate active editing complexes (394, 416).

Analysis of the peak Q1 fraction shows efficient co-IP of the p40, p50, p60, and p100 kDa cross-links by anti-KREPA3 antibodies (Fig. A-9). Relative to a control lane showing the starting cross-linked sample (‘‘C’’), the unbound lane (‘‘U’’) shows a significant decrease in three cross-links, p40, p60, and p100, and their corresponding enrichment in the bound material (‘‘B’’) after two washes (‘‘W2’’). Most cross-linking activity at ~50 kDa remains in the unbound fraction, but a significant amount (above background levels) co-IPs with the editing complex, as compared with a mock assay with no antibodies.

We interpret this as indicative of at least two proteins comigrating at ~50 kDa, one corresponding to a stably bound component (p50) of editing complexes and another representing a mitochondrial protein that is presumably abundant but not tightly associated with editing complexes. Consistent with this notion, the latter cross-link may account for the prominent ~50-kDa band in the flow through and first few fractions in the initial chromatographic step (Fig. A-6C), and apparent trailing into the peak editing fractions. The same cross-linking protein is significantly reduced or lost in the D and Q2 peak fractions (Fig. A-8A), and in most gels, it appears to migrate slightly above the proposed p50 cross-link (e.g., Figs. A-6C, A-8A, A-9).

Co-IP assays were also performed with antibodies against two other editing subunits, KREPA2 and KREL1, and in cases, p40, p50, p60, and p100 selectively immunoprecipitate with editing complexes (not shown). Additional analyses were performed to confirm the specificity of the p40–100 interaction with editing complexes. These include a positive control showing efficient co-IP of radiolabeled RNA ligase subunits (via ^{32}P-adenylylation; (394, 416)) and a negative control with a nonrelated antibody (not shown). The virtual absence of the ~150- and ~250-kDa cross-links in Q2 fractions (Fig. A-7A) and their reduction to near background levels in co-IP assays (Fig. A-9) suggest that the cross-linking proteins are either weakly or not bound to editing complexes.

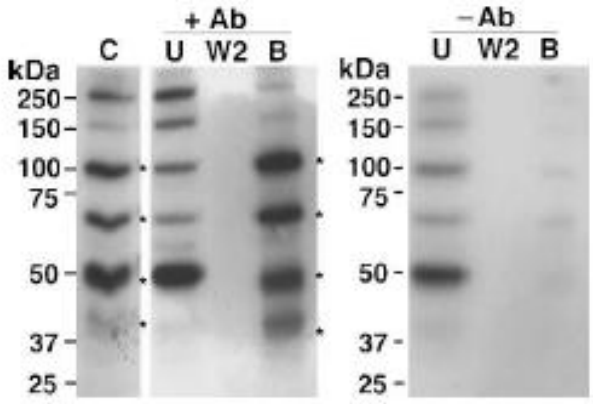


FIGURE A-9. p40, p50, p60, and p100 co-immunoprecipitate with RECCs. Protein–RNA cross-links in a peak Q1 fraction before (C lane) and after a co-IP assay with anti-MP42 antibodies (+Ab), including the unbound (U), second wash (W2), and bound immunoprecipitated (B) fractions. A parallel mock co-IP assay with no antibodies (-Ab) is shown.

Combined, our extensive chromatography purification and immunoprecipitation analyses show at least four RNA–proteins cross-links between one or more stably bound subunits of editing complexes and a site for full-round deletion in an A6 substrate. Notably, these cross-links specifically target the [^{32}P]-labeled photoreactive 4-thioU positioned at the scissile bond of this functional substrate.

To determine whether or not the polypeptides that bind ES1 also contact other positions of the A6 pre-mRNA/gRNA substrate, we moved the [^{32}P]-labeled photo-reactive 4-thioU a few nucleotides away from the scissile bond at ES1 (bond 45; Fig. A-10A). In one case, we tested the upstream bond 34 that corresponds to the second deletion site (ES4) in the natural A6 substrate, and in another, the downstream bond 51 in the never-edited region of this transcript. Both positions are located within the predicted upstream and downstream duplexes formed by the partially complementary gRNA D33, respectively (Fig. A-10A, top and middle RNA pairs). Notably, all four protein–RNA interactions detected by cross-linking at functional ES1 (bond 45) are absent at either duplex position (Fig. A-10B). This suggests that the observed RNA–protein cross-linking interactions may exhibit structural selectivity for single-strandedness of the editing site. To confirm this apparent preference for single-stranded residues adjoining the photo-reactive 4-thioU, we annealed the pre-mRNA to a gRNA derivative (31.dx) that extends the upstream and downstream duplexes into a single contiguous duplex (Fig. A-8A, bottom pair). We found that base-

pairing of the ES1 region with 31.dx strongly inhibits all cross-links observed with the parental gRNA D33 (Fig. A-10C).

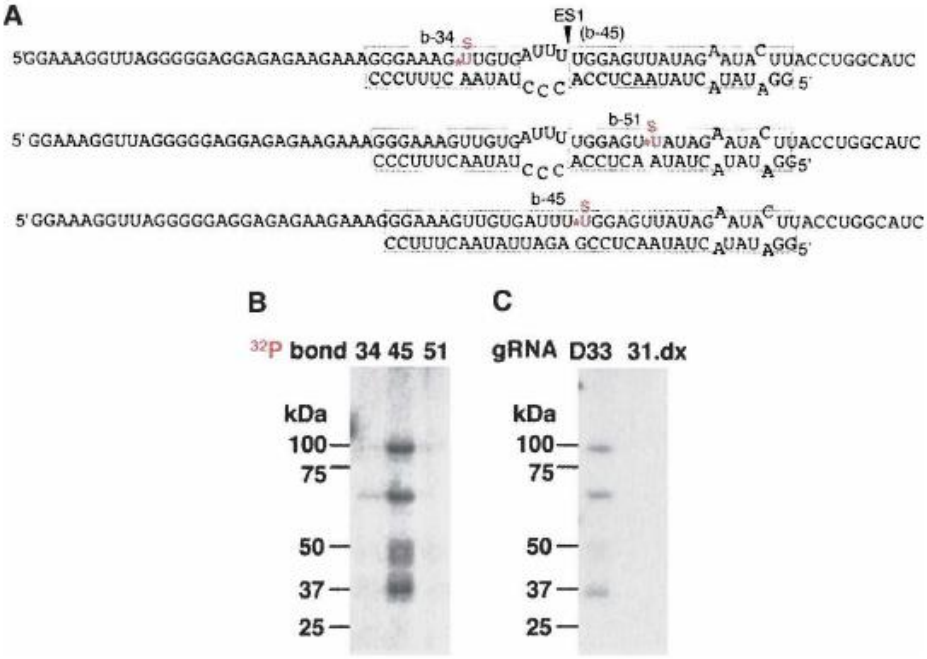


FIGURE A-10. All four RNA–protein interactions detected by cross-linking in Pf editing complexes are favored by single-strandedness at the editing site. (A) Diagrams of A6 pre-mRNA/ D33 pairs as in Fig. A-6A, but with the [5'- 32P] thiolated U at upstream (b-34) or downstream (b-51) bonds (top and middle RNA pairs, respectively). The position of ES1 (b-45) is also indicated. The A6 pre-mRNA modified at b-45 was also paired to a gRNA D33 derivative (31-dx) that forms a continuous duplex across ES1 (bottom pair). (B) Parallel crosslinking assays in a Q2 peak fraction of radiolabeled pre-mRNA at each of three indicated bonds above, paired with gRNA D33. (C) Cross-links of pre-mRNA modified at b-45 and annealed with either D33 or a D33-like derivative (31-dx) that fully base-pairs the ES1 and directs no deletion.

Together, our data indicate that all four cross-linking proteins observed at ES1 are favored by the single-strand character of the editing site. Importantly, precise gRNA base-pairing across ES1 inhibits *in vitro* U deletion at this site (389).

To assess the specificity of the interaction between editing complexes and A6 pre-mRNA/D33 substrate, we supplemented the cross-linking assay with a molar excess of various non-radiolabeled RNA competitors (Fig. A-11A–C). Interestingly, addition of 10- and 25-fold excess (relative to radiolabeled A6 pre-mRNA) of the homologous A6 pre-mRNA virtually abolished all cross-linking (Fig. A-11A, lanes 1–3), whereas another pre-mRNA (CYb; lanes 4–5) and tRNA (lanes 6–7) were only slightly inhibitory at the same concentration. The partial effect of the latter heterologous competitors seems specific to these transcripts, as further addition (25-fold) of gRNA D33 did not affect the crosslinking efficiency (lanes 8–9). Note that the assay includes pre-mRNA ((392); see Methods section). The inhibition by the A6 pre-mRNA competitor is consistent with its ability to base-pair with gRNA D33. Additional heterologous transcripts including the noncomplementary gRNA gRPS12, viral RNA H121 (25- to 50-fold excess), and several homopolymers (100-fold excess) were slightly or not inhibitory (Fig. A-11B, C; data not shown). Up to 100-fold further addition of gRNA D33 (i.e., ~200-fold excess overall) in the latter assays was not inhibitory (Fig. A-11B, lanes 5, 6).

We also tested the above RNA competitors on full-round U deletion. As expected, the homologous pre-mRNA was fully inhibitory at 25-fold excess, whereas all other competitors in Fig. A-11A–C were little or not inhibitory at the same concentration (Fig. A-11D; data not shown). Combined, the above competition analyses on cross-linking and editing assays suggest that editing complexes may be able to distinguish the pre-mRNA/gRNA duplex from individual substrate strands and from nonrelated structured or relatively nonstructured transcripts. Additional studies are currently under way in our laboratory to further address this question.

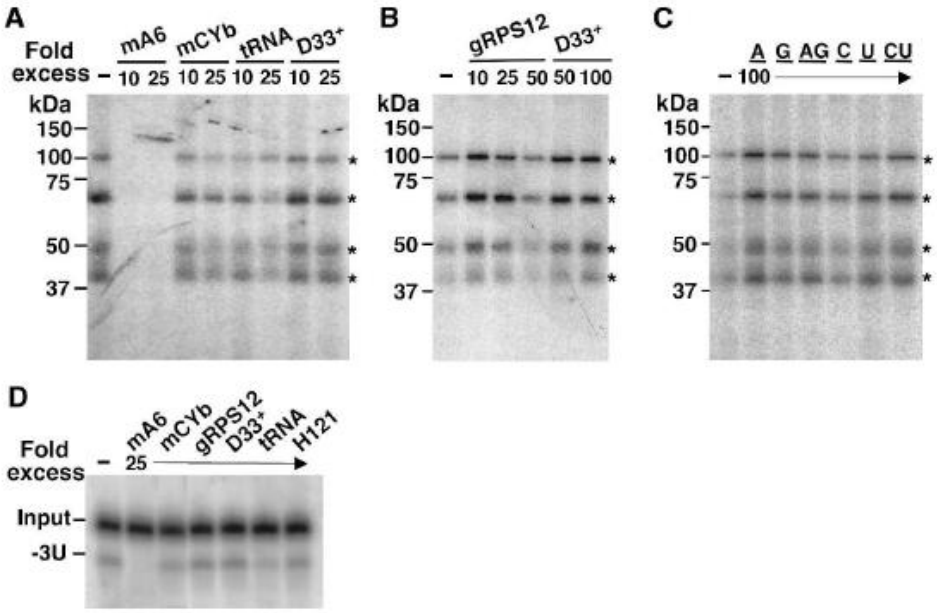


FIGURE A-11. Homologous and heterologous RNA competitors in cross-linking and editing assays. Cross-linking with or without (A) 10 and 25-molar excess of homologous A6 pre-mRNA (mA6) or heterologous CYb pre-mRNA (mCYb) and tRNA, or complementary gRNA D33. (+) Additional D33 over the standard amount (~100-fold excess) present in the cross-linking assay. (B) Ten-, 25-, and 50-fold excess of noncomplementary gRNA gRPS12 or 50- and 100-fold excess of complementary gRNA D33 (over its standard level in the assay, as in A). (C) Hundred-fold excess of 15-nt oligomers. (D) U-deletion assay with or without 25-fold excess of the indicated transcripts (~125-fold overall in the case of gRNA D33).

Based on the observed gel mobility of p40, p50, p60, and p100, we suspected that one or more of them could correspond to known subunits of editing complexes. To test this possibility, we transferred the reactions to a membrane after cross-linking and performed Western analysis using available monoclonal antibodies to identify the co-localizing proteins. Our initial analysis showed a precise co-localization between p60 and KREPA2 (~60 kDa; band III; LC-4), whereas p40 did not precisely match with KREPA3 (~40 kDa; band VI; LC-7b) (Fig. A-12). Furthermore, p40 and p50 do not comigrate with the editing RNA ligases (^{32}P-labeled by adenylylation; (417); data not shown). MS analyses of the protein bands matching the crosslinks are under way, but due to the possibility of cross-contamination between similar-size subunits (particularly in the ~90–100 kDa and ~40–55 kDa size ranges; (383)) additional work using epitope-tagging of *Candida*te subunits will be required to establish definite subunit assignments for p40, p50, and p100, and confirm that p60 corresponds to KREPA2.

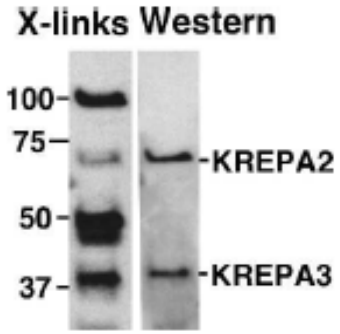


FIGURE A-12. p60 colocalizes with the KREPA2 subunit. A cross-linking reaction (X-links lane) and subsequent Western blot analysis of the same gel (Western lane) with anti-KREPA2 and anti-KREPA3 antibodies.

Overall, the extensive biochemical co-purification and co-immunoprecipitation of p40, p50, p60, and p100 with active editing complexes indicates that the cross-links involve one or more stably bound components of editing complexes. Moreover, our analysis of substrate features and response to RNA competitors suggests that editing complexes and possibly these particular RNA–protein interactions exhibit structural selectivity for the editing substrate used in our studies.

Discussion

The specific RNA–protein interactions in editing complexes that lead to their activation and catalysis of faithful RNA editing cycles in trypanosomes are unknown. The purpose of this study was to identify specific pre-mRNA/protein contacts using assembled editing complexes and an A6 pre-mRNA/gRNA substrate for full-round editing *in vitro*.

We found at least four protein interactions, p40, p50, p60, and p100, in direct contact with ES1 for U deletion. These interactions revealed by protein–RNA cross-linking involve one or more tightly bound subunits of editing complexes since they precisely co-purify with editing activity upon extensive ion-exchange chromatography in three consecutive columns and co-IP using monoclonal antibodies raised against known editing complex subunits. The ion-exchange chromatography (414) and immunoprecipitation (394, 402, 416) approaches applied in this study were previously exploited to efficiently purify active editing complexes and study their protein composition. All major protein components of the complexes originally observed by (393) are also present in the complexes prepared by immunoprecipitation and similar chromatography or affinity purifications (394-396, 416).

The identification of the cross-linking polypeptides reported is evidently necessary to begin dissecting their potential in editing. The protein banding pattern of our purified editing complexes is remarkably similar to others previously reported using related biochemical purification schemes (393-395),

and associations between specific subunits and protein bands in those patterns have been proposed (for reviews, see (383, 385)). Based on the co-localization of p60 with band III (KREPA2; LC-4) in both silver-stained gels (Fig. A-7D) and immunoblots (Fig. A-12; data not shown) we speculate that p60 may indeed correspond to band III. The precise molecular function of this subunit has not been defined, but it has been found associated with KREPC2 and KREL1 in a purified subcomplex that catalyzes partial (precleaved) deletion editing (418). These authors have speculated that KREPA2 could use its potentially regulatory OB-fold to coordinate the sequential enzymatic steps of U deletion. Furthermore, this subunit has also been proposed to play a critical structural role in the formation or stability of entire editing complexes (419-420). Other reported subunits of predicted molecular size similar to p60, although not found during the peptide sequencing of band III (by Edman degradation; (419)), include KREN2 and KREPB2, an essential insertion-specific endonuclease and a potential endonuclease, respectively (403-404). At least the essential KREN2 is expected in our purified complexes, either migrating with band III (possibly at sub-stoichiometric levels) or near to it. Another reported subunit, KRET2, appeared to be sub-stoichiometric (397) in similarly purified complexes.

p100 precisely colocalizes with the prominent band I (393), which corresponds to an (~99 kDa) exonuclease proposed to function in U deletion (KREPC2; LC-3; (383, 385). However, we cannot exclude the possibility that p100 may be the closely migrating KREN1, an essential U deletion-specific

endonuclease (395) expected in our purified active complexes, or alternatively KREPC1 (~100 kDa), a *Candida*te editing exonuclease (395) potentially present in our preparation. Any of the above likely p100 *Candida*tes is consistent with our search for subunits that bind and cross-link a deletion site.

Several known editing complex subunits could account for the p40 and p50 cross-links we observe (383, 385), including five (~41- to 49-kDa) subunits with a conserved U1-like Zn-finger domain potentially involved in macromolecular interactions with RNA substrates or other proteins in the complex. Two of these proteins also exhibit a C-terminal Pumilio RNA-binding domain and less conserved RNase III motifs potentially involved in endonuclease cleavage. Our Western blot analysis revealed that p40 is not KREPA3 (~42 kDa; Fig. A-12). Moreover, the RNA ligases KREL1 (~52 kDa) and KREL2 (~45 kDa) migrate between the p40 and p50 cross-links in high-resolution acrylamide gels and therefore are different proteins (not shown). It is also conceivable that one or more of these proteins, p40, p50, and/or p100, correspond to novel subunits of editing complexes. Further work is under way to identify these proteins and their potential roles in deletion.

KREPA3 (~42-kDa subunit) and five related subunits exhibit apparent Zn-finger domains and/or an OB fold. The former are found in many regulatory proteins and could mediate interactions with nucleic acids or with other proteins, whereas the latter typically provides a nonspecific binding platform for single- and double-stranded nucleic acids (421). KREPA3 is the only subunit known so

far to bind RNA (405). Surprisingly, a recombinant version of this protein was reported to exhibit endonuclease and (3'–5') exonuclease activities on a stretch of unpaired uridylates in a partial RNA hybrid, although KREPA3 lacks recognizable nuclease domains. While these activities are editing-like, the substrate used in that study is not functional, and the proposed protein–RNA interaction remains to be confirmed in assembled editing complexes. RNAi knockdown of KREPA3 does not appreciably disassemble editing complexes, but reduces *in vivo* and *in vitro* editing (405). Thus, the reported properties of rKREPA3 suggest that this subunit has important roles in editing. Whether or not KREPA3 is functionally similar or even redundant to any structurally related subunit remains to be determined. KREPA3 was not detected in our analysis at ES1, however this may reflect a limitation of our ‘‘zero-distance’’ cross-linking approach. That is, even if a protein specifically binds the targeted site, the thiolated uridylate and adjacent amino acid side chain may not be properly orientated with each other for efficient photoreaction.

A double-strand/single-strand junction just 5' of the downstream ‘‘anchor’’ duplex is a critical feature of functional editing sites (388-389). Interestingly, the cross-links we observe are strongly inhibited by gRNA base-pairing of the editing site (Fig. A-10). This observation suggests that the p40–100 interactions with the substrate exhibit structural selectivity for the mismatched preedited ES1, but are inhibited by gRNA complementarity across the edited site. In addition to simple mRNA/gRNA mismatches at editing sites, structural studies have

indicated that other features of functional pre-mRNA/gRNA pairs may determine the basis for endonuclease recognition (422-423). Nevertheless, it is feasible that p40, p60, p50, and p100 may play important roles during recognition and/or catalysis at editing sites. A previous study of U insertion in Leishmania proposed that two RNA cross-linking proteins, ~80 and 100 kDa, from highly enriched editing extracts may be associated with editing site recognition, but the RNA substrate positions cross-linked remain to be determined (424).

Our competition analyses also suggest that editing complexes may preferentially recognize features of the pre-mRNA/gRNA hybrid (Fig. A-11). gRNA D33 is supplemented at ~100-fold the level of the radiolabeled A6 premRNA, in both standard cross-linking and editing assays, although we have seen that a ~200-fold excess affects neither activity (Fig. A-11B,D). Importantly, we have seen in native gels that during the pre-incubation step in our assays virtually all radiolabeled A6 pre-mRNsA anneals to gRNA D33 (see Materials and Methods section; data not shown). Addition of non-radiolabeled A6 pre-mRNA at 10-fold excess (or less) strongly inhibits cross-linking and editing (Fig. A-11A; data not shown), whereas 25- to 100-fold excess of other transcripts that should not hybridize with gRNA D33 have little or no effect. Interestingly, significantly structured transcripts such as tRNA (25-fold) appear relatively more inhibitory than predicted low-structured sequences, including the gRNA constructs (50-fold) and short RNA homopolymers (100-fold) tested (Fig. A-11;

data not shown). This apparent binding preference of editing complexes for RNA substrates *in vitro* is under further investigation in our laboratory.

Our observation of multiple cross-linking interactions at the ES1 for deletion in the A6 pre-mRNA/gRNA substrate may reflect that this site is dense with protein contacts in editing complexes (possibly not all detected by our cross-linking approach). Also the natural dynamics of interacting subunits, variable RNA substrate conformations, or protein breakdown may account for the multiple cross-links detected. These possibilities will be further studied in our laboratory. Furthermore, we observed the same cross-linking pattern in immunoprecipitated editing complexes enriched from bloodstream form trypanosomes ((425); data not shown). Together with our extensive purification of the procyclic complexes, this suggests that these proteins are part of the core complex and may not directly account for developmental regulation.

Finally, editing complexes contain subgroups of apparently related subunits sharing similar conserved motif (383). This may reflect the proposed functional and structural partition of insertion and deletion components in editing complexes (418, 426-429), and functions outside editing, including polycistronic mRNA, gRNA, and rRNA processing (430-431). Whether the editing complex cross-links reported here and/or other subunits occur at different deletion or insertion sites and in other substrates is currently under investigation in our laboratory.

Methods

Pre-mRNA and gRNA substrates

The ATPase 6 (A6) pre-mRNA editing substrates (Seiwert et al.1996) for deletion with gRNA D33 (392) were prepared as previously described. The site-specific radio-labeled and 4-thioU modified pre-mRNAs were obtained by ligation of two fragments as in (432). Acceptor RNAs were transcribed using the Uhlenbeck single-stranded T7 transcription method (433) and gel-purified. Donor thiolated RNAs were chemically synthesized by Dharmacon. The 4-thioU residue of the donor piece was radiolabeled to high-specific activity with polynucleotide kinase and [γ-^{32}P]ATP (using a 1:2 molar ratio of 5' ends:ATP), gel-purified, and ligated to the acceptor piece using the following DNA oligonucleotide bridges. A 3:1:2 molar ratio of acceptor/donor/bridge molecules was used.

Preparation of crude mitochondrial extracts and fractions containing enriched or purified editing complexes

Procyclic form (Pf) *T. brucei* strain TREU667 was grown in Cunningham media, and mitochondrial crude extracts were prepared as in (434) with modifications as in (414). Mitochondrial crude extracts were fractionated by ion-exchange chromatography in consecutive Q-sepharose (Q1) DNA-cellulose (D), and Q-sepharose (Q2) columns, as described by (393) and (414). The elution

fractions with the peak of editing complexes determined by Western blot analysis or editing activity also contained the peak of cross-linking activity in all purification steps.

Editing, adenylylation, and cross-linking analysis

Full-round editing reactions assembled in 20-μL mixtures with pre-annealed 3'-end labeled A6 pre-mRNA (~10 fmol) and gRNA D33 (~1.2 pmol) and adenylylation of RNA ligases in editing complexes were performed as in (426-427) and (417), respectively. For photocross-linking analysis, editing reactions were assembled as above, but in the absence of nucleotides, which somewhat improves cross-linking. The complete mixtures were incubated for 10 min at 26°C and an additional 10 min on ice prior to irradiation with 365-nm UV light (on ice for 10 min, ~5 cm below a Spectroline 150-V lamp) and subsequent treatment with RNases A and T1 (50 mg/mL and 120 U/μL) for 10 min at 37°C. After addition of 7 μL of 4X Laemmli buffer, the samples were analyzed by SDS-PAGE and autoradiography. RNA competitors at the indicated molar excess were included in the reaction mixture supplemented to the pre-annealed pre-mRNA/gRNA duplex in both cross-linking and editing assays. The 15-nt homopolymers were synthesized by IDT. The 121-nt viral RNA H121 was a gift from Cheng C. Kao (435). We have determined in native gels that our pre-annealing step yields >95% of the pre-mRNA in a duplex with gRNA D33 (not

shown), so further gRNA addition in Fig. A-11A,B should hybridize virtually all pre-mRNA.

Immunoprecipitation and Western blot analysis

Immunoprecipitations were performed essentially as described by (394) with minor modifications. For immunoprecipitation analysis of cross-linking proteins, editing reactions were scaled up 10 times and cross-linked as described above. One hundred micro-liters of Immunomagnetic beads (Dynabeads M-450; Dynal) were coupled with 225 µL of monoclonal antibodies (kindly provided by the laboratory of Ken Stuart, SBRI Seattle) and 1% BSA. Editing reactions were incubated with antibody-coated beads for 1 h at 4°C using a bi-directional shaker and occasional tapping. After washing two times with 100 µL of immunoprecipitation buffer (10 mM Tris at pH 7.2, 10 mM $MgCl_2$, 200 mM KCl, 0.1% Triton-X 100) the beads were resuspended with 100 µL of TE buffer and incubated in the presence of RNases A and T1 as described above. Upon the addition of 30 µL of 4X Laemmli buffer, the bead suspension was boiled at 100°C for 5 min and the supernatant analyzed by SDS-PAGE and autoradiography. The entire 200 µL unbound fraction and 100 µL washes mixed with 60 µL and 30 µL of 4X Laemmli buffer, respectively, boiled as well as analyzed. For Western blot analysis with the indicated monoclonal antibodies, protein samples (cross-linked to RNA or not) were separated by SDS-PAGE, blotted, and probed with the indicated mouse monoclonal antibodies at a dilution

of 1:25–1:50. The secondary antibody was applied at a 1:5,000 dilution and the blot developed using the ECL plus system (Amersham).

APPENDIX III

SUBSTRATE DETERMINANTS FOR RNA EDITING AND EDITING COMPLEX INTERACTIONS AT A SITE FOR FULL-ROUND U INSERTION*

Summary

Multisubunit RNA editing complexes catalyze uridylate insertion/deletion RNA editing directed by complementary guide RNAs (gRNAs). Editing in trypanosome mitochondria is transcript-specific and developmentally controlled, but the molecular mechanisms of substrate specificity remain unknown.

Here we used a minimal A6 pre-mRNA/gRNA substrate to define functional determinants for full-round insertion and editing complex interactions at the editing site 2 (ES2). Editing begins with pre-mRNA cleavage within an internal loop flanked by upstream and downstream duplexes with gRNA. We found that substrate recognition around the internal loop is sequence-independent and that completely artificial duplexes spanning a single helical turn are functional.

Furthermore, after our report of cross-linking interactions at the deletion ES1 (35), we show for the first time editing complex contacts at an insertion ES.

---

*Reprinted with permission from: "Substrate determinants for RNA editing and editing complex interactions at a site for full-round U insertion by Cifuentes-Rojas C, Pavia P, Hernandez A, Osterwisch D, Puerta C, Cruz-Reyes J. 2007 *The Journal of Biological Chemistry*, 282, 4265-4276.

Our studies using site-specific ribose 2' substitutions defined 2'-hydroxyls within the (a) gRNA loop region and (b) flanking helixes that markedly stimulate both pre-mRNA cleavage and editing complex interactions at ES2. Modification of the downstream helix affected scissile bond specificity. Notably, a single 2'-hydroxyl at ES2 is essential for cleavage but dispensable for editing complex cross-linking. This study provides new insights on substrate recognition during full-round editing, including the relevance of secondary structure and the first functional association of specific (pre-mRNA and gRNA) riboses with both endonuclease cleavage and cross-linking activities of editing complexes at an ES. Importantly, most observed cross-linking interactions are both conserved and relatively stable at ES2 and ES1 in hybrid substrates. However, they were also detected as transient low-stability contacts in a non-edited transcript.

Introduction

The single-mitochondrion containing kinetoplastid protozoa, including species of Trypanosoma and Leishmania, use cycles of uridylate insertion or deletion at numerous editing sites (ESs) within pre-mRNAs to generate mature mRNAs (for recent reviews, see (383-385)). This post-transcriptional mRNA maturation is catalyzed by a multisubunit editing complex (399, 403-404, 436-437) with specificity for the ESs being directed by small transacting guide RNAs (gRNAs) that are partially complementary to pre-mRNA (386-387, 436, 438).

A significant body of information has been accumulated on the functional and structural composition of editing complexes, including the identity of the subunits catalyzing the three steps of each editing cycle; they are mRNA cleavage at deletion and insertion ESs (403-404), U addition or U removal (398-399, 439) and RNA ligation at deletion and insertion ESs (400-402, 418, 428-429). The complexes are heterogeneous in protein composition but share most of the approximately 20 subunits identified (440). Several factors are also known or proposed to play auxiliary roles in editing (390-391, 395, 406-411, 441), although they are dispensable *in vitro* (393, 395-396, 412).

Much less is known about the mechanisms of substrate recognition including the protein subunits and substrate determinants that distinguish pre-edited (pre-) mRNAs from other transcripts and DNA in mitochondria. We recently reported the first observations of direct editing complex interactions with a functional site for full-round U deletion, showed preferential association with the editing substrate, and provided evidence for one of the interacting subunits corresponding to KREPA2 (442). However, editing complex interactions at insertion sites have not been reported. Other recent reports showed that bacterially expressed recombinant versions of the subunits KREPA3 and KREPA4 bind RNA (405, 443). The latter exhibited specificity for a gRNA 3'-oligo(U) tail.

In pre-mRNA/gRNA substrates, unpaired pre-mRNA uridylates or unpaired gRNA purines are landmarks of deletion or insertion sites, respectively

(386) and the number of such residues dictates the extent of U removal or addition (386-388). The two kinds of editing are likely to be differentially regulated as they involve separate activities and enzymes (403-404, 426-429) and there is evidence for their physical separation in heterogeneous complexes and subcomplexes (418, 440). Interestingly, efficient deletion and insertion editing have distinct requirements for a proposed pre-mRNA/gRNA ligation bridge (392, 444) and artificially interconverted sites use differing pre-mRNA lengths (445). The above observations suggest that the editing complex recognitions in and near an ES may also differ between the two editing types.

Our interconversion of functional ESs from deletion to insertion and vice versa experimentally demonstrated that the basic determinants that commit editing complexes into full-cycle deletion or insertion reside within the internal loop containing the targeted ES (445). However, additional features proximal and/or distal to an ES may modulate the efficiency of editosome assembly and catalysis. For example, discrete sequence changes affecting the pairing potential of residues adjoining an ES can significantly impact the specificity and efficacy of full-round and partial ("pre-cleaved") editing (444, 446).

The current model of trypanosome RNA editing postulates that natural sites should be flanked by a proximal upstream duplex between a purine-rich pre-mRNA sequence and a gRNA 3' poly-U tail (386-387, 436, 438) and an adjacent pre-mRNA/gRNA downstream "anchor" duplex that directs cleavage (386-389). Mutational analysis of the gRNA 3' region that stabilizes the

upstream duplex can significantly enhance full-round editing *in vitro* (392, 447). In *Leishmania tarentolae*, an upstream duplex was used in model U-insertion substrates by one group (448-449) but was not essential according to another (413). The latter group proposed that pre-mRNA purine sequences have a role in editing that is independent of base-pairing with gRNA (448-449). In a CYb pre-mRNA substrate, a 34-nt A/U element appeared to modulate gRNA-directed and gRNA-independent insertion (450) and a discrete 5' determinant near an editing site in a ND7 substrate was proposed (413, 449). In *Trypanosoma brucei*, the three model systems that currently recreate a full-round editing *in vitro*, A6, CYb, and RPS12 (387-388, 444-445, 451) are based on natural purine-rich pre-mRNA fragments.

We showed in A6 constructs that natural pre-mRNA extensions protruding from the upstream and downstream duplexes could be replaced by unnatural stretches without significant effects on editing. In one such construct, about half of a 5' poly-purine run proposed to stimulate editing *in vitro* (387) was substituted by pyrimidines (445). However, whether or not a specific pre-mRNA (or gRNA) sequence or its natural base composition modulates editing remains unclear. Previous structural studies in solutions of different natural-like mRNA/gRNA pairs have proposed that a common secondary structure may be important for editing complex recognition (452) but this has not been tested in functional *in vitro* systems.

Here we performed systematic sequence mutagenesis and ribose 2'-deoxynucleoside substitutions of a minimal A6 pre-mRNA/gRNA substrate to define functional determinants for both full-round U insertion and editing complex interactions at the targeted ES2. Our competition analyses of editing and RNA-protein interactions showed evidence of preferential association of editing complexes with the hybrid substrate. We observed that the requirement for the duplexes flanking the internal loop is sequence-independent, and artificial helices spanning a single turn support efficient editing. We also found that specific ribose 2 -hydroxyls in both strands of the downstream helix and, surprisingly, in the gRNAloop region strongly stimulate both pre-mRNA cleavage and editing complex interactions at the targeted insertion site. Moreover, 2'-deoxy substitution of the downstream helix affected scissile bond selectivity, whereas the tested modifications in either pre-mRNA or gRNA strand had relatively moderate effects. Notably, the 2 -hydroxyl moiety adjoining the scissile bond is an essential determinant of insertion, potentially involved in cleavage catalysis.

The current studies of trypanosome full-round insertion editing provide significant insights on the relevance of the substrate secondary structure rather than its specific sequence and suggest that specific pre-mRNA and gRNA riboses significantly affect both pre-mRNA cleavage and editing complex interactions at the targeted bond.

Results

Analysis of the Natural A6 Pre-mRNA Features Proximal to ES2 for Full-round Insertion

Features in the RNA substrate that are specifically recognized during full-round editing are not fully defined in trypanosomes. These may include the native pre-mRNA sequence, purine richness, and higher-order structure of the pre-mRNA/gRNA bimolecular substrate. To address this question we characterized the proximal features of an A6 RNA pair (Pair-1) for ES2 insertion that uses a 45-nt pre-mRNA fragment (Fig. A-13A; Ref. 41). We have shown that minimal 43–45-nt pre-mRNA substrates support efficient fullround insertion in the A6, CYb, and RPS12 systems (445). For simplicity, we will refer to the upstream and downstream duplexes (relative to the pre-mRNA) flanking the internal loop containing ES2 as “left” and “right” helices. The terminal 5’ extensions of Pair-1 will be termed pre-mRNA and gRNA protrusions, respectively (Fig. A-13A).

We first analyzed the left helix of Pair-1. Our previous studies showed that virtually all natural 5’ purines in the pre-mRNA protrusion could be substituted by pyrimidines (445). It was also reported that natural pre-mRNA sequence beyond the residues forming the right duplex were dispensable for ES2 insertion (447). To assess the importance of the natural pre-mRNA sequence in the left helix and the requirement for the pre-mRNA protrusion, we designed Pair-2 containing an artificial 15-bp blunt-ended left duplex (Fig. A-

13A). Furthermore, the pre-mRNA/gRNA-paired residues in this duplex were flipped to alternate all purines and pyrimidines (except the first two residues needed for T7 *in vitro* transcription). Pair-2 supported insertion at a level comparable with the parental Pair-1 (Figs. A-13 B and C), thus showing that neither the natural pre-mRNA sequence nor purine richness in the left duplex or the pre-mRNA protrusion is required for efficient insertion. Note that to preserve both a pre-mRNA functional length (445) and its purine content in Pair-1 (77.8%), all 25 upstream purines in the parental substrate were moved to the 3' end of Pair-2 (75% purines).

To further analyze the functional length of the left helix, we tested Pair-2 derivatives (Fig. A-13A) containing 12-bp (Pair-3) or 11-bp (Pair-4) helices with predicted stabilities similar to the parental duplex (data not shown). Interestingly, Pair-3 was edited less efficiently than Pair-1. This may reflect a partial influence of nucleotide composition of the helix. Notably, the artificial 11-bp upstream duplex in Pair-4, which represents ~one turn of helical RNA (453), efficiently replaced the complete 5' region of the parental Pair-1. Because the minimal length of the natural A6 pre-mRNA for efficient full-round ES2 insertion is ~43–45 nt (445), we trimmed the rightward region of Pair-4 to generate Pair-5 (45-nt pre-mRNA; Fig. A-14A). This substrate, with reduced purine content (71%), supported less accurate editing (i.e. 3U addition) than the parental Pair-1 and accumulated inaccurate insertion by 1U addition (Fig. A-14B). However, relocation of the protruding 3' purines of Pair-5 to the 5' end (Pair-6; 71%

purines) re-established 3U insertion nearly to the level of the parental Pair-1 (lane 3). Although the latter 45-nt constructs imply that a short protrusion may be more stimulatory 5' than 3' to ES2, alternative structural reasons are also feasible. To determine whether or not the pre-mRNA extension in Pair-6 must be purine-rich, we substituted most of the protruding purines by pyrimidines (Pair-7). Interestingly, Pair-7 was about as efficient as Pair-1 despite its relatively low (49%) purine content (Fig. A-14BC, compare the first and last lanes).

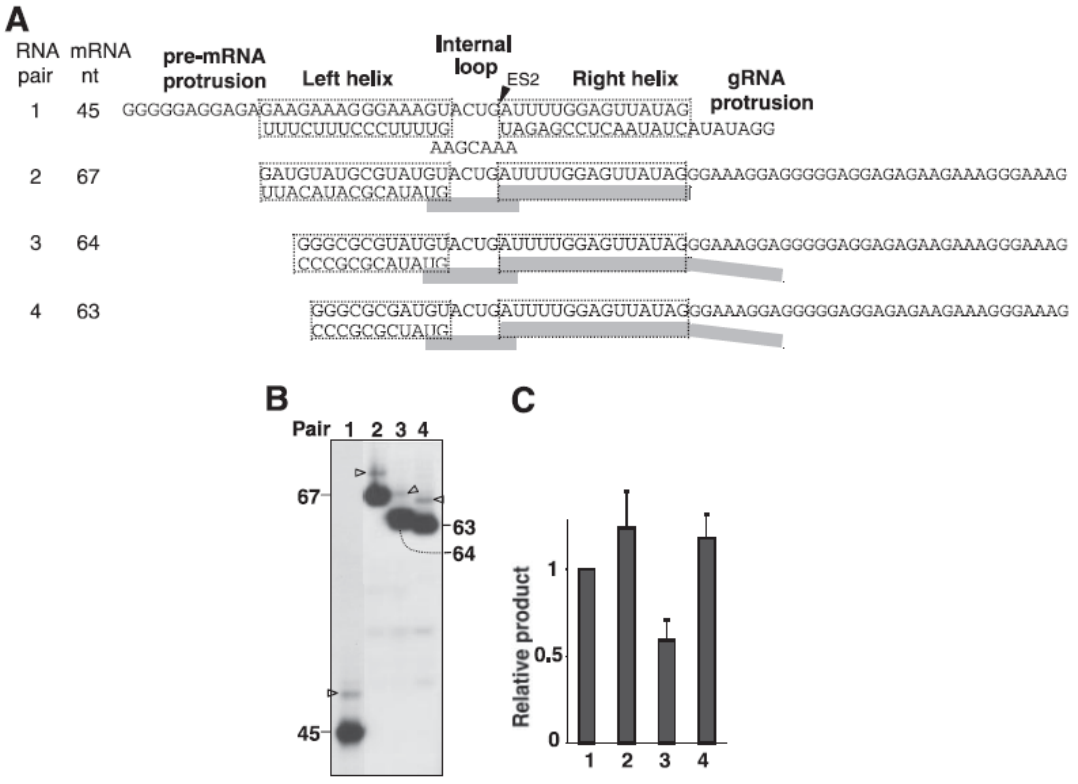


FIGURE A-13. Analysis of the left helix in the parental A6 pre-mRNA/gRNA substrate for full-round insertion at ES2 (Pair-1) (41). A, the starting Pair-1 substrate and nomenclature of the analyzed regions. Derivative constructs are aligned. The gRNA gA6[14]USD-3A (447) in Pair-1 (in blue) and the ES2 (arrowhead) are indicated. The predicted stability (3.0 Mfol, M. Zuker program (342)) of left and right duplexes (boxed) of Pair-1 are -20.5 and -21.5 kcal/mol. The based composition of the derivative duplexes was adjusted to conserve a similar $\Delta G°$ as in Pair-1. In the left columns the RNA pair assigned number and pre-mRNA size are indicated. Parental unmodified gRNA sequences are depicted by filled boxes. B, full-round insertion assays using 3'-end-labeled pre-mRNA constructs. Accurate insertion by addition of three Us is indicated by an arrowhead. The pre-mRNA size and assigned pair number are indicated. C, plots of relative accumulation of accurately edited product.

Together, the above results indicate that neither the natural polypurine run nor the overall purine richness of A6 pre-mRNA is a critical determinant of ES2 insertion *in vitro*. Combined with our previous observations (445), these features upstream of the internal loop is sequence-independent. Furthermore, a single turn of helical RNA was sufficient for full-round insertion. Additional work will be required to test whether smaller and less stable left helixes are efficient.

We then examined the features of the right helix of Pair-1 required for editing. To this end we used variants of Pair-6 (Fig. A-14A) containing 14-bp (Pair-8) and 13-bp (Pair-9) right duplexes with a predicted stability comparable with the parental helix (data not shown). Both Pair-8 and Pair-9 supported editing nearly as efficiently as the parental Pair-1 (Fig. A-14C). We then tested derivatives bearing either 12-bp (Pair-10) or 11-bp (Pair-11) right helices.

Notably, whereas the base composition of the 3' duplex significantly deviated from the parental helix, these substrates were appreciably more efficient than Pair-1 (Figs. A-14D and E). It is also worth noting that the predicted stability the right helix in Pair-11 is ~10% lower than in Pair-1 (see the legend to Fig. A-14).

Finally, we generated Pair-12 bearing a blunt-ended right duplex of 11 bp, and a change of the 5'-terminal two gRNA residues to facilitate *in vitro* transcription starting with G (433). This last construct was less efficient than Pair-11. It is possible that a branch structure downstream of ES2 favors insertion editing. Consistent with this notion, a derivative of Pair-10 with a right blunt-

ended duplex was also less efficient (data not shown). Altogether, these data show that substrates with one turn of helical RNA at both the left and right sides of ES2 support efficient full-round U insertion and that the natural sequence of the minimal (~43–45 nt) pre-mRNA outside the internal loop is not required. This suggests that the basic editing complex recognitions flanking the internal loop involve sequence-independent features of the pre-mRNA/gRNA pair.

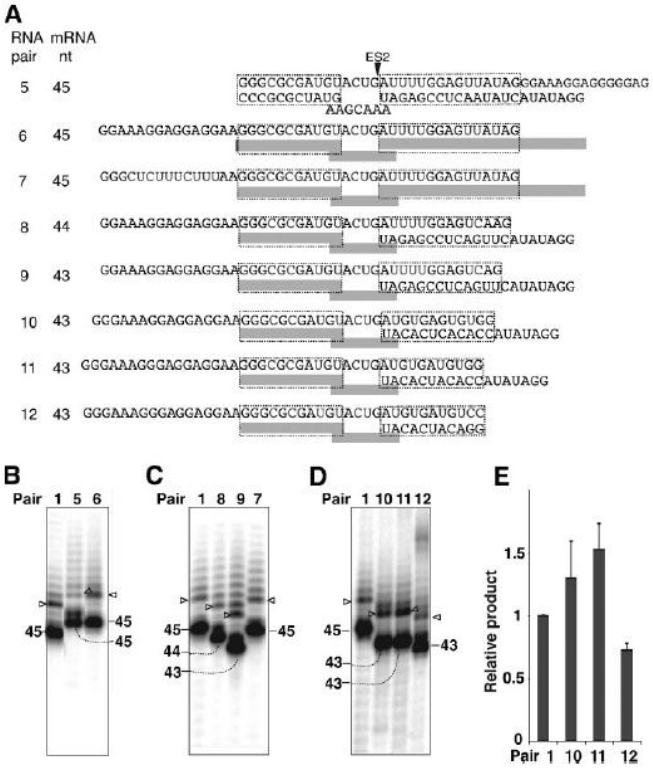


FIGURE A-14. Analysis of the right helix. A, the starting Pair-5 and derivatives are aligned. All labeling is as in Fig. A-13A. B–D, full-round insertion assays are as in Fig. A-13B. The derivative duplexes conserved a similar predicted ΔG°, except for the right helix in Pair-11 and Pair-12 that dropped to -18.3 kcal/mol. The pre-mRNA size in these pairs was adjusted to a minimum of 43 nt (41), as the 3' end was truncated. Some sequence-dependent gel migration differences were observed (e.g. the 3' end of the last two 43-nt pre-mRNAs differ by two residues). E, plot as in Fig. A-13C.

Effect of Ribose 2' Substitutions on Full-round Insertion

We analyzed the contribution of ribose 2'-hydroxyl groups to substrate recognition by editing complexes, by incorporating 2'-deoxy substitutions and other 2' modifications in and around ES2 in the parental Pair-1 (summarized in Fig. A-18).

pre-mRNA Residues Upstream of the Internal Loop

We first analyzed the pre-mRNA 5' region (Fig. A-15A) using derivatives containing 11 deoxynucleotides that replaced either the entire pre-mRNA protrusion (Pair-13) or most bases in the left helix (Pair-14). Interestingly, both constructs supported insertion at about half the level of the parental Pair-1 (Fig. A-15B, upper panel). This decrease in editing was largely paralleled by a reduction in ES2 cleavage (Fig. A-15B, lower panel, and C). Importantly, pre-mRNA cleavage was scored in absence of RNA ligase activity using editing complexes that were pretreated with 10 mM PPi (Ref. (429); see "Experimental Procedures"). These two constructs showed that 2'-hydroxyl groups in the pre-mRNA protrusion and the left helix partially stimulate insertion.

Our previous work showed that truncation of the protrusion in the minimal 45-nt A6 pre-mRNA (see 34-nt RNA 8 in Cifuentes-Rojas et al. (445)) abrogates full-round insertion. This suggests that editing complexes make sufficient contacts with the all-DNA protrusion to support an appreciable insertion level, and thus, the 2'-hydroxyls of the pre-mRNA protrusion are significantly

stimulatory but not essential. Furthermore, the left RNA/DNA heteroduplex in Pair-14 should adopt a shape that is more similar to the A-form (of dsRNA) than the mainly reflect a requirement for 2'-hydroxyls, although shape dependent recognitions may also be important. We then tested the effect of a single 2'-deoxynucleoside substitution adjoining the scissile bond (Pair-15). Notably, this 2'-H abolished both insertion and cleavage (Figs. A-15B and C). This may reflect a direct role of the 2'-OH at ES2 in catalysis and/or site recognition or an indirect role due to a change in the sugar pucker (from ribose C3'-endo to deoxyribose C2'-endo conformation (453).

To address these two possibilities we tested other ribose 2' modifications such as 2'-O-methyl (-OCH_3; Pair-16) that favor the RNA-like sugar pucker but cannot act as a hydrogen bond donor (57). 2'-OCH_3 supported neither insertion nor cleavage (Figs. A-15B and C), but it is conceivable that the added bulk of this group, compared with a 2'-OH, caused steric hindrance. We then tested the smaller 2'-fluorine (-F) modification (Pair-17), which should also promote RNA-like C3'-endo conformation even more than 2'-OCH_3 and is highly unlikely to accept a proton (454). This substituent also completely inhibited cleavage (Fig. A-15C and D; see the legend). Unfortunately, a 2'-NH_2 modified guanosine at ES2 is not available (Dharmacon).

Based on these results, it is conceivable that formation of a hydrogen bond by the ribose 2'-OH group at ES2 is required for insertion. The ribose 2'-

hydroxyl group at ES2 could mediate either catalysis at pre-mRNA cleavage, editing site recognition, or editing complex assembly onto the substrate.

To attempt distinguishing between these possibilities, we modified our recently developed photo-crosslinking assay to analyze direct editing complex contacts at ES1 in A6 pre-mRNA (442). To similarly assay ES2, we placed a single ^{32}P-labeled 4-thioU immediately 3' of the scissile bond. Pairing of this residue with a guiding adenosine should extend the right helix by 1 base pair (Pair-15'; Fig. A-15E). Notably, the thiolated ES2 supported protein crosslinking interactions that both are similar to those reported for ES1 (442) and specifically co-purify and co-immunoprecipitate with editing complexes (see Fig. A-19). Substrates with either a 2'-OH (Pair-1') or the inhibitory 2'-H modification (Pair-15) at ES2 exhibited identical cross-linking patterns (Fig. A-15E), suggesting a similar editing complex association with both the 2'-H-substituted and unmodified ES2. Thus, the single 2'-deoxy substitution at ES2 does not seem to prevent editing complex interactions at ES2, adjoining the scissile bond may play a role in catalysis.

Pre-mRNA Residues in the Right Helix

We then examined the pre-mRNA residues in Pair-1 that are part of a 15-bp right duplex. Three deoxynucleotide patches were initially compared, 10, 14, and 15 nt long (Pairs 18, 19, and 20, respectively; Fig. A-15A). Note that these pre-mRNAs contain a 3'-most ribonucleoside to allow end-radiolabeling with T4

RNA ligase (see “Materials and methods” (454). Interestingly, these ribose substitutions decreased both insertion and cleavage compared with Pair-1 (Figs. A-15F and G). This inhibition increased with the extent of deoxy substitution. Notably, Pair-20 with all upstream pre-mRNA residues modified was most inhibited.

To determine the importance of the 2’-hydroxyl immediately 3’ of the cleavage site, the singly substituted Pair-21 was tested. This substrate was about half inhibited in both editing and cleavage assays (Figs. A-15F and G). Combined, these observations indicate that the 2’-hydroxyls just 3’ of ES2 and further downstream in the duplexed pre-mRNA are significantly stimulatory. These deoxy substitutions did negatively impact insertion, primarily at pre-mRNA cleavage.

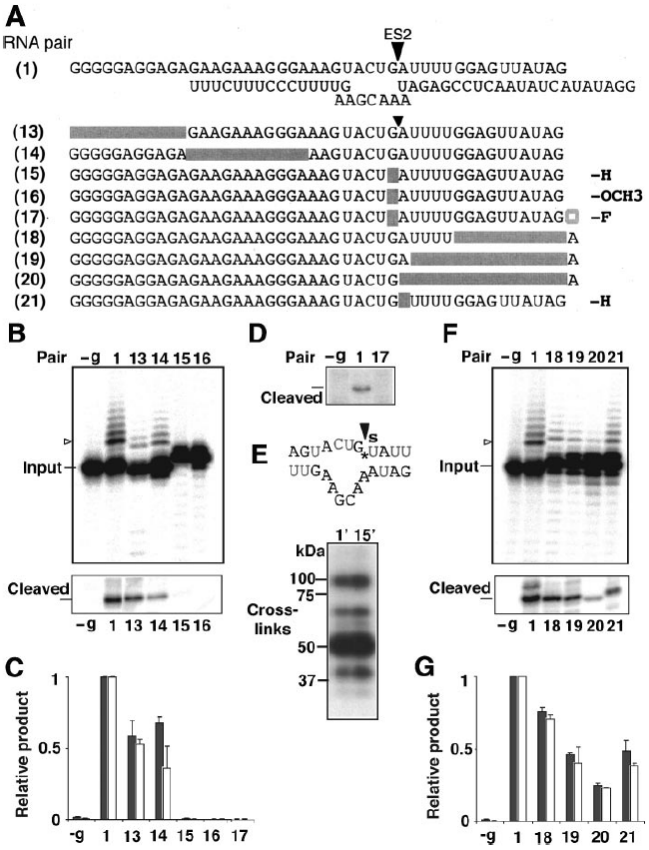


FIGURE A-15. Ribose 2'-substitution of upstream and downstream residues in the pre-mRNA of Pair-1. A, Pair-1 and derivatives are aligned (pre-mRNA strand is shown). Single or multiple 2'-substitutions are indicated in boxes. The modifications are 2'-deoxy (-H), 2'methoxy (-OCH_3) or 2'-fluorine (-F). The latter substrate was manufactured with a 3'deoxynucleoside (empty box). All other labeling is as in Fig. A-13B, insertion and cleavage (upper and lower panels, respectively) of 5'-substituted pre-mRNAs. A control lane devoid of gRNA (-g) was included. C and G, relative accumulation of accurately edited (black bars) and ES2 cleavage (white bars) of 3'-end labeled pre-mRNA. D, cleavage of 2'-fluorine modified, 5'-end labeled mRNA (see "Experimental Procedures"). Precise cleavage at ES2 was confirmed using ribonuclease T1 and hydroxyl partial pre-mRNA digestions (not shown). E, editosome photocross-linking with pre-mRNAs containing a single ^{32}P and 4-ThioU at ES2 of Pair-1 and Pair-15. Diagram indicates the labeled bond (*) and thiolated U (s). The right duplex was extended by one base pair between Thio-U and a guiding adenosine. The size of molecular markers is indicated in kDa. F, insertion and cleavage assay of 3'-substituted pre-mRNAs.

gRNA Residues and Duplexes Flanking the Internal Loop

Apart from the critical ribose 2'-OH at the editing site, most pre-mRNA 2'-hydroxyl groups tested were stimulatory but not essential for ES2 insertion. We then examined the effect of proximal gRNA substitutions and DNA duplexes (Fig. A-16A). A 10-deoxynucleotide patch on the gRNA strand at either side of the internal loop (Pair-22 and Pair-24, respectively) had a slight negative effect on insertion (Fig. A-16B) comparable with that observed with corresponding pre-mRNA patches (Pair-14 and Pair-18). In contrast, DNA duplexes formed by the complementary patches at left (Pair-23) or right (Pair-25) of ES2 were more inhibitory, particularly the Pair-25 (Fig. A-16B). In both cases, insertion and pre-mRNA cleavage at ES2 were similarly inhibited (Fig. A-16D).

Interestingly, the right DNA duplex also significantly affected the scissile-bond selectivity. That is, the pre-mRNA of Pair-25 was cleaved at several residues flanking ES2; the upstream cuts are in the loop, whereas the downstream cuts are in the right duplex. All these cleavages were gRNA-dependent (not shown). We assessed whether the inhibitory DNA duplexes affected the editing complex photo-cross-linking interactions with ES2. Interestingly, thiolated versions of these substrates (Pair-23' and Pair-25') reduced the level of all cross-linking subunits (Fig. A-16E). This effect was particularly severe with Pair-25'. However, because protein-RNA cross-linking can be quite sensitive to conformational changes, inhibition of cross-linking activity may reflect local structural changes of the editing site rather than

reduced editing complex assembly onto the RNA substrate. A native gel analysis of DNA duplex-containing pairs and Pair-1 confirmed complete annealing of these substrates (Fig. A-16F).

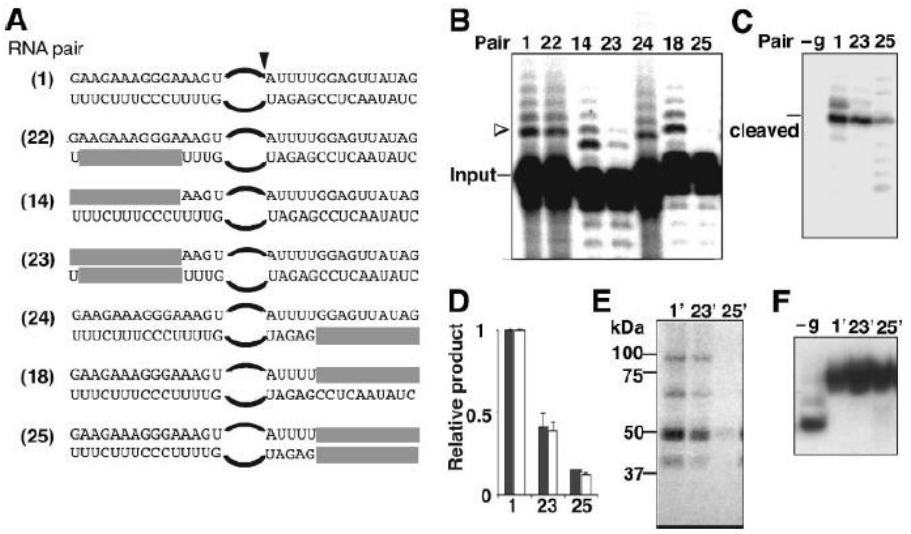


FIGURE A-16. Ribose 2' substitution of Pair-1 pre-mRNA and gRNA strands. A, pair-1 and derivatives are aligned. Internal loop residues around ES2 are depicted as curved lines with an arrowhead pointing to ES2. Deoxynucleoside substituted residues are boxed. B, insertion assay. C, cleavage assay. The inaccurate cleavages with Pair-25 are also gRNA dependent (data not shown). Plots (D) and photo-cross-linking assays (E) are as in Fig. A-15F, annealing assays of RNA pairs used in E. Control lanes without gRNA (-g) are indicated.

Combined, the parallel inhibition of insertion, cleavage, and cross-linking activities at ES2 suggest that proximal DNA duplexes negatively impact a productive interaction of editing complexes with the substrate and, thereby, catalysis.

Internal Loop Residues

Finally, we examined the importance of 2'-hydroxyl groups in the internal-loop residues containing ES2 (Fig. 5A). Deoxy substitutions in the three residues 5' of the guanosine at ES2 (Pair-26) had virtually no effect on either insertion or cleavage (Fig. 5, B and C). In contrast, 2'-H substitution in all seven loop gRNA nucleotides (Pair-27) significantly reduced editing and cleavage (Fig. 5, C and D). Notably, modification of both strands of the internal loop (Pair-28) further inhibited both editing and cleavage. Furthermore, editing complex cross-linking at ES2 was also moderately and strongly reduced in the corresponding Pair-27' and Pair-28' substrates, respectively (Fig. 5E). A native gel analysis confirmed the complete annealing of these pairs (Fig. 5F).

Together, these observations indicate that several hydroxyl groups in the internal loop are relevant determinants of insertion that largely influence both the efficiency of pre-mRNA cleavage and cross-linking by editing complexes. Interestingly, hydroxyls on the gRNA strand appeared to compensate for deoxy substitutions on the pre-mRNA strand but not vice versa. That is, one or more 2'-hydroxyls in the gRNA internal loop residues significantly stimulate *in trans*

pre-mRNA cleavage and/or site recognition by editing complexes. The analyses in Figs. 3–5 are summarized in Fig. 6.

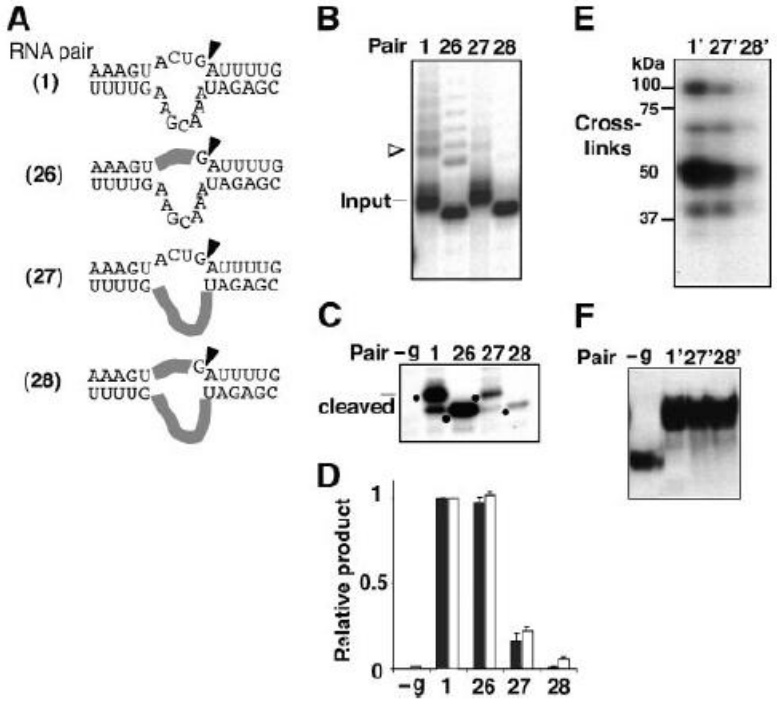


FIGURE A-17. Ribose 2' substitution of internal loop residues around ES2 in Pair-1. A, Pair-1 and derivatives. B, insertion assay. C, cleavage assay. ES2-cleavage products are marked with filled dots. Note that the offset gel mobility reflects the use of the same pre-mRNA in Pairs 1 and 27 as well as in Pairs 26 and 28. D, plots and E, photo-cross-linking assays are as in Fig. A-15 F, annealing assays of RNA pairs used in D. Control lanes without gRNA are indicated.

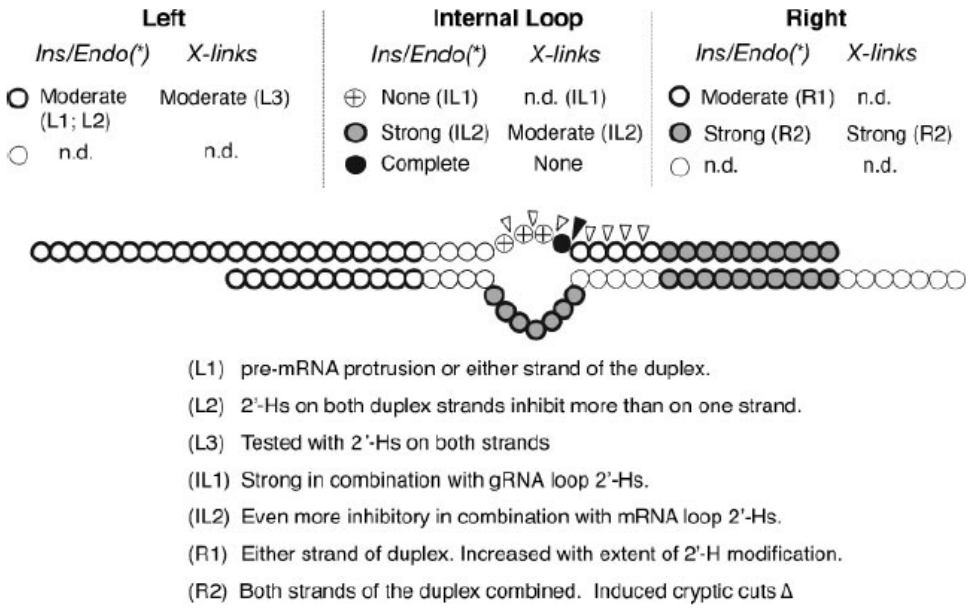


FIGURE A-18. Summary of ribose 2'-deoxy substitutions tested on Pair-1 for full-round insertion at ES2. Upper panel, indicates the level of inhibition for the left and right sides of the internal loop and within the loop (each region is separated by a vertical line). Full-round insertion and pre-mRNA cleavage (Ins/Endo) are on the left, and editosomes cross-linking (X-links) are at the right. The asterisk indicates that insertion and cleavage are affected at comparable levels. Circle types representing the level of inhibition: thick line, moderate; thin line, not determined (n.d.); gray (in addition to thick line), strong; with a pattern, no effect; black, complete. Parentheses indicate clarification notes (lower panel). A filled arrowhead points to the natural ES2 for full-round insertion. dsDNA within the right duplex induced cryptic cuts at several residues (open arrowheads) flanking the editing site. Middle panel, diagram of Pair-1 with individual residues shown as circles. Lower panel, explanatory notes on the effect of the indicated modifications.

Specificity of Cross-linking Interactions at the Insertion ES2

Several observations indicate that the protein-RNA photo-crosslinking interactions at the insertion ES2 represent direct contacts of editing complexes with the substrate. For example, the cross-links are specifically co-immunoprecipitated by antibodies raised against protein subunits of the complex (Fig. A-19A and data not shown). Also, native complexes purified by either two or three steps of consecutive ion-exchange chromatography exhibit comparable cross-linking (Fig. A-19B) and editing activities at ES2 in the minimal substrate analyzed (445). Thus, although the latter preparation has a simpler protein composition (Fig. A-19C), the two preparations of editing complexes appear functionally equivalent with the substrate analyzed. Furthermore, the presence of representative subunits (Fig. A-19D and data not shown) as well as all critical catalytic activities including editing endonucleases suggest that the functional and protein composition of our complexes is similar to that reported by other groups (383, 385).

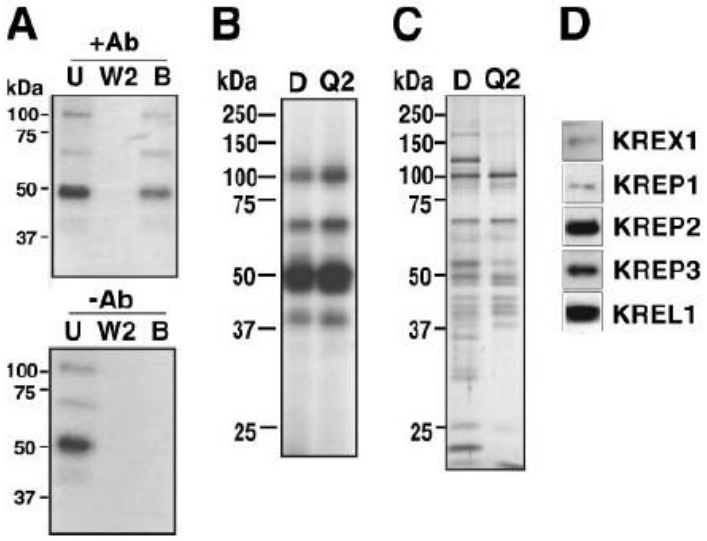


FIGURE A-19. The RNA-protein cross-linking interactions at ES2 co-immunoprecipitate and co-purify with native editing complexes obtained by sequential steps of chromatography. A, an immunoprecipitation assay using anti-KREPA3 monoclonal antibodies (Ab), including the unbound (U), second wash (W2), and bound immunoprecipitated (B) fractions (upper panel). A parallel mock assay without antibodies (-Ab) is also shown (lower panel). B, the cross-linking pattern is conserved in the peak fractions of consecutive chromatography steps; D and Q2 represent two-step and three-step purification protocols, respectively (see “Experimental Procedures”). C, silver-staining of the fractions in B. D, Western blot analysis of the Q2 fraction with available antibodies against five different stably-bound subunits of editing complexes.

Together, these results indicate that the cross-linking interactions at ES2 are specific to editing complexes. This is consistent with our recently reported observations of cross-linking interactions by editing complexes at the deletion site ES1 in a similar A6 substrate (442). We analyzed the specificity of the cross-linking and pre-mRNA cleavage activities at ES2 in competition analyses with homologous and heterologous transcripts (Figs. A-20 A and B). In both cases the bimolecular A6 substrate was readily competed out by a 5- and 10-fold excess of homologous pre-mRNA that can hybridize with free cognate gRNA (~120 times the pre-mRNA concentration in the standard mixture; see "Materials and methods").

In contrast, heterologous transcripts including another pre-mRNA, tRNA, and a non-complementary gRNA were partially or not inhibitory at greater (10- and 25-fold) excess. Finally, further addition of cognate gRNA was not inhibitory. Thus, both cross-linking and cleavage activities of editing complexes at ES2 exhibit specificity for the hybrid substrate. These observations are also consistent with our reported preferential interactions of editing complexes at ES1 (442). Interestingly, the highly structured tRNA competitor partially affected both cross-linking and editing activities more at ES2 than at ES1 (Ref. (442); Fig. A-20). The similar gel mobility of mayor cross-links at ES2 and ES1 (Fig. A-21) suggested that the same subunits of editing complexes make these contacts. However, the cross-linking efficiency at ES1 is significantly greater than at ES2,

consistent with the former substrate supporting a much higher level of editing *in vitro* (392).

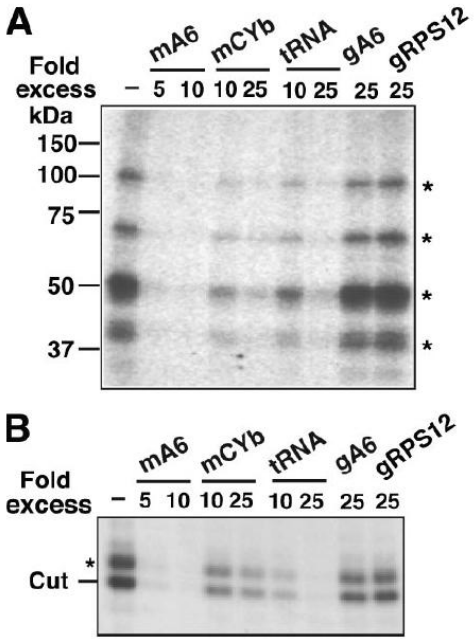


FIGURE A-20. Competition analysis of cross-linking and pre-mRNA endonuclease activities of editing complexes. Cross-linking (A) and cleavage (B) assays with or without a 5-, and 10-molar excess of homologous A6 pre-mRNA (mA6) or 10- and 25-fold excess of heterologous CYb mRNA (mCYb), tRNA, and non-complementary gRNA gRPS12. Extra cognate gRNA (gA6) was also supplemented to the standard mixture containing 120-fold excess of this transcript (see “Experimental Procedures”). The cleaved pre-mRNA piece (Cut) is mixed with a 1-nt 3 -extended homologue fragment (asterisk) derived from T7 RNA polymerase *in vitro* transcription.

Because the cross-linking patterns at ES2 and ES1 are similar, we asked whether a transcript that does not undergo editing can cross-link with editing complexes. To this end we tested gRNA D33, which does not inhibit ES1 cross-linking at a ~200-fold excess and exhibits virtually no predicted structure (442). Surprisingly, such a transcript containing a single photoreactive 4-thioU supported a level and pattern of cross-linking comparable with that of ES2. Nevertheless, as expected, an excess of the homologous competitor inhibited all cross-linking by thiolated D33, whereas the same (20–40-fold) or greater (200-fold) excess of competitor had no effect on the cross-linking interactions at ES2 and ES1 (Fig. A-21 and data not shown).

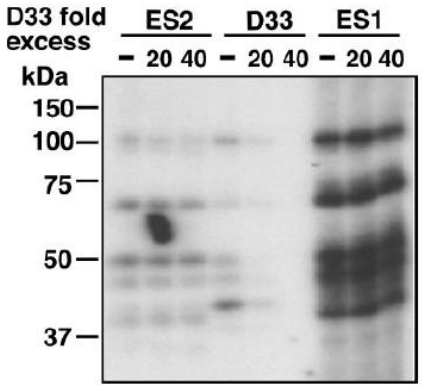


FIGURE A-21. Competition analysis of cross-linking interactions with bimolecular editing substrates and a non-edited transcript. Lanes 1–3, 2–5, and 7–9 show interactions with ES2, ES1, and D33 photo-reactive substrates, respectively. The D33 competitor was used at the indicated molar excess.

Overall, these observations suggest that the cross-linking subunits of editing complexes can make transient nonspecific contacts with RNA; however, the similar interactions at ES2 and ES1 are significantly more stable.

Discussion

The molecular basis of substrate recognition by editing complexes and the regulation of RNA editing in the single mitochondrion of trypanosomes are still poorly understood. The purpose of these studies was to dissect functional substrate determinants proximal to a site for full-round U insertion catalyzed by purified editing complexes. Combined, these observations and our previous study (442) have important implications on the mechanisms of substrate recognition by editing complexes. First, the overall recognition of the insertion substrate outside the internal loop is sequence independent. This notion is consistent with structural probing studies suggesting that related secondary structures of different mRNA/gRNA pairs may be important for editosome recognition (452).

Our analysis of a minimal A6 RNA pair for full-round insertion at ES2 (445) showed that the sequence and base composition of the parental helices flanking the editing site, including the pre-mRNA purine-richness (93%) in the left duplex, are not required for efficient editing. We had recently shown that the all-purine pre-mRNA protrusion could be replaced by a pyrimidine-rich stretch (445). Although natural 5' poly-purine runs in the A6 pre-mRNA are dispensable

for the basic insertion reaction *in vitro*, it is conceivable that these structures are specifically recognized *in vivo* by factors to promote nucleation with the complementary uridylate tail of gRNAs or stabilization of the duplex. In line with this notion, an editing complex subunit (KREPA4) was recently reported to exhibit binding specificity for a gRNA 3' U-tail (37) (443). Furthermore, a proposed accessory factor (REAP-1) preferentially binds to purine-rich transcripts such as pre-mRNAs (455). Notably, substrates with completely artificial 11-bp duplexes (i.e. one helical turn of RNA) flanking ES2 support efficient insertion (e.g. Pair-11 and Pair-12; Fig. A-14).

We speculate that one or both duplexes flanking ES2 is recognized by KREN2, an endonuclease subunit that specifically serves in insertion (404) and bears an RNase III-like domain and one double-stranded RNA binding motif (383, 395). Interestingly, the 11-bp artificial helices in our substrates may be minimal in length, as structural studies of highly conserved double-stranded RNA binding motifs in other systems indicate that these proteins typically interact with 16-bp (~1.5 helical turns of dsRNA; Ref. (456)). Moreover, a recent study proposed that the smallest dsRNA substrate for the single double-stranded RNA binding motif in a bacterial RNases III is 11 base pairs (457). KREN2 may dimerize (24) like other class 1 RNase III enzymes (437), so that each double-stranded RNA binding motif could contact one of the 11-bp helices flanking ES2 in our constructs. Typical double-stranded RNA binding motifs specifically bind the A-form of dsRNA through interactions that are adapted to

the shape of the helix, are sequence-independent, and primarily involve hydrogen bonds with ribose 2'-hydroxyls (458). This is consistent with the observed inhibition of both U insertion and editing complex photo-cross-linking interactions at ES2 by the presence of (B-form) dsDNA in the flanking helices (Fig. A-16B). Also in line with this notion, 2' deoxy substitutions in one strand caused more moderate negative effects (Figs. A-16 B and E) possibly because DNA/RNA heteroduplexes retain half of the hydroxyls and conserve more the RNA than the DNA helical shape (459). An 11-bp DNA duplex 5 bp downstream of ES2 (Pair-25) markedly reduced accurate cleavage and, instead, stimulated low-level cryptic pre-mRNA cuts near ES2 (Fig. A-16C). Such an effect on scissile-bond selection suggests that the modified riboses in the right duplex help position KREN2 to precisely cleave the bond at the single-/double-strand junction (ES2). Interestingly, the significant cleavage inhibition correlates with a dramatic reduction of editing complex cross-linking at the same site (Fig. A-16E). Similarly, the presence of helical DNA 4 bp upstream the internal loop moderately inhibited both cleavage and RNA-protein interactions at ES2.

Together, these data suggest that the editing complex makes important contacts with both helices flanking ES2. Such contacts with the downstream anchor duplex appear particularly relevant for efficient and accurate pre-mRNA cleavage as well as for cross-linking interactions at the editing site. 2'–hydroxyl groups in the pre-mRNA protrusion (i.e. Pair-13; Fig. A-15A) also contribute to the insertion reaction but more moderately. However, a construct with the

combined modifications of Pairs 13 and 14 significantly inhibited cleavage (data not shown) and, thus, further emphasize the impact of the upstream pre-mRNA structure on this editing step. Single ribose 2' substitutions at ES2 including 2'-H, 2'-OCH_3, and 2'-F showed that the 2'-hydroxyl at this site is critical for either pre-mRNA cleavage or prior ES2 recognition. Interestingly, ES2 with either a 2'-OH or 2'-H supported the same pattern of editing complex cross-linking interactions at the editing site. Because protein-RNA photo-cross-linking is sensitive to substrate conformational changes, we suspect that most protein interactions involved in ES2 recognition are unaffected at the 2'-H modified site.

A potential role of the 2'-hydroxyl at ES2 in catalysis could involve 1) formation of a critical hydrogen bond with the KREN2 endonuclease, 2) coordination of a divalent cation cofactor, or 3) direct nucleophilic attack on the scissile bond. Editing complex-catalyzed cleavage of pre-mRNA containing a single ^{32}P at ES2 produces a 5'-labeled downstream fragment (data not shown) rather than a 3'-labeled upstream fragment as would be expected if the 2'-OH group forms a 2',3' cyclic phosphate upon cleavage, as occurs with RNase A (427). This is consistent with previous RNase mapping that deduced the 5'-P and 3'-OH nature of the termini produced at ESs (388-389) and also is in line with an RNase III Mg^{2+}-dependent type of processing (388-389). Further work will be needed to dissect the precise role/s of the 2'-hydroxyl group at ES2 in pre-mRNA cleavage and/or editing site recognition. Our deoxy substitutions within the ES2 internal loop were also quite informative. Interestingly,

substitution of the three residues 5’ of the guanosine (bearing a critical 2’-OH) at ES2 affected neither cleavage nor U-insertion; however, substitution of all apposing loop gRNA residues strongly inhibited both insertion at pre-mRNA cleavage and cross-linking activities at ES2. These results were unanticipated as they reveal that the gRNA strand in the internal loop is an important determinant of full-round U-insertion at the level of pre-mRNA cleavage and suggest that the proximal pre-mRNA 2’-OH groups upstream of the guanosine at ES2 are less relevant for substrate cleavage.

Because the relative level of insertion and cleavage were similar, these 2’-deoxy modifications appear to have little or no effect on either U addition (by the terminal U-transferase, KRET2) or RNA ligation (by KREL2) in the insertion cycle. It is intriuing that the combined pre-mRNA/gRNA substitutions in the loop were more inhibitory (in all assays tested) than gRNA modifications alone because pre-mRNA substitutions had no effect. It is possible that the pre-mRNA deoxynucleotides facilitate a conformation of the substituted gRNA loop that is particularly inhibitory. Several editing complex subunits contain conserved motifs that may bind single-stranded RNA around an ES. For example, KREPA1 (also termed band II, LC-1; Refs.(383, 393, 396)) has an oligonucleotide/oligosaccharide binding (OB)-fold that could bind single-stranded RNA in interactions independent of ribose 2’-hydroxyls (460). KREPA1 was speculated to have a role in recognition of the editing substrate and possible coordination of an insertion cycle (383). Consistent with this concept, we

recently proposed that the related OB-fold containing KREPA2 (band III, LC-4) directly binds a site for full-round U deletion (442). Additional related subunits of the editing complex were also proposed to conserve an OB-fold. Three subunits including KREPA1 and KREPA2 also bear C_2H_2 zinc-finger domains that could potentially bind single-stranded RNA (458, 461). Furthermore, the RNase III-like insertion endonuclease KREN2 is expected to cleave single-stranded RNA at a single-/double-strand junction, unlike typical RNase III enzymes that cleave dsRNA. KREN2 has one double-stranded RNA binding motif, one RNase III domain, and also a U1-like zinc finger (383). Perhaps one or more of these motifs specifically interact with internal loop determinants involved in scissile-bond selection at insertion sites.

All ribose 2'-deoxy substitutions tested in this study are summarized in Fig. A-18. Notably, proximal changes that significantly decreased pre-mRNA cleavage were also associated with a parallel inhibition of editing complex cross-linking at ES2. Thus, the cross-linking assay we introduced here not only revealed for the first time direct editing complex interactions at an ES for full-round insertion but also can help define ribose 2'-OH groups in cis (pre-mRNA) or in trans (gRNA) that affect the efficiency of both cleavage catalysis and photo-cross-linking at sites targeted by editing complexes.

Work is in progress to identify the photo-cross-linking subunits; however, due to the similar size and gel mobility of several subunits, identification of the cross-linking proteins is not straightforward. We are currently combining the use

of mass spectroscopy techniques and epitope-tagging to produce confirmatory gel-shifts of *Candida*te subunits. Interestingly, the cross-linking pattern at the insertion ES2 and deletion ES1 (442) in A6 model substrates are similar. Whether or not they play a role in the distinction of editing sites is uncertain, although the cross-linking efficiency at ES1 is significantly greater than at ES2. Notably, the former is the most efficient model ES known for *in vitro* editing (392).

Our competition analyses showed a similar response of both cross-linking and pre-mRNA cleavage activities and suggest a preferential association of editing complexes with the A6 substrate. These results also implied a functional relevance of the cross-linking interactions. Surprisingly, a transcript (D33) that does not interfere with cross-linking at an ~200-fold excess (442) cross-linked in its thiolated version with editing complexes. However, as expected, these cross-links were very sensitive to low concentrations of unlabeled D33 (competitor), whereas the similar ES2 and ES1 cross-links were resistant to the same or greater concentrations of D33. This suggests that most if not all crosslinking subunits can make transient nonspecific contacts with RNA, which may be detected by the sensitive 4-thioU photoreagent. In contrast, the associations at ES2 and ES1 are relatively stable. These stabilized interactions, however, are sequence-independent as they are conserved at both ESs tested and may reflect recognition of helical irregularities (e.g. potential ESs) in hybrid substrates. Our laboratory is currently exploring these possibilities.

Materials and methods

Pre-mRNA, gRNA, RNA/DNA Chimeric, and Photoreactive Substrates

The starting substrate in these studies was the minimized ATPase 6 (A6) 45-nt pre-mRNA (445) paired with a variant of the enhanced gRNA gA6[14]USD-3A (447). This substrate directs full-round insertion of 3Us at ES2 and uses pre-edited ES1 to increase the stability of the downstream duplex. RNAs were transcribed from a DNA template as described by Milligan et al. (433) gel-purified, and quantified. Deoxynucleoside-substituted transcripts were made by (IDT, Inc.), and 2'-F and 2'-OCH_3 modified transcripts were by (Dharmacon, Boulder, CO).

Photoreactive Substrates—Each pre-mRNA was obtained by ligation of two pieces (462). The acceptor pieces were synthesized by IDT®, and the thiolated donors were synthesized by Dharmacon®. The donor pieces were radiolabeled to high specific activity with T4 polynucleotide kinase and [γ-32P]ATP (MP Biomedicals) using a 1:2 molar ratio of ends:ATP, gel-purified, and ligated to the acceptor piece as described (442)

Purification of Editing Complexes

Procyclic *T. brucei* strain TREU667 was grown in Cunningham media and mitochondrial extracts were prepared as described (434). Editing complexes were enriched by Q-Sepharose ion exchange chromatography and further

purified by DNA-cellulose affinity chromatography as reported (393, 442). Additional enrichment can be achieved by using another step of Q-Sepharose; however, both editing and cross-linking activities were equivalent in the minimal substrate for full-round insertion (445) by complexes from the two-step and three-step purifications (see Fig. A-19 and data not shown). Fractions with the peak of editing activity were used for all the experiments.

Editing and Cleavage Assays and Quantitation Analysis

Full-round U insertion was performed as described (426). Briefly, a 2-µl mixture with pre-annealed 3'-end-labeled pre-mRNA (~10 fmol) and gRNA (1.2 pmol) was completed to 20 µl with 10 mM MRB buffer (10 mM magnesium acetate, 10 mM KCl, 1 mM EDTA, pH 8, 25 mM Tris-HCl, pH 8, and 5% glycerol), 150 µM UTP, 3 µM ATP, and 2 µl of peak editing fraction. The reaction was incubated at 26 °C for 60 min, deproteinized, and resolved in 9% acrylamide, 7 M urea gels. Editing complexes were pretreated with 10 mM PPi to score total cleavage in absence of RNA ligase activity (429). Neither ATP nor UTP were added to this assay, and the cleavage products were resolved in 15% PAGE with 7 M urea. Ribonuclease T1 and hydroxyl ladders were used to confirm the cleavage at ES2 (not shown). All pre-mRNAs for editing were 3'-end-radiolabeled with [^{32}P]cytidine 3',5'-(bis)phosphate except for the 2'-F-modified transcript (Pair-17), which had to be made with a 3'-terminal deoxynucleoside (Dharmacon). Such a terminus prevents radiolabeling with T4

RNA ligase (454) so this transcript was 5'-end-labeled with T4 polynucleotide kinase. Data were visualized by phosphorimaging and/or x-ray autoradiography, and quantitation was p

erformed using a STORM PhosphorImager (ImageQuant 5.0, GE Healthcare). Each panel in the figures corresponds to one of two replica series performed simultaneously (i.e. one experiment). At least two independent experiments were performed for each figure, and the data shown are representative. The editing activity varied between editosome preparations, but the relative efficiency of the constructs was always consistent. The abundance of accurately edited and cleavage product for each construct tested was initially calculated as the percentage of total input RNA and then normalized to the abundance of the corresponding product by the parental Pair-1 substrate. Mean and error bars were plotted on a linear scale.

Annealing and Photo-cross-linking Assays

The efficiency of pre-mRNA/gRNA annealing was scored in native gels. ES2-radiolabeled mRNA(~10 fmol) and gRNA (1.2 pmol) were pre-annealed in a 2-µl mixture for 10 min at 37 °C followed by 1h at 26 °C, as for editing assays. 20-µl mixtures were completed with 10 mM MRB buffer (see above) containing xylene cyanol and bromphenol blue, incubated for an additional 10 min at 26 °C, loaded directly onto a 6% native PAGE, and resolved at 180 V for 6 h at 4 °C. 0.5 X Tris-borate EDTA buffer and 1 mM $MgCl_2$ were used in both the gel and

running buffer. The photo-cross-linking assays were performed using pre-annealed RNA pairs and under editing reaction conditions (but without nucleotides) as recently reported (442). Also, coimmunoprecipitation and competition analyses were carried out as described (442). All competitor transcripts were supplemented to mixtures and incubated for extra 10 min at 26 °C to allow annealing (i.e. of homologous competitor with free cognate gRNA) before the addition of complexes and irradiation.

APPENDIX IV

DETERMINANTS FOR ASSOCIATION AND gRNA-DIRECTED ENDONUCLEASE CLEAVAGE BY PURIFIED RNA EDITING COMPLEXES FROM *TRYPANOSOMA BRUCEI**

Summary

U-insertion/deletion RNA editing in the single mitochondrion of ancient kinetoplastids is a unique mRNA maturation process needed for translation. Multi-subunit editing complexes recognize many pre-mRNA sites and modify them via cycles of three catalytic steps: guide-RNA (gRNA) mediated cleavage, insertion or deletion of uridylates at the 3' terminus of the upstream cleaved piece, and ligation of the two mRNA pieces. While catalytic and many structural protein subunits of these complexes have been identified, the mechanisms and basic determinants of substrate recognition are still poorly understood.

The current study defined relatively simple single- and double-stranded determinants for association and gRNA-directed cleavage. To this end, we used an electrophoretic mobility shift assay to directly score the association of purified editing complexes with RNA ligands, in parallel with U.V. photocross-linking and

*Reprinted with permission from: "Determinants for association and gRNA-directed endonuclease cleavage by purified RNA editing complexes from *Trypanosoma brucei*" by Alfredo Hernandez, Aswini Panigrahi, Catherine Cifuentes-Rojas, Anastasia Sacharidou, Kenneth Stuart, and Jorge Cruz-Reyes. 2008. J Mol Biol. August 1; Vol. 381 No. 1: 35–48.

functional studies. The cleaved strand required a minimal 5' overhang of 12-nt and a ~15-bp duplex with gRNA to direct the cleavage site. A second protruding element in either the cleaved or the guide strand was required unless longer duplexes were used. Importantly, the single-stranded RNA requirement for association can be upstream or downstream of the duplex, and the binding and cleavage activities of purified editing complexes could be uncoupled. The current observations together with our previous reports (445, 463) show that association, cleavage and full-round editing by purified editing complexes have distinct determinants that increase in complexity as these editing stages progress. Finally, we found that the endonuclease KREN1 in purified complexes photo-crosslinks with a targeted editing site. A model is proposed whereby one or more RNase III-type endonucleases in editing complexes mediate the initial binding and scrutiny of potential ligands, and subsequent catalytic selectivity triggers either insertion or deletion editing enzymes.

Introduction

The majority of primary mRNA transcripts in the single mitochondrion of kinetoplastids, including species of Trypanosoma and Leishmania, are plagued with frame-shifts and stop codons. Protein-encoding sequences are produced via an extraordinary maturation process involving specific insertion and deletion of uridylates at often hundreds of editing sites (ESs) in a single transcript. This

process is catalyzed by megadalton multi-subunit assemblies known as L-complexes, 20S editosomes, or editing complexes that contain between 16 and 20 known subunits and target ESs specified by the partial complementarity of pre-edited mRNA (pre-mRNA) and guide RNAs (gRNAs). For recent reviews, see (464-465).

RNA editing has been recreated *in vitro* at single model ESs in either natural-like (387-388) or completely artificial (463) substrates. Early mechanistic studies indicated that all steps of deletion and insertion editing were catalyzed by distinct enzymatic activities (389, 392, 426, 428-429). More recently, it was shown that a deletion cycle involves the consecutive action of endonuclease KREN1, 3' exo-uridylylase KREX1 and/or KREX2, and ligase KREL1 (399, 403, 418, 429, 466). Similarly, an insertion cycle involves endonuclease REN2 or REN3, terminal uridylyl transferase KRET2, and preferentially, ligase KREL2 (398, 404, 418, 429, 467). Yet, KREL1 may be used in absence of KREL2 *in vitro* and *in vivo* (418, 429, 468-469). Potentially, KREN1 and KREX enzymes could also help proofread misedited insertion ESs bearing extra Us; i.e., mis-edited insertion sites could be targeted and repaired by deletion editing (389). Additional observations also suggest that deletion and insertion activities may occur at individual ESs *in vivo*. Namely, RNAi of KREN1 down-regulates editing of CYb and COII pre-mRNAs *in vivo*, which only contain insertion ESs11. Also, RET2 was shown to add Us at deletion sites *in vitro* (470).

Pre-mRNA/gRNA hybrids are proposed to form two helical regions flanking an internal loop. The downstream (relative to the scissile bond) “anchor” duplex directs endonuclease cleavage immediately 5’ to it, whereas the upstream duplex is thought to tether the cleaved 5’ piece during U-specific processing and re-ligation. The mechanisms of substrate recognition in assembled editing complexes are currently been addressed (for a recent review see (471)).

Previous studies in our laboratory using purified native complexes have shown that secondary structure rather than sequence-specific features are primarily required for full-round insertion editing (445, 463). In a completely artificial 43-nt pre-mRNA/gRNA model substrate with single-helical turns flanking the central loop, simple features of this loop were manipulated to interconvert sites between insertion and deletion editing. Important insights on the specificity of substrate association with purified editing complexes were obtained in competition studies using parallel U.V. photo-crosslinking and full-round catalytic editing assays. Such studies, using a single photo-reactive 4-thioU and a ^{32}P atom at targeted ESs, showed a preferential association of complexes with deletion and insertion substrates, particularly with the most efficient model substrate currently available for full-round editing (A6 pre-mRNA/D33 gRNA hybrid) (392, 463). The native complexes also exhibited a level of non-specific binding to unrelated transcripts. Interestingly, ribose 2’-H substitutions on the downstream helix and gRNA-side of the central loop significantly inhibited both

pre-mRNA cleavage and photo-cross-linking activities at a targeted ES. Furthermore, a single 2'-H substitution adjoining the scissile bond obliterated the endonucleolytic activity but had no effect on photo-crosslinking, suggesting that the ribose 2'-hydroxyl at this position is relevant for catalysis not association of editing complexes (463).

One of the photo-crosslinking subunits in assembled editing complexes was proposed to be KREPA2 (MP63) (442), which as several other subunits, contains conserved domains that predict interaction with nucleic acids (383, 385). Studies of purified recombinant proteins established that KREPA3 (MP42), KREPA4 (MP24) and KREPA6 (MP18) exhibit RNA-binding activity (405, 443, 472), but their precise function in assembled editing complexes remains to be determined. KREPA4 and KREPA6 exhibited preferential binding to poly (U) homopolymers, suggesting a role in the recognition of the natural 3'-poly (U) extension of gRNAs. These recombinant proteins also showed a general low-affinity binding for RNA.

While previous photo-crosslinking analyses provided insights on the specificity of the editing enzyme/substrate association, absence of crosslinking with certain mutant substrates could not be interpreted with certainty. Furthermore, whether purified editing complexes form transient or stable ribonucleoprotein complexes (RNPs) with cognate substrates is unknown. In the current study we used an electrophoretic mobility shift assay (EMSA) to directly examine, for the first time, RNPs formed by purified editing complexes. We

applied EMSA, photo-crosslinking and endonuclease analyses to define substrate determinants for association and endonuclease cleavage, the first catalytic step of RNA editing. Both single-stranded (ssRNA) and double-stranded (dsRNA) RNA were required for these two stages of editing, but ssRNA for association can be satisfied in different ways, whether or not endonuclease cleavage activity is observed. Importantly, the determinants for association and cleavage can be uncoupled, and the determinants for endonuclease cleavage are more complex than for association but less intricate than for full-round editing.

Finally, we compared preparations of native and affinity-purified editing complexes in association and catalytic assays, and established that one subunit that photo-crosslinks at a targeted ES is the essential endonuclease KREN1. The subunit KREPA2 (MP63) was also confirmed to photo-crosslink. A model is proposed whereby recognition of basic determinants including those defined here, leads to a preferential association of editing complexes with potential substrates. Such initial interactions may precede subsequent specialized contacts that trigger catalysis by either deletion or insertion editing.

Results

Our previous RNA-protein photo-crosslinking studies showed that purified native editing complexes preferentially associate with a model A6 substrate for full-round editing (Fig. A-20A) via recognition of secondary structure not

sequence-specific features (442, 463). However, absence of cross-linking due to certain substrate modifications or reaction conditions leaves uncertainties about the editing enzyme/substrate association.

To directly score substrate binding by editing complexes, we established an electrophoretic mobility shift assay (EMSA). A standard reaction mixture for full-round editing or photocrosslinking studies, using purified editing complexes and an ES1-labeled substrate (Fig. A-20A) 24, was briefly incubated and loaded onto a native agarose gel. A fraction of radiolabeled substrate exhibited delayed electrophoretic mobility only in the presence of editing complexes (Fig. A-20B). This shifted product comigrated with complexes that were radiolabeled by adenylylation of ligase subunits (Fig. A-20C) (417) and was specifically immunodepleted by monoclonal antibodies to editing subunits (Fig. A-20D, upper). As expected, adenylylatable editing ligases were enriched in the antibody-conjugated IgG beads but not in beads without antibodies (lower).

To further confirm that these ribonucleoprotein assemblies (RNPs) include editing complexes, we examined their substrate specificity using competition analysis as those performed in photo-crosslinking and full-round editing studies (442). Importantly, the competition profiles in photo-crosslinking (that we reported in (442)) and EMSA assays were equivalent, i.e., the homologous A6 competitor was strongly inhibitory at 5–10 fold excess whereas tRNA and CYb were significantly less inhibitory at 25-fold excess (Figs. A-20E and A-20F, respectively; and data not shown). Moreover, a similar competition

pattern was observed in assays of gRNA-directed endonuclease cleavage, the first enzymatic step of a full-round editing cycle (Fig. A-20G).

Together, these data indicate that the EMSA directly scores the editing enzyme/substrate association and specificity of editing complexes. The data using EMSA also mirrors the observations in parallel studies of RNA-protein photo-crosslinking and editing enzymatic activities. Furthermore, all these activities of editing complexes can be examined using common substrates and reaction conditions. Based on these observations, we sought to define substrate determinants for association and guide-directed cleavage by editing complexes.

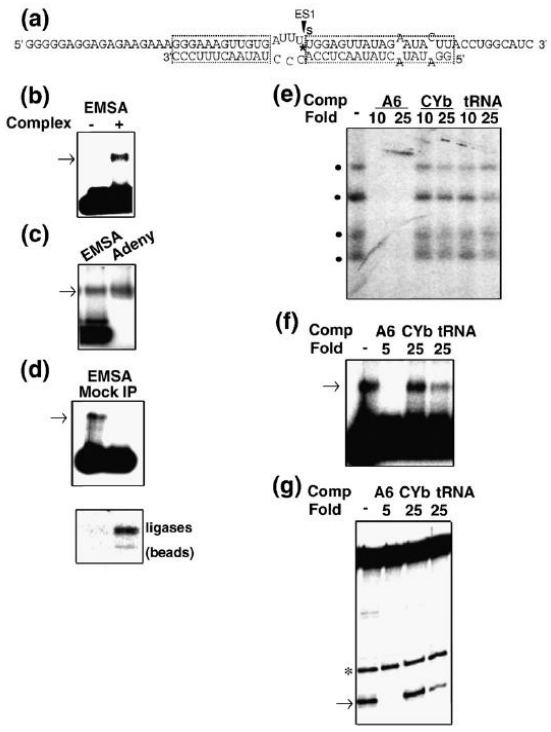


FIGURE A-20. The association of purified editing complexes with substrates can be directly scored by EMSA, in parallel with UV photocrosslinking and pre-mRNA cleavage assays. (a) ES1-32P labeled (*) and thio-labeled (s) model A6 substrate for EMSA, UV photocrosslinking, and full-round U-deletion editing. (b) EMSA in a native agarose gel showing a shifted band (arrow) only in the presence of editing complexes. (c) Co-migration of the shifted substrate with editing complexes that were radiolabeled by auto-adenylylation (Adeny). (d) Specific depletion of the shifted product by immunoprecipitation (IP; upper) and recovery of self-adenylylatable ligase subunits in the beads (lower). A mock reaction devoid of antibodies. (e) Preferential association of editing complexes with a substrate (A6) for full-round editing in competition studies using UV photocrosslinking (dots indicate four major crosslinks) or parallel assays of (f) EMSA and (g) endonuclease cleavage (arrow). A spurious cut (*) serves as loading control. The fold excesses of unlabeled homologous A6 (5- or 10-fold) and heterologous CYb and tRNA competitors are indicated. No competitor is (−).

We performed competition analyses, as in Figs. A-20E–G, to examine the effects of unlabeled derivatives of the homologous (A6) competitor (diagramed in Fig. A-21A). Our standard editing mixtures include gRNA at ~120-fold excess over radiolabeled A6 pre-mRNA to ensure quantitative annealing (463). In order to form competitor duplexes, the abundant free gRNA ("guide strand") in the standard mixture was allowed to pre-anneal with each pre-mRNA derivative ("substrate strand") added at a small, 5–10 fold, excess over radiolabeled pre-mRNA (Fig. A-21B). All constructs in Fig. A-21A used the same guide strand, and quantitative annealing was confirmed in native gels 5 (463) (see methods).

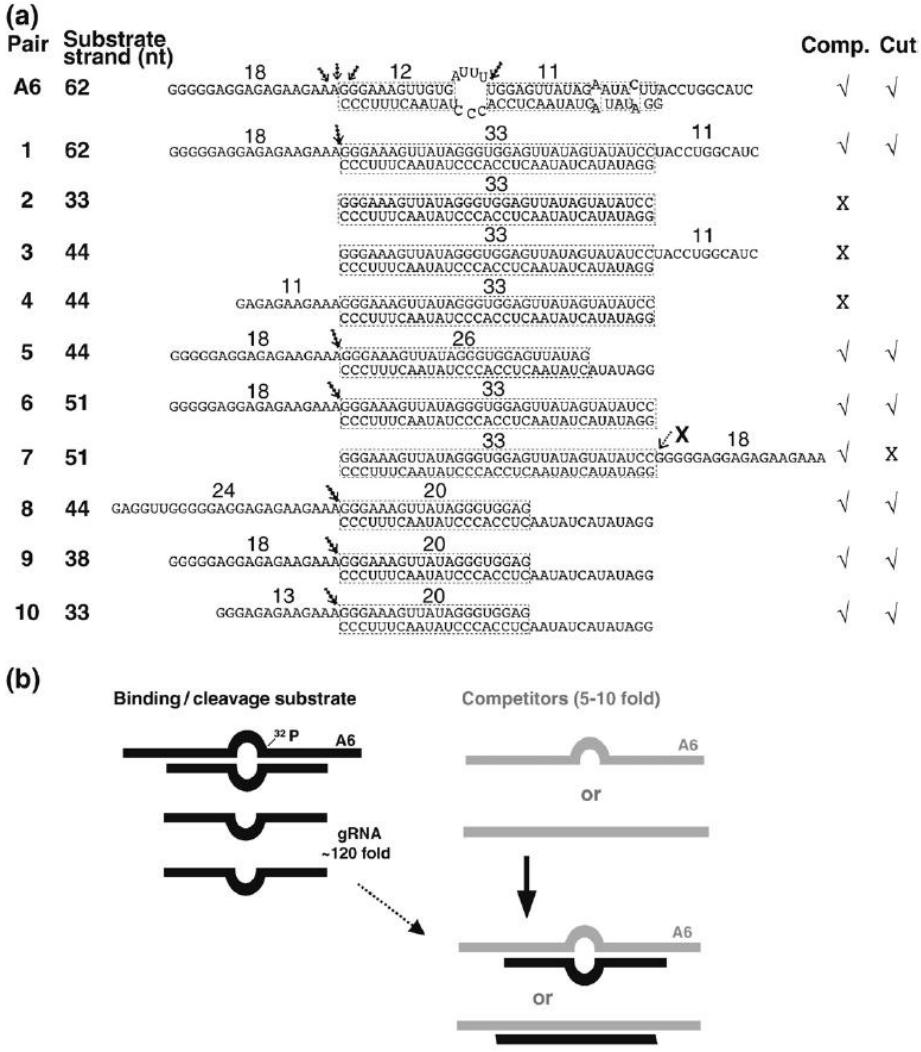


FIGURE A-21. Constructs tested in competition assays. (a) A6 and derivative competitors (top "substrate" strand) paired with gRNA D33 (lower "guide" strand). The assigned number of each competitor RNA pair and size (nt) of the substrate strand are indicated, as well as the size of the predicted helix and overhangs. Cleavage sites are noted with an arrow. Evident (✓) or weak-to-undetected (X) competition and cleavage activity for each construct are indicated at the right. (b) Cartoon of model RNA construct in standard functional (left) and modified competition (right) assays.

Such analysis in binding and catalytic assays performed in parallel is illustrated in Fig. A-22. In this example, both the homologous A6 pair and derivative Pair-1, whose guide strand is fully based paired (i.e., it forms a continuous 33-bp duplex), were strong competitors in photo-crosslinking, EMSA and cleavage assays (Figs. A-22A–C, respectively). However, a second derivative that conserves the 33-bp duplex but lacks overhangs (Pair-2) was a poor competitor in all assays. These data suggest that editing complexes associate with Pair-1 but not Pair-2. Thus the presence or absence of the central loop region in the parental A6 construct does not significantly affect the binding efficiency of editing complexes, although ssRNA seems required for association.

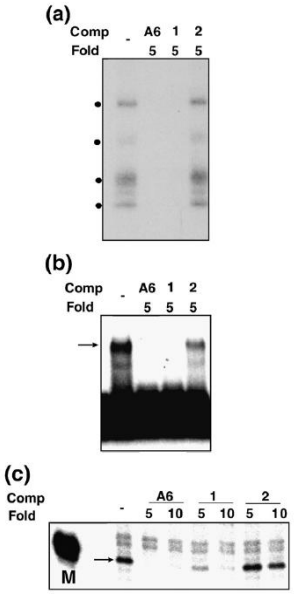


Figure A-22. Parallel competitions in (A) UV photo-cross-linking, (B) EMSA and (C) RNA cleavage assays, as in Fig. A-20. A6 and variants (Comp) diagrammed in Fig. A-21A were examined at the indicated fold-excess. Cleavage assays typically include a size marker (M) such as the ^{32}P-kinased donor fragment used to prepare the parental A6 substrate (Fig. A-20A) or control lanes with and without gRNA.

To dissect RNA requirements for association with editing complexes that distinguish Pair-1 from Pair-2, we designed competitors based on Pair-2 that contain upstream and/or substrate strand downstream overhangs of various lengths (Fig. A-23; diagrammed in Fig. A-21A). While 13-nt, 18-nt and 24-nt extensions favored association of editing complexes (Pairs 5–10), 11-nt extensions at either side of the duplex (Pairs 3–4) did not. Furthermore, constructs with shorter duplexes, 26-bp (Pair-5) and 20-bp long (Pairs 8–10), were also effective competitors. Most of these constructs used a 44-nt substrate strand, however, Pair-10 with a 33-nt substrate strand was also a significant

competitor. Some competitions are more evident in cross-inking and EMSA than in cleavage studies (Figs. A-23A–C). This difference may reflect different dynamics in the assays; that is, the former two score RNP complexes that either are present at the time of cross-inking or that withstand gel electrophoresis, respectively, whereas the latter scores accumulation of cleaved product over time, regardless of the relative stability of RNPs. Together, the competition studies in Fig. A-20–Fig. A-23 suggest that association with editing complexes requires recognition of a relatively simple structure bearing discrete ssRNA and dsRNA determinants.

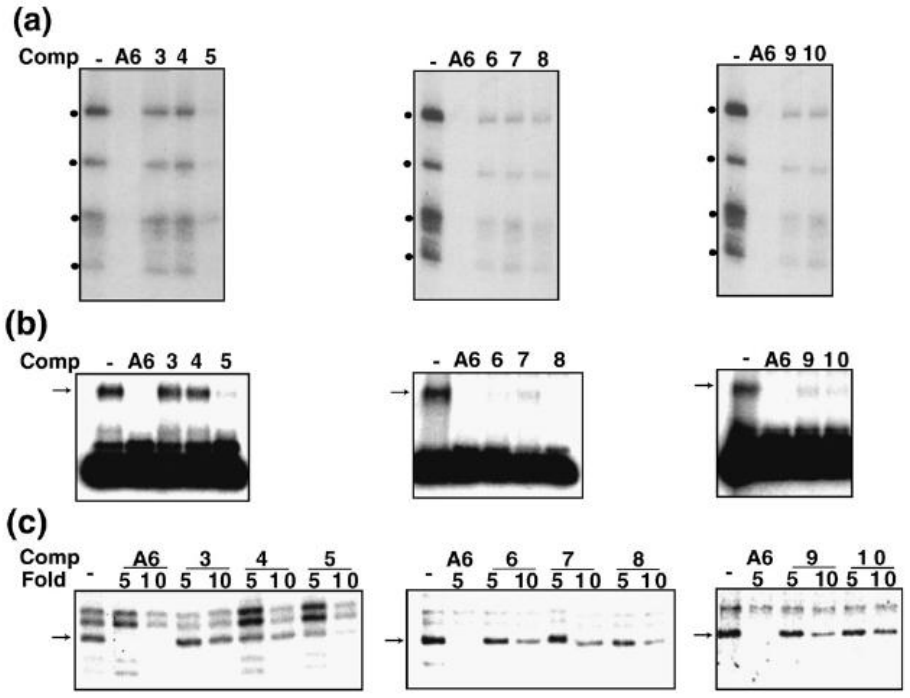


FIGURE A-23. (a–c) Parallel competitions as in Fig. A-22. The homologous A6 and derived competitors are diagrammed in Fig. A-21a.

Several constructs examined so far were effective competitors, indicating that are bound by editing complexes, but it was unclear whether they were also active in enzymatic assays. To directly address this, we tested these constructs for specific gRNA-directed cleavage by editing complexes (Fig. A-24). Since the guide strand in these pairs fully complements the substrate strand we assayed for potential guide-directed cleavage at the phosphodiester bond just 5' of the duplex (388). We have reported that this particular bond is cleaved just 5' of the upstream duplex in the parental construct (Fig. A-21A, top construct; and ahead in Fig. A-24B) (392). Pairs 1, 5 and 6 generated a predicted 18-nt cleaved

product (Fig. A-24A) that corresponds to the 5'-end labeled overhang. This cleavage occurred only in presence of the guide strand. Furthermore, Pairs 8–10 which form a shorter 20-bp duplex were also cleaved with comparable efficiencies to the parental A6 construct (Fig. A-24B). The expected 24-nt, 18-nt and 13-nt cleavage products, respectively, were gRNA-dependent. In the parental A6 construct, gRNA-directed cleavages occur 5' of both downstream (ES1) and upstream duplexes: the 5' end-labeled substrate strand accumulates a 31-nt product, as a result of consecutive cleavage and removal of 3Us by U-specific exonuclease activity at ES1 (399); also, multiple cuts 5' of the upstream duplex are observed probably due to misannealing of this helix. Spurious fragments of the substrate strand often accumulate due to breakage or RNase contamination that preferentially target Us in absence of guide strand, and are more evident with 5' labeled substrates.

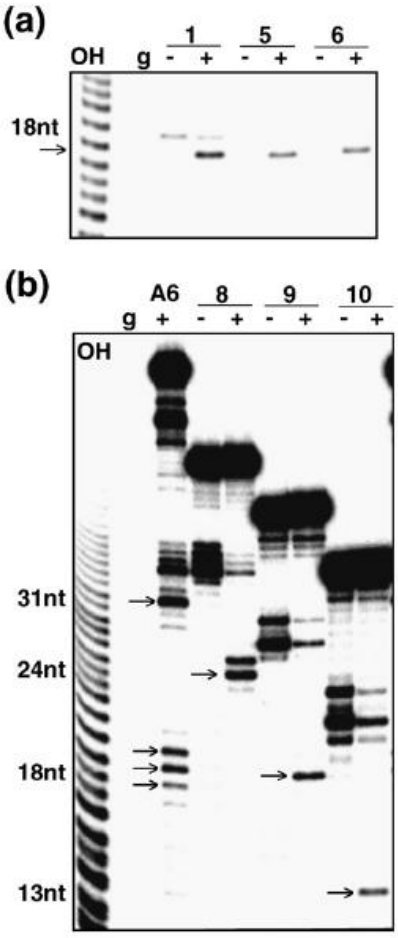


FIGURE A-24. (a and b) Direct cleavage assays of 5′-endlabeled substrate-strand transcripts paired with the parental gRNA.D33. The homologous A6 and derived competitors are diagrammed in Fig. A-21a. Lanes with "+" and without "−" gRNA (g) are shown. Specific cleavage only occurs in the presence of gRNA (marked by an arrow). Spurious fragmentation of these transcripts occurs without gRNA but is inhibited by annealing of gRNA. Partial alkaline RNA hydrolysis "OH" was used as sizing ladder. Guide-directed cleavage of the A6 construct is directed by the downstream duplex (ES1) and by the upstream duplex. The latter occurs at three adjacent positions (∼18-nt products) possibly due to alternative pairing. The short upstream duplex may be stabilized by coaxially stacking with the downstream duplex.

Among constructs found to associate with editing complexes, Pair-7 was not subject to guide-directed endonuclease cleavage as its substrate strand lacks a 5’ overhang and its 3’ ssRNA extension does not undergo cleavage (Fig. A-21A; data not shown). The 18-nt protrusion of Pair-7 rescued the inactive Pairs 2–4 in crosslinking (Fig. A-21A) and EMSA (data not shown) assays.

In summary, all efficient competitors in EMSA and photo-crossliking assays were also functional for endonuclease cleavage, except for Pair-7. While both association and endonuclease cleavage activities of editing complexes have ssRNA and dsRNA requirements, these could be present in a way that promotes association but not cleavage. Thus, association and catalysis by editing complexes can be uncoupled.

We decided to further analyze derivatives of Pair-10, the shortest construct tested that supported editing complex association and specific endonuclease cleavage activity. This symmetric construct with 13-nt overhangs flanking a 20-bp duplex was ideal to dissect determinants involved in selection of the substrate strand. That is, how are the substrate and guide strands distinguished in a duplex? We tested Pair-10 derivatives (Pairs 11–16) bearing progressively shortened 5’-overhangs in the guide strand (Figs. A-25A and ahead in A-25D). In these reduced structures, a 3-nt 5’-overhang in the guide strand promoted efficient cleavage of the substrate strand, but 1-nt and 2-nt extensions were strongly inhibitory (Pairs 15–16). Also, the latter constructs were not rescued by longer (18-nt) 5’-overhangs in the substrate strand (not

shown). This suggests that 5' overhangs in the substrate and guide strands are not compensatory.

Analysis of constructs bearing shorter 5' extensions in the substrate strand (Pairs 17–20; Fig. A-25B) showed that 12-nt are minimally required for endonuclease cleavage activity (Figs. A-25D–E; and data not shown). Constructs with 11-nt 5'-overhangs in the substrate strand were inactive and not rescued by the presence of longer guide-strand overhangs (e.g., Pairs 19–20).

To determine whether constructs with duplexes shorter than 20-bp are functional we examined Pairs-21–26 (Figs. A-25C and A-25F). Efficient endonuclease cleavage was supported by Pair-21, which forms a 15-bp duplex, but progressive truncations of the guide-strand 5' overhang were increasingly inhibitory (Pairs 22–24). Pair-21 also showed that the substrate strand can be shorter that the guide strand, and that a ~27-nt substrate strand bearing a 12-nt 5'-overhang supports efficient endonuclease cleavage. In the above constructs the substrate-strand 5'-extension appears to be separately recognized, as inactivating truncations of this element were not compensated by a longer duplex or extended guide-strand 5'-ssRNA. In contrast, the guide-strand 5'-overhang could be replaced by using either an extended double-stranded terminus (e.g., Pair-6; Fig. A-21A), or a 3'-overhang of the substrate-strand (Pair-25; Fig. A-25C). The latter pair also showed that an 18-nt guide strand, largely annealed with the substrate strand, directs efficient endonuclease cleavage activity. Seiwert et al. reported that an 18-nt guide strand directs

endonuclease cleavage of a complementary 73-nt A6 mRNA (388). Pair-25 and Pair-5, both of which generate the same cleaved product, were nearly as efficient as the parental A6 construct (Fig. A-25G; see also Fig. A-24A). Finally, we found that an 11-bp duplex in Pair-26 failed to direct detectable cleavage of the substrate strand (not shown). Such 11-bp duplex seems relatively stable (−18.4 kcal/mol) and we confirmed efficient annealing with the substrate strand in native gels (463). Although this simple pair is not cleaved, it binds editing complexes in an EMSA (see the site-specific labeled Pair-27 in Fig. A-26A).

Importantly, the ssRNA overhang was essential for binding, and substitution of a paired strand with DNA was inhibitory (Pair-28 and Pair-29, respectively). We examined additional constructs for association, whether or not they are cleaved, (Fig. A-26B). In this case, we prepared derivatives of the thiolated parental A6 (diagrammed in Fig. A-20) and tested their ability to photo-crosslink with editing complexes. For example, Pair-30 photo-crosslinks and is also cleaved (Fig. A-26B; and data not shown). Other derivatives with an ssRNA overhang that crosslinked are not cleaved, whereas a blunt helix did not exhibit detectable crosslinking (Pairs 31–33, respectively). The parental A6 substrate generates more robust signals in association assays that most derivatives tested in our study.

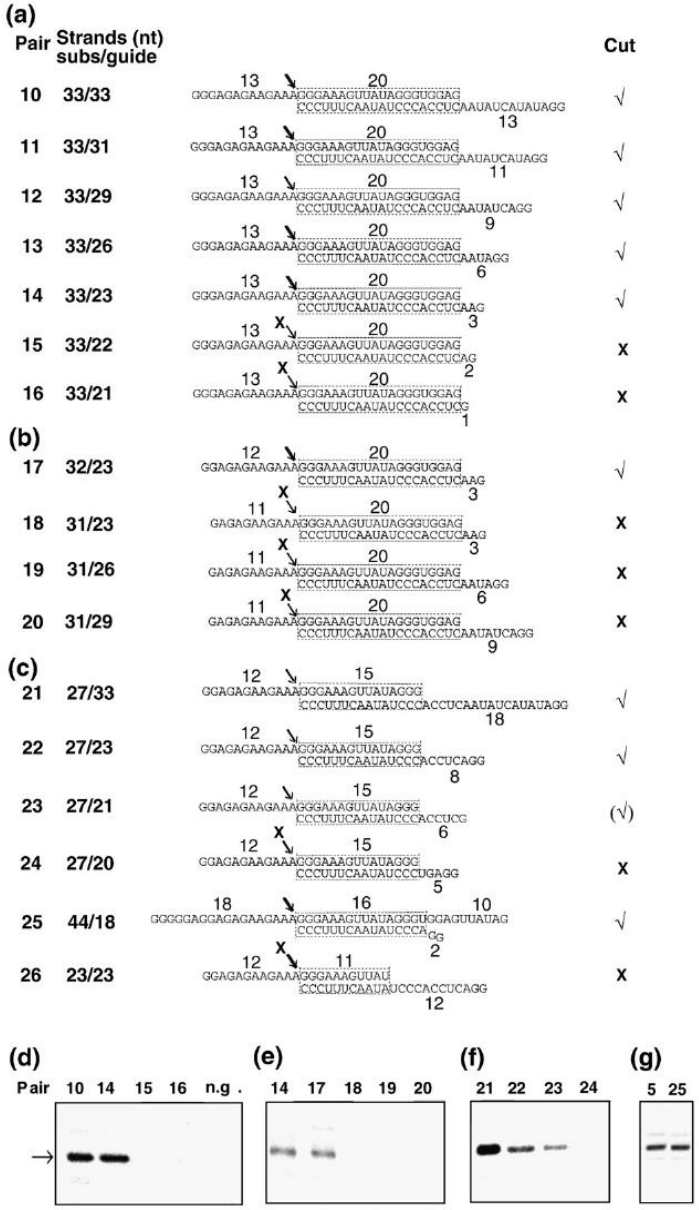


FIGURE A-25. Diagram of minimized substrates for endonuclease cleavage by purified editing complexes. (a–c) A6 and derivative competitors (substrate strand) paired with parental gRNA.D33 or shorter versions (guide strand). The size of both strands in each pair is indicated. All other labels are as those in Fig. A-21a. Detected (✓) or undetected (X) cleavage activity is indicated for each construct. Cleavage activity on Pair 23 was relatively weak. (d–g) Cleavage assays using 3′-end-labeled substrate strand derivatives.

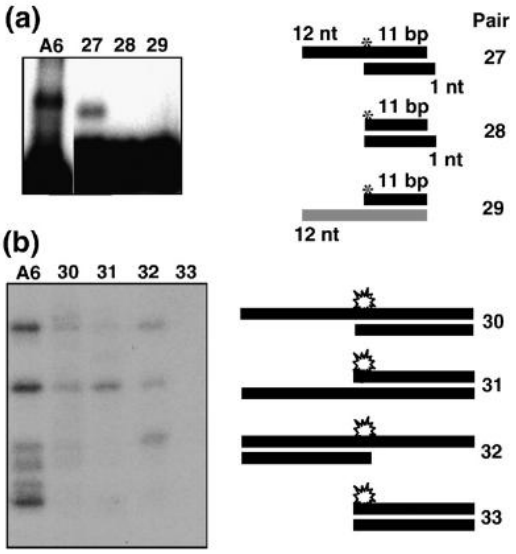


FIGURE A-26. Additional RNA pairs that associate with purified editing complexes but are not cleaved. (a) EMSA; Pair 27 (derived from Pair 26) forms an RNP but is not cleaved. This RNP exhibits a faster electrophoretic mobility compared with the parental A6, but the reason for this is unclear. Duplexes without the 12-nt overhang or bearing a DNA strand failed to form an RNP (Pairs 28 and 29, respectively). (b) UV photocross-linking assays of the A6 parental construct in Fig. A-20 and derivatives with or without an ssRNA overhang (Pairs 30–33). The site-specific ^{32}P label in (a) and the ^{32}P and thio labels in (b) are depicted by an asterisk and a star, respectively.

In summary, the construct series in Fig. A-25 showed that purified editing complexes only cleave substrate strands bearing a minimal 5'-overhang of 12-nt. The minimal duplex directing specific cleavage was not determined to the nucleotide but it could be ~15-bp long, if not smaller. In addition to these two features, cleavage activity required the presence of (a) either a substrate 3'-overhang or a guide 5'-overhang when using a 15-bp duplex, or (b) a larger duplex without additional ssRNA. Fig. A-26 confirmed that association and cleavage can be uncoupled although an ssRNA overhang is essential for both these two stages of editing. Importantly, association exhibits simpler determinants than cleavage.

It is feasible that some if not all determinants defined in the current study may be recognized by one or more RNA-binding subunits of editing complexes, including RNase III-type, OB-fold and zinc-finger domain-bearing subunits. At least three RNase III-type endonucleases identified in editing complexes are thought to catalyze pre-mRNA cleavage in insertion and deletion editing (403-404, 466-467). However, the composition of the native editing complexes used here, including the presence of reported endonucleases, was unclear.

A mass spectrometric analysis of this protein preparation revealed nearly all reported subunits of affinity-purified ~20S editing complexes in *T. brucei* and *L. tarentolae* (403-404, 466), in addition to subunits of the MRP complex which are thought to transiently associate with ~20S editing complexes via an RNA linker (Fig. A-27A) (396). Three other proposed editing subunits, KREPA5,

KREPA6 and KREH1, were not detected likely because they were either sub-stoichiometric, insufficiently ionized in our preparation or absent. However, KREPA6 was recently reported to be essential (472) and was most likely undetected in our samples.

Since our previous photo-croslinking studies indicated that at least four subunits of purified ~20S native complexes make intimate contact with model editing sites (Fig. A-27B) (442, 463) we attempted the identification of a cross-linking subunit that migrates at about 100 kDa, where the endonuclease KREN1 was expected. To this end, we made a TAP-KREN1 construct and expressed it in *T. brucei* procyclic cells (see Materials and Methods section) based on a reported protocol that generated the same cell line (440).

Tagged-editing complexes were purified through IgG and calmodulin-binding peptide (CBP) coupled resins and then examined by photo-crosslinking. We found that CBP-KREN1 complexes produced a shift of the ~100 kDa crosslink due to the mass added by the tag (~5kDa; Fig. A-27C). These complexes also exhibited the crosslink by endogenous KREN1 and the other major crosslinks observed in native complexes. As far as we know this is the first evidence that at least two copies of KREN1 are present in editing complexes. Previous characterization of KREL1, KREN2 and KREN3 (KREPB2) affinity-purified complexes showed that endogenous and ectopic copies of these subunits were also present (396, 418, 440). Importantly, the shifted crosslink is specific of our tagged-KREN1 cell line, and not associated

with the cell culture or protein purification conditions, as affinity-purified complexes using a different tagged subunit (TAP-KREPB5; i.e., MP44) exhibited the same cross-linking pattern of native complexes (Fig. A-27C), as well as a similar silver staining pattern (Fig. A-27D) and full-round insertion and deletion editing activity (not shown).

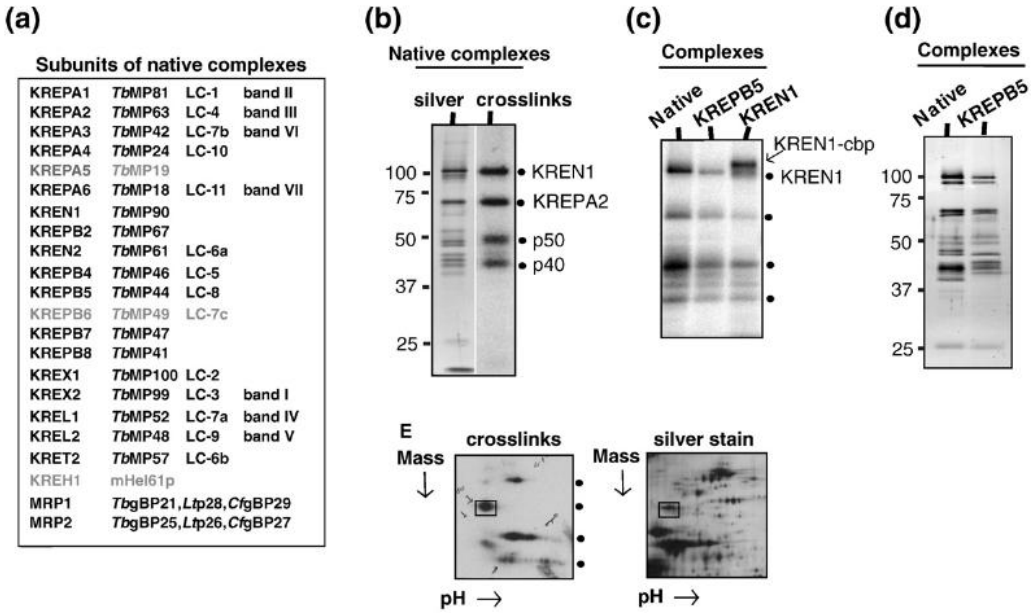


FIGURE A-27. Composition of native editing complexes and identification of two photocross-linking subunits: RNase III-type endonuclease KREN1 and structural KREPA2 (MP63). (a) Listing of subunits detected by mass spectrometry. Alternative nomenclature is indicated. Three subunits were not detected (faded). (b) Native editing complexes stained with silver (lane 1) or exposed onto an X-ray film after UV photocrosslinking (lane 2). The cross-links (dots) by KREN1 and KREPA2 and two more subunits to be identified (p50 and p40) are indicated. (c) Cross-links by native (lane 1) or affinity-purified KREPB5 (MP44) (lane 2) and KREN1 (lane 3) complexes. Both CBP-tagged (upshift) and endogenous KREN1 are indicted. (d) Silver staining of native and affinity-purified KREPB5 complexes. (e) Two-dimensional gel of partially purified complexes after photocrosslinking (left) or silver staining (right). Cross-linked KREPA2 (boxed) was excised from the gel and identified by mass spectrometry.

Consistent with the identification of KREN1 in the current study, our preliminary cross-linking analysis using aliquots of KREN1 and KREN2 complexes purified and characterized in another study (440) showed that the former but not the latter forms the 100-kDa crosslink (data not shown). The presence of these KREN proteins was mutually exclusive in the reported purified complexes (440).

Our previous 1D-analyses suggested that the crosslink at ~60 kDa was KREPA2 (442). We confirmed this identification by performing a 2D-gel analysis of a partially purified protein preparation exhibiting significant cross-linking activity by editing complexes (Fig. A-27D, lower panel). The ~60-kDa crosslink was resolved in a discrete region of the gel, and mass spectrometric analysis of the excised region only contained KREPA2. The cross-linking subunits at about 50 and 40 kDa were more disperse and mass spectrometric analyses of these gel regions were unsuccessful. Thus, they remain to be identified.

Overall, the native editing complexes used in the current studies contain most subunits previously observed in purifications by other labs including the RNase III-type endonuclease KREN1, which we showed directly photo-croslinks with model editing sites. This subunit may be involved in the editing complex recognition of the substrate determinants defined here for association and endonuclease cleavage, but additional work is needed to explore this possibility.

Finally, we compared the substrate specificity of native editing complexes and KREPB5 affinity-purified complexes in parallel EMSA, photo-crosslinking

and endonuclease cleavage assays (Fig. A-27A–C). Native and tagged-KREPB5 editing complexes exhibited similar substrate specificity, in presence of the homologous A6 (5-fold excess) and tRNA (25-fold), as positive control and relatively poor competitors, respectively. Thus, the approaches adopted in these studies should be useful in further comparisons of native and affinity-purified editing RNPs that exhibit different protein and functional composition.

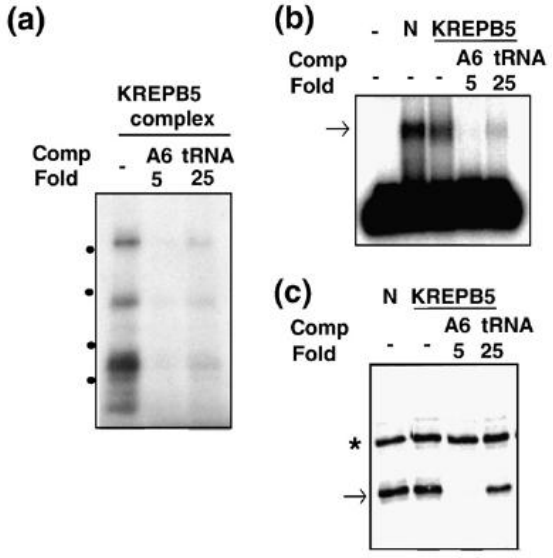


FIGURE A-28. Association and endonuclease cleavage activity of affinity-purified editing complexes. Parallel (a) photocross-linking assay, (b) EMSA, and (c) cleavage assay. All labels are as those in Fig. A-20. KREPB5-tagged complexes were directly compared with native “N” complexes.

Discussion

The goal of the current work was to define substrate requirements for association of purified editing complexes and gRNA-directed cleavage, the first catalytic step of an editing cycle. To this end, we used an EMSA, for the first time, in parallel with U.V. photo-crosslinking and gRNA-directed cleavage assays. Importantly, these assays were performed under comparable reaction conditions and the data obtained was complementary. The RNP assemblies detected by EMSA contained adenylylatable ligases and co-immunoprecipitated with known editing subunits (Figs. A-20B–D), and their substrate specificity was conserved in the association and catalytic assays (Figs. A-20E–G). Our combined EMSA, photo-crosslinking and enzymatic studies defined ssRNA and dsRNA determinants for association and cleavage, summarized in Fig. A-21. Three main combinations of ssRNA and dsRNA determinants supported endonuclease cleavage are represented by the following pairs: Pair-22 (27-nt substrate and 23nt guide strands) exhibits minimal 5' overhangs and ~15-bp duplex for cleavage. In this context, a 12-nt 5' overhang in the substrate strand was minimally required, whereas truncations of the 8-nt 5' overhang in the guide strand were gradually inhibitory. The size of one overhang did not compensate for the size of the other, and thus appear to involve separate recognitions. Pair-25, a long substrate-strand annealed to a minimal guide-strand of 18-nt (16-nt in a duplex) supports efficient cleavage. This confirms the observation by Seiwert et al., that an 18-nt guide strand directed endonuclease cleavage of a

complementary 73-nt A6 mRNA 3 (388). Thus, a substrate 3'-overhang can substitute for a guide 5'-overhang. Pair-6, a long duplex overrides a requirement for ssRNA rightward of the duplex. Thus neither these overhangs are essential but an ssRNA extension, abutting a short duplex, may suffice. In this type of construct, the size of the substrate 5'-overhang was also tested. 12-nt or more supported cleavage (e.g., Pair-6, and data not shown) but 11-nt was inactivating (i.e., Pair-4; data not shown). Additional pairs were bound but not cleaved by editing complexes, showing that these two aspects of editing can be uncoupled. Pair-27 is the simplest construct of this kind. Competition studies or straight association assays by crosslinking or EMSA showed that pairs bearing blunt-ended helices or insufficient ssRNA cannot associate with editing complexes. Pair-2 and Pair-28 reproducibly failed to form detectable RNPs and Pair-3 was significantly less effective than the parental A6 substrate (data not shown). Some constructs that bind but are not cleaved were examined by photo-croslinking or EMSA using 5'-end labeled rather than more sensitive site-specific labeled RNAs (Fig. A-26B; and data not shown).

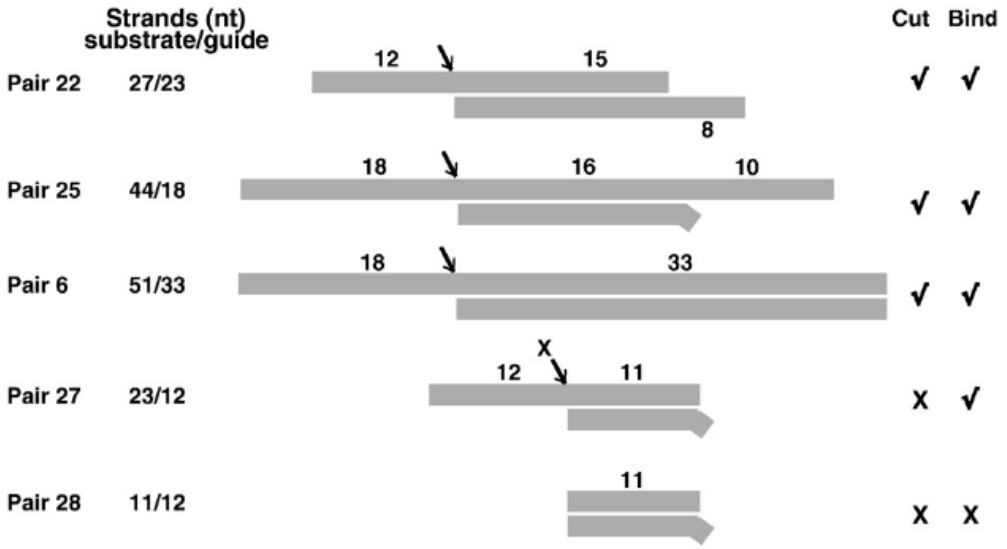


FIGURE A-29. Summary of defined ssRNA and dsRNA determinants for endonuclease cleavage and association by purified editing complexes. Important variations were observed depending on the secondary structure context. Three main types of cleaved constructs are illustrated by Pair 22: It bears minimal substrate 5′ and guide 3′ overhangs. In this context, further shortening of either element was strongly inhibitory and not rescued by lengthening of the other. Pair 25: Its long substrate strand allowed reducing the guide strand to 18 nt. Thus, a substrate 3′ overhang can substitute for a guide 5′ overhang. Pair 6: Its long duplex can substitute for either substrate 3′ overhang or guide 5′ overhang. Thus, neither of these overhangs is essential but one may suffice in cleaved constructs. Importantly, association can occur without cleavage although it also requires an essential overhang either upstream or downstream of the helix. This is illustrated by Pair 27 and Pair 28. Detected (✓) or undetected (X) binding (bind) and cleavage (cut) are indicated.

Together, these constructs as well as others examined indicated that an appropriate combination of dsRNA and ssRNA determinants, rather than overall size of the bi-molecular structure, is required for both association and endonuclease cleavage by purified editing complexes. The ssRNA requirement (12-nt) 5' of the scissile bond and the dsRNA/ssRNA combinations 3' of it seem to involve separate recognitions. The smallest helix tested that directed endonuclease cleavage was 15-bp long (~1.5 helices) but shorter versions similar to Pair-26 may be feasible (Fig. A-25C). Although the shortest functional guide strand tested was 18-nt long, a functional guide strand may be longer than the substrate strand (e.g., Pair-21).

Importantly, the requirements for association and for catalysis can be uncoupled. This was shown by Pair-7 (Fig. A-21), Pair-27 and A6 thio-lated derivatives (Fig. A-26) that bind editing complexes but are not cleaved. In Pair-7, the substrate-strand forms a 3'-overhang but not a 5'-overhang. Its substrate-strand 3' ssRNA stimulates association (compare with the inactive Pair-2) but, as expected, is not cleaved since editing endonucleases specifically target the phosphodiester bond immediately 5' of the guiding "anchor" duplex (388-389). On the other hand Pair-27 bears the critical 12-nt 5'-overhang but either insufficient duplex or overall length for cleavage. Furthermore, while a substrate 5' 12-nt overhang is minimally required for cleavage, whether all residues need be unpaired or some may partially complement apposing guide-strand residues was not determined. In full-round editing substrates, single-strandedness of

residues near the downstream “anchor” duplex is strongly stimulatory. More distal residues can engage in formation of a proposed upstream “tether” duplex in deletion or insertion *in vitro* (392, 444). Furthermore, the presence and/or the nature of gRNA residues in the internal loop may stimulate full-round editing. Consistent with this idea, the lack or inappropriate number of such residues inhibited full-round deletion and insertion editing (392, 444) (and unpublished data), and 2’-deoxy substitutions on the gRNA-side of the internal loop inhibited both photo-crosslinking and cleavage at the scissile bond (463).

Previously, our lab defined a minimal 43-nt pre-mRNA/gRNA hybrid for efficient full-round editing, which formed 10-bp helices flanking the ES. These nearby helices may be stabilized by coaxial stacking interactions, resembling a continuous helix. The smaller hybrid for endonuclease cleavage activity (including a ~27-nt substrate strand) and even simpler structure for binding imply that editing complexes require more extensive RNA contacts for the complete editing reaction, than for the intermediate cleavage step and the initial association step. Consistent with this concept, the artificially enhanced A6 parental substrate (392) is more efficient in all EMSA, photo-crosslinking and cleavage assays than most simpler derivatives tested here.

The fact that only one shifted product is reproducibly detected in the EMSA of the constructs examined suggests binding by a single editing complex, whether dimeric or of higher-order composition consistent with the co-purification of endogenous and ectopically expressed editing subunits, i.e., KREN1 in the

current study (Fig. A-27C) and KREN2, KREN3 (KREPB2) and KREL1 in previous studies (395-396, 418). A mass spectrometric analysis revealed that the native complexes used in this study contain most known subunits of catalytic ~20S editing complexes, as expected from similar biochemical purifications (473). In addition, we found subunits of the MRP sub-complex as it was reported in purified L-complexes (385, 396), suggesting that at least some purified assemblies represent holoenzyme rather than core complexes. Several observations lead us to suggest that some if not all determinants defined in this study may be recognized by one or more RNase III-type proteins. Namely, (a) the shortest duplex tested that directed efficient endonuclease cleavage activity spanned ~1.5 turns. This is also the size of the smallest substrate identified that binds bacterial RNase III (457); (b) the critical role of 5' and 3' overhangs for cleavage at ssRNA-dsRNA junctions by the RNase III family member Drosha (474), and (c) the fact that KREN1 photo-crosslinks with a site for full-round editing (Fig. A-27). This photo-crosslink was defined at a deletion site (Fig. A-27C) but most likely also corresponds to a co-migrating crosslink at insertion sites (463). KREN1 endonuclease was proposed to specifically cleave deletion sites (403), however since association and cleavage are uncoupled we propose a model whereby KREN1 and related RNA-binding subunits may help scrutinize potential ligand determinants in the earliest checkpoint of RNA editing. Subsequent to the binding step, catalytic selectivity based on additional specific substrate recognitions may activate either the deletion or insertion enzymes,

including proofreading of mis-edited insertion sites. A role of REN1 in an early checkpoint of ligand binding may explain why KREN1 down-regulation inhibits editing of CYb and COII pre-mRNAs *in vivo*, which only have insertion sites. It is known that bacterial RNase III can undertake a modulatory role as a general dsRNA-binding protein regardless of its catalytic action (475). Importantly, the crosslinking activity of KREN1, KREPA2 (MP63) and at least two other major cross-linking subunits is conserved in both native and tap-tagged affinity-purified complexes. Such conservation further suggests that the interactions are relevant, and independent of purification protocols and cell lines used. The conserved OB-fold and zinc fingers of KREPA2 may also be involved in recognition of single-stranded determinants defined here.

Finally, while the current study shows that RNPs formed by purified editing complexes can be directly visualized, it is currently unclear if the fraction of substrate that remains unbound in association assays reflects the concentration and/or affinity of either total complexes or functional complexes. Also, not all RNPs formed in solution may be stable enough to withstand the forces of gel electrophoresis. These and related questions will be addressed in separate studies.

Methods

Synthesis and labeling of RNA

The ES1-radiolabeled A6 mRNA substrate was prepared by splint ligation as described (442). All other RNAs were synthesized *in vitro* by the Uhlenbeck single-stranded enzymatic method (433) and gel-purified.

For the preparation of 5'-end labeled substrates, gel-purified RNA was dephosphorylated by treatment by alkaline phosphatase at 37°C for one hour, followed by addition of SDS, EDTA and proteinase K to a final concentration of 1.5%, 5 mM, and 40 µg/mL, respectively and additional incubation at 50°C for 30 minutes. RNA was purified by phenol/chloroform extraction and precipitated with ethanol. 5 pmols of dephosphorylated RNA were incubated at 37°C for 30 minutes with [γ-^{32}P] ATP (1:2 ratio of 5'-ends to ATP) and T4 polynucleotide kinase and gel-purified. For 3'-end labeling, 5 pmoles of gel-purified RNA were incubated at 4°C for 12 hours with an equimolar amount of [5'-^{32}P] Cytidine 3', 5'-Bis (Phosphate) and 15 units of T4 RNA ligase in RNA ligase buffer (50 mM Tris-HCl pH 7.5, 10 mM $MgCl_2$, 10 µg/mL BSA, 50 µM ATP, 10 mM DTT, 2U/µL anti-RNase (Ambion), 10% DMSO) and gel-purified.

Cloning, cell culture and transfection

ORFs were amplified from *T. brucei* genomic DNA, kindly provided by Larry Simpson. The primers for KREN1 were designed as reported (440). For KREPB5, the primers were: forward CCC aagctt ATG AGA CGG GCT GTG

GTA CTC CGT AC; and reverse CGC ggatcc CCG CCC TCC CAG TGC CAG CGC AAC TA (Hind III and Bam HI sites are in small case letters, respectively). The amplified products using Pfu DNA polymerase were treated with HindIII, BamHI and ligated to the pLEW79TAP expression vector, kindly provided by Achim Schnaufer. Constructs were linearized with NotI and used to transfect *T. brucei* strain 29.13 as described (476). Selection of transfectants was applied with 2.5 µg/mL phleomycin. KREN1 and KREPB5 expression was induced with 100 ng/mL and 1 µg/mL tetracycline, respectively, and confirmed by immunoblotting with the PAP reagent (Sigma).

Purification and protein composition determination of Editing Complexes

Chromatographic purification of RNA editing complexes

Mitochondrial extracts were prepared from procyclic *T. brucei* strain TREU667 as described (414, 434). Editing complexes were purified from mitochondrial extracts by consecutive anion exchange and DNA-affinity chromatography as described (393, 414).

Tandem affinity purification of RNA editing complexes

Four liters of culture at a density of ~2.0×107 cells/mL were pelleted and lysed in 25 mL of 10 mM Tris-HCl, pH 8.0, 150 mM KCl, 0.1% NP-40, 1% Triton-X-100 and one tablet of EDTA-free complete protease inhibitors (Roche) for 30 minutes on ice. Lysis was confirmed by microscopy. Lysates were spun at

6000 X g for 15 minutes and the clarified extract purified by sequential IgG and Calmodulin affinity chromatography as described (477).

Mass spectrometric analysis of native RNA editing complexes

Proteins in gel bands and complex mixtures were identified by LC-MS/MS analysis as described (473).

Photo-crosslinking, RNA cleavage, Electrophoretic Mobility Shift, Competitions and Adenylylation assays

All assays are variations of the standard editing assay in our lab which consists of a mixture of a pre-annealed mixture of 10 fmols ^{32}P-labeled RNA and 1.25 pmoles unlabeled gRNA, completed to 20 µL with MRB [25 mM Tris-HCl, pH 8, 10 mM $Mg(OAc)_2$, 10 mM KCl, 1 mM EDTA, pH 8, 50 µg/mL hexokinase and 5% glycerol] and, if applicable, competitor RNA at the indicated molar excess relative to the ^{32}P-labeled substrate. The mixture was pre-equilibrated for 10 minutes at 26°C and 2 µL of peak editing or TAP fraction was added. Prior to the assays, quantitative annealing of the RNA pairs tested was confirmed in native gels [as in (463)]. The sample was incubated at 26°C for 10 minutes then treated in an assay-specific manner. For cross-linking, samples were irradiated for 10 minutes under a 365 nm UV lamp, treated with RNase A and RNase T1 (50 µg/mL and 125 units/mL final concentrations respectively) at 37°C for 15 minutes, supplemented with SDS loading dye and loaded onto an SDS-polyacrylamide gel. For mRNA cleavage, purified editing complexes were

pre-treated with 10 mM tetrapotassium pyrophosphate, pH 8, in MRB, for 5 minutes on ice to inhibit ligase activity (429); after incubation the mixture was deproteinized and RNAs were resolved on denaturing polyacrylamide gels. For electrophoretic mobility shift assays, the reaction mixture was loaded directly (no loading dye) onto a 1.5% agarose gel in 0.5 X TBE (45 mM Tris-borate and 1 mM EDTA) and run for 2 hours at ~5 V/cm at 4°C. Following electrophoresis, the agarose gel was dried under vacuum. EMSA with site-specific labeled transcripts were significantly more sensitive and reproducible than with end-labeled substrates, since the splint-ligation method used to generate the former (see above) exclusively incorporates phosphorylated fragments. Only the parental A6 substrate and Pair-27 were site-specific labeled using synthetic donor fragments [e.g., as in Fig. A-20; (442)], although 5'-end labeled A6 parental and other constructs were also compared side-by-side in shift assays. Immunodepletions were carried out as described for the immunoprecipitation of RNA cross-linking proteins (442) using a monoclonal antibody against KREPA2 immobilized on goat anti-mouse IgG resin (Dynal). Adenylylation assays were performed as described (417). All assays can be scaled-up linearly to enhance signal. The data were reproducible in at least two independent experiments. Each experiment included repeat assays, and those shown are representative. Data were visualized by phosphorimaging and/or autoradiography.

APPENDIX V

APPENDIX REFERENCES

383. Stuart KD, Schnaufer A, Ernst NL, & Panigrahi AK (2005) Complex management: RNA editing in trypanosomes. *Trends Biochem Sci* 30(2):97-105.

384. Madison-Antenucci S, Grams J, & Hajduk SL (2002) Editing machines: the complexities of trypanosome RNA editing. *Cell* 108(4):435-438.

385. Simpson L, Aphasizhev R, Gao G, & Kang X (2004) Mitochondrial proteins and complexes in Leishmania and Trypanosoma involved in U-insertion/deletion RNA editing. *RNA* 10(2):159-170.

386. Blum B, Bakalara N, & Simpson L (1990) A model for RNA editing in kinetoplastid mitochondria: "guide" RNA molecules transcribed from maxicircle DNA provide the edited information. *Cell* 60(2):189-198.

387. Kable ML, Seiwert SD, Heidmann S, & Stuart K (1996) RNA editing: a mechanism for gRNA-specified uridylate insertion into precursor mRNA. *Science* 273(5279):1189-1195.

388. Seiwert SD, Heidmann S, & Stuart K (1996) Direct visualization of uridylate deletion in vitro suggests a mechanism for kinetoplastid RNA editing. *Cell* 84(6):831-841.

389. Cruz-Reyes J & Sollner-Webb B (1996) Trypanosome U-deletional RNA editing involves guide RNA-directed endonuclease cleavage, terminal U exonuclease, and RNA ligase activities. *Proc Natl Acad Sci U S A* 93(17):8901-8906.

390. Blom D*, et al.* (2001) Cloning and characterization of two guide RNA-binding proteins from mitochondria of Crithidia fasciculata: gBP27, a novel protein, and gBP29, the orthologue of Trypanosoma brucei gBP21. *Nucleic Acids Res* 29(14):2950-2962.

391. Muller UF, Lambert L, & Goringer HU (2001) Annealing of RNA editing substrates facilitated by guide RNA-binding protein gBP21. *EMBO J* 20(6):1394-1404.

392. Cruz-Reyes J, Zhelonkina A, Rusche L, & Sollner-Webb B (2001) Trypanosome RNA editing: simple guide RNA features enhance U deletion 100-fold. *Mol Cell Biol* 21(3):884-892.

393. Rusche LN, Cruz-Reyes J, Piller KJ, & Sollner-Webb B (1997) Purification of a functional enzymatic editing complex from Trypanosoma brucei mitochondria. *EMBO J* 16(13):4069-4081.

394. Panigrahi AK*, et al.* (2001) Four related proteins of the Trypanosoma brucei RNA editing complex. *Mol Cell Biol* 21(20):6833-6840.

395. Panigrahi AK*, et al.* (2003) Identification of novel components of Trypanosoma brucei editosomes. *RNA* 9(4):484-492.

396. Aphasizhev R*, et al.* (2003) Isolation of a U-insertion/deletion editing complex from Leishmania tarentolae mitochondria. *EMBO J* 22(4):913-924.

397. Law JA, Huang CE, O'Hearn SF, & Sollner-Webb B (2005) In Trypanosoma brucei RNA editing, band II enables recognition specifically at each step of the U insertion cycle. *Mol Cell Biol* 25(7):2785-2794.

398. Ernst NL*, et al.* (2003) TbMP57 is a 3' terminal uridylyl transferase (TUTase) of the Trypanosoma brucei editosome. *Mol Cell* 11(6):1525-1536.

399. Kang X*, et al.* (2005) Reconstitution of uridine-deletion precleaved RNA editing with two recombinant enzymes. *Proc Natl Acad Sci U S A* 102(4):1017-1022.

400. McManus MT, Shimamura M, Grams J, & Hajduk SL (2001) Identification of candidate mitochondrial RNA editing ligases from Trypanosoma brucei. *RNA* 7(2):167-175.

401. Rusche LN*, et al.* (2001) The two RNA ligases of the Trypanosoma brucei RNA editing complex: cloning the essential band IV gene and identifying the band V gene. *Mol Cell Biol* 21(4):979-989.

402. Schnaufer A*, et al.* (2001) An RNA ligase essential for RNA editing and survival of the bloodstream form of Trypanosoma brucei. *Science* 291(5511):2159-2162.

403. Trotter JR, Ernst NL, Carnes J, Panicucci B, & Stuart K (2005) A deletion site editing endonuclease in Trypanosoma brucei. *Mol Cell* 20(3):403-412.

404. Carnes J, Trotter JR, Ernst NL, Steinberg A, & Stuart K (2005) An essential RNase III insertion editing endonuclease in Trypanosoma brucei. *Proc Natl Acad Sci U S A* 102(46):16614-16619.

405. Brecht M*, et al.* (2005) TbMP42, a protein component of the RNA editing complex in African trypanosomes, has endo-exoribonuclease activity. *Mol Cell* 17(5):621-630.

406. Aphasizhev R, Aphasizheva I, Nelson RE, & Simpson L (2003) A 100-kD complex of two RNA-binding proteins from mitochondria of Leishmania tarentolae catalyzes RNA annealing and interacts with several RNA editing components. *RNA* 9(1):62-76.

407. Vondruskova E*, et al.* (2005) RNA interference analyses suggest a transcript-specific regulatory role for mitochondrial RNA-binding proteins MRP1 and MRP2 in RNA editing and other RNA processing in Trypanosoma brucei. *J Biol Chem* 280(4):2429-2438.

408. Pelletier M & Read LK (2003) RBP16 is a multifunctional gene regulatory protein involved in editing and stabilization of specific mitochondrial mRNAs in Trypanosoma brucei. *RNA* 9(4):457-468.

409. Missel A, Souza AE, Norskau G, & Goringer HU (1997) Disruption of a gene encoding a novel mitochondrial DEAD-box protein in Trypanosoma brucei affects edited mRNAs. *Mol Cell Biol* 17(9):4895-4903.

410. Madison-Antenucci S, Sabatini RS, Pollard VW, & Hajduk SL (1998) Kinetoplastid RNA-editing-associated protein 1 (REAP-1): a novel editing complex protein with repetitive domains. *EMBO J* 17(21):6368-6376.

411. Vanhamme L*, et al.* (1998) Trypanosoma brucei TBRGG1, a mitochondrial oligo(U)-binding protein that co-localizes with an in vitro RNA editing activity. *J Biol Chem* 273(34):21825-21833.

412. Allen TE*, et al.* (1998) Association of guide RNA binding protein gBP21 with active RNA editing complexes in Trypanosoma brucei. *Mol Cell Biol* 18(10):6014-6022.

413. Pai RD, Oppegard LM, & Connell GJ (2003) Sequence and structural requirements for optimal guide RNA-directed insertional editing within Leishmania tarentolae. *RNA* 9(4):469-483.

414. Sollner-Webb B, Rusche LN, & Cruz-Reyes J (2001) Ribonuclease activities of trypanosome RNA editing complex directed to cleave specifically at a chosen site. *Methods Enzymol* 341:154-174.

415. Oppegard LM & Connell GJ (2002) Direct visualisation of RNA editing within a Leishmania tarentolae mitochondrial extract. *Int J Parasitol* 32(7):859-866.

416. Panigrahi AK*, et al.* (2001) Association of two novel proteins, TbMP52 and TbMP48, with the Trypanosoma brucei RNA editing complex. *Mol Cell Biol* 21(2):380-389.

417. Sabatini R & Hajduk SL (1995) RNA ligase and its involvement in guide RNA/mRNA chimera formation. Evidence for a cleavage-ligation mechanism of Trypanosoma brucei mRNA editing. *J Biol Chem* 270(13):7233-7240.

418. Schnaufer A*, et al.* (2003) Separate insertion and deletion subcomplexes of the Trypanosoma brucei RNA editing complex. *Mol Cell* 12(2):307-319.

419. Huang CE, O'Hearn SF, & Sollner-Webb B (2002) Assembly and function of the RNA editing complex in Trypanosoma brucei requires band III protein. *Mol Cell Biol* 22(9):3194-3203.

420. Kang X*, et al.* (2004) Disruption of the zinc finger motifs in the Leishmania tarentolae LC-4 (=TbMP63) L-complex editing protein affects the stability of the L-complex. *J Biol Chem* 279(6):3893-3899.

421. Suck D (1997) Common fold, common function, common origin? *Nat Struct Biol* 4(3):161-165.

422. Leung SS & Koslowsky DJ (2001) Interactions of mRNAs and gRNAs involved in trypanosome mitochondrial RNA editing: structure probing of an mRNA bound to its cognate gRNA. *RNA* 7(12):1803-1816.

423. Leung SS & Koslowsky DJ (2001) RNA editing in Trypanosoma brucei: characterization of gRNA U-tail interactions with partially edited mRNA substrates. *Nucleic Acids Res* 29(3):703-709.

424. Oppegard LM, Hillestad M, McCarthy RT, Pai RD, & Connell GJ (2003) Cis-acting elements stimulating kinetoplastid guide RNA-directed editing. *J Biol Chem* 278(51):51167-51175.

425. Halbig K, De Nova-Ocampo M, & Cruz-Reyes J (2004) Complete cycles of bloodstream trypanosome RNA editing in vitro. *RNA* 10(6):914-920.

426. Cruz-Reyes J, Rusche LN, & Sollner-Webb B (1998) Trypanosoma brucei U insertion and U deletion activities co-purify with an enzymatic editing complex but are differentially optimized. *Nucleic Acids Res* 26(16):3634-3639.

427. Cruz-Reyes J, Rusche LN, Piller KJ, & Sollner-Webb B (1998) T. brucei RNA editing: adenosine nucleotides inversely affect U-deletion and U-insertion reactions at mRNA cleavage. *Mol Cell* 1(3):401-409.

428. Huang CE*, et al.* (2001) Roles for ligases in the RNA editing complex of Trypanosoma brucei: band IV is needed for U-deletion and RNA repair. *EMBO J* 20(17):4694-4703.

429. Cruz-Reyes J, Zhelonkina AG, Huang CE, & Sollner-Webb B (2002) Distinct functions of two RNA ligases in active Trypanosoma brucei RNA editing complexes. *Mol Cell Biol* 22(13):4652-4660.

430. Koslowsky DJ & Yahampath G (1997) Mitochondrial mRNA 3' cleavage/polyadenylation and RNA editing in Trypanosoma brucei are independent events. *Mol Biochem Parasitol* 90(1):81-94.

431. Grams J, McManus MT, & Hajduk SL (2000) Processing of polycistronic guide RNAs is associated with RNA editing complexes in Trypanosoma brucei. *EMBO J* 19(20):5525-5532.

432. Reichert VL, Le Hir H, Jurica MS, & Moore MJ (2002) 5' exon interactions within the human spliceosome establish a framework for exon junction complex structure and assembly. *Genes Dev* 16(21):2778-2791.

433. Milligan JF, Groebe DR, Witherell GW, & Uhlenbeck OC (1987) Oligoribonucleotide synthesis using T7 RNA polymerase and synthetic DNA templates. *Nucleic Acids Res* 15(21):8783-8798.

434. Harris ME & Hajduk SL (1992) Kinetoplastid RNA editing: in vitro formation of cytochrome b gRNA-mRNA chimeras from synthetic substrate RNAs. *Cell* 68(6):1091-1099.

435. Hema M & Kao CC (2004) Template sequence near the initiation nucleotide can modulate brome mosaic virus RNA accumulation in plant protoplasts. *J Virol* 78(3):1169-1180.

436. Koslowsky DJ, Reifur L, Yu LE, & Chen W (2004) Evidence for U-tail stabilization of gRNA/mRNA interactions in kinetoplastid RNA editing. *RNA Biol* 1(1):28-34.

437. Blaszczyk J*, et al.* (2001) Crystallographic and modeling studies of RNase III suggest a mechanism for double-stranded RNA cleavage. *Structure* 9(12):1225-1236.

438. Leung SS & Koslowsky DJ (1999) Mapping contacts between gRNA and mRNA in trypanosome RNA editing. *Nucleic Acids Res* 27(3):778-787.

439. Aphasizhev R, Aphasizheva I, & Simpson L (2003) A tale of two TUTases. *P Natl Acad Sci USA* 100(19):10617-10622.

440. Panigrahi AK*, et al.* (2006) Compositionally and functionally distinct editosomes in Trypanosoma brucei. *RNA* 12(6):1038-1049.

441. Miller MM, Halbig K, Cruz-Reyes J, & Read LK (2006) RBP16 stimulates trypanosome RNA editing in vitro at an early step in the editing reaction. *RNA* 12(7):1292-1303.

442. Sacharidou A*, et al.* (2006) RNA editing complex interactions with a site for full-round U deletion in Trypanosoma brucei. *RNA* 12(7):1219-1228.

443. Salavati R*, et al.* (2006) KREPA4, an RNA binding protein essential for editosome integrity and survival of Trypanosoma brucei. *RNA* 12(5):819-831.

444. Igo RP, Jr., Lawson SD, & Stuart K (2002) RNA sequence and base pairing effects on insertion editing in Trypanosoma brucei. *Mol Cell Biol* 22(5):1567-1576.

445. Cifuentes-Rojas C, Halbig K, Sacharidou A, De Nova-Ocampo M, & Cruz-Reyes J (2005) Minimal pre-mRNA substrates with natural and converted sites for full-round U insertion and U deletion RNA editing in trypanosomes. *Nucleic Acids Res* 33(20):6610-6620.

446. Lawson SD, Igo RP, Jr., Salavati R, & Stuart KD (2001) The specificity of nucleotide removal during RNA editing in Trypanosoma brucei. *RNA* 7(12):1793-1802.

447. Igo RP, Jr., Palazzo SS, Burgess ML, Panigrahi AK, & Stuart K (2000) Uridylate addition and RNA ligation contribute to the specificity of kinetoplastid insertion RNA editing. *Mol Cell Biol* 20(22):8447-8457.

448. Kapushoc ST & Simpson L (1999) In vitro uridine insertion RNA editing mediated by cis-acting guide RNAs. *RNA* 5(5):656-669.

449. Kabb AL, Oppegard LM, McKenzie BA, & Connell GJ (2001) A mRNA determinant of gRNA-directed kinetoplastid editing. *Nucleic Acids Res* 29(12):2575-2580.

450. Oppegard LM, Kabb AL, & Connell GJ (2000) Activation of guide RNA-directed editing of a cytochrome b mRNA. *J Biol Chem* 275(43):33911-33919.

451. Igo RP, Jr.*, et al.* (2002) Role of uridylate-specific exoribonuclease activity in Trypanosoma brucei RNA editing. *Eukaryot Cell* 1(1):112-118.

452. Yu LE & Koslowsky DJ (2006) Interactions of mRNAs and gRNAs involved in trypanosome mitochondrial RNA editing: structure probing of a gRNA bound to its cognate mRNA. *RNA* 12(6):1050-1060.

453. Arnott S (1970) The geometry of nucleic acids. *Prog Biophys Mol Biol* 21:265-319.

454. Nandakumar J & Shuman S (2004) How an RNA ligase discriminates RNA versus DNA damage. *Mol Cell* 16(2):211-221.

455. Madison-Antenucci S & Hajduk SL (2001) RNA editing-associated protein 1 is an RNA binding protein with specificity for preedited mRNA. *Mol Cell* 7(4):879-886.

456. Tian B, Bevilacqua PC, Diegelman-Parente A, & Mathews MB (2004) The double-stranded-RNA-binding motif: interference and much more. *Nat Rev Mol Cell Biol* 5(12):1013-1023.

457. Gan J*, et al.* (2006) Structural insight into the mechanism of double-stranded RNA processing by ribonuclease III. *Cell* 124(2):355-366.

458. Stefl R, Skrisovska L, & Allain FH (2005) RNA sequence- and shape-dependent recognition by proteins in the ribonucleoprotein particle. *EMBO Rep* 6(1):33-38.

459. Wang AH*, et al.* (1982) Molecular structure of r(GCG)d(TATACGC): a DNA--RNA hybrid helix joined to double helical DNA. *Nature* 299(5884):601-604.

460. Messias AC & Sattler M (2004) Structural basis of single-stranded RNA recognition. *Acc Chem Res* 37(5):279-287.

461. Hudson BP, Martinez-Yamout MA, Dyson HJ, & Wright PE (2004) Recognition of the mRNA AU-rich element by the zinc finger domain of TIS11d. *Nat Struct Mol Biol* 11(3):257-264.

462. Moore MJ & Sharp PA (1992) Site-specific modification of pre-mRNA: the 2'-hydroxyl groups at the splice sites. *Science* 256(5059):992-997.

463. Cifuentes-Rojas C*, et al.* (2007) Substrate determinants for RNA editing and editing complex interactions at a site for full-round U insertion. *J Biol Chem* 282(7):4265-4276.

464. Cruz-Reyes J. & Hernandez A (2008) Protein-protein and RNA-protein interactions in U-insertion/deletion RNA editing complexes. *RNA and DNA Editing*, ed Smith HC (John Wiley & Sons, Inc., New Jersey), pp 71–98.

465. Carnes J & Stuart KD (2007) Uridine insertion/deletion editing activities. *Methods Enzymol* 424:25-54.

466. Kang X*, et al.* (2006) Reconstitution of full-round uridine-deletion RNA editing with three recombinant proteins. *Proc Natl Acad Sci U S A* 103(38):13944-13949.

467. Carnes J, Trotter JR, Peltan A, Fleck M, & Stuart K (2008) RNA editing in Trypanosoma brucei requires three different editosomes. *Mol Cell Biol* 28(1):122-130.

468. Gao G & Simpson L (2003) Is the Trypanosoma brucei REL1 RNA ligase specific for U-deletion RNA editing, and is the REL2 RNA ligase specific for U-insertion editing? *J Biol Chem* 278(30):27570-27574.

469. O'Hearn SF, Huang CE, Hemann M, Zhelonkina A, & Sollner-Webb B (2003) Trypanosoma brucei RNA editing complex: band II is structurally critical and maintains band V ligase, which is nonessential. *Mol Cell Biol* 23(21):7909-7919.

470. Zhelonkina AG*, et al.* (2006) T. brucei RNA editing: action of the U-insertional TUTase within a U-deletion cycle. *RNA* 12(3):476-487.

471. Cruz-Reyes J (2007) RNA-protein interactions in assembled editing complexes in trypanosomes. *Methods Enzymol* 424:107-125.

472. Tarun SZ, Jr.*, et al.* (2008) KREPA6 is an RNA-binding protein essential for editosome integrity and survival of Trypanosoma brucei. *RNA* 14(2):347-358.

473. Panigrahi AK, Schnaufer A, & Stuart KD (2007) Isolation and compositional analysis of trypanosomatid editosomes. *Methods Enzymol* 424:3-24.

474. Zeng Y & Cullen BR (2005) Efficient processing of primary microRNA hairpins by Drosha requires flanking nonstructured RNA sequences. *J Biol Chem* 280(30):27595-27603.

475. Calin-Jageman I & Nicholson AW (2003) RNA structure-dependent uncoupling of substrate recognition and cleavage by Escherichia coli ribonuclease III. *Nucleic Acids Res* 31(9):2381-2392.

476. Wirtz E, Leal S, Ochatt C, & Cross GA (1999) A tightly regulated inducible expression system for conditional gene knock-outs and dominant-negative genetics in Trypanosoma brucei. *Mol Biochem Parasitol* 99(1):89-101.

477. Rigaut G*, et al.* (1999) A generic protein purification method for protein complex characterization and proteome exploration. *Nat Biotechnol* 17(10):1030-1032.

VITA

Name: Catherine Cifuentes Rojas

Address: Texas A&M University

Department of Biochemistry and Biophysics

College Station, TX, 77843-2128

Email Address: cathycifr@tamu.edu

Education: Ph.D., Genetics, Texas A&M University, 2010

M.Sc., Microbiology (emphasis Immunology), Pontificia Universidad Javeriana, Bogotá, Colombia, 2004

B.Sc., Bacteriology and Clinical Laboratory, Universidad del Valle, Cali, Colombia, 2000