BIOGENESIS OF TELOMERASE RNA IN FISSION YEAST

BY

Wen Tang

B.S., Wuhan University, 2007

Submitted to the graduate degree program in Molecular & Integrative Physiology and the Graduate Faculty of the University of Kansas medical center in partial fulfillment of the requirements for the degree of Doctor of Philosophy

Committee members:

Peter Baumann (Advisor)

Peter Baumann (Co-chairperson)

Steven M. LeVine (Co-chairperson)

Joan W. Conaway

Christophe P. Nicot

Kausik Si

Date defended: 7/26/2012

The Dissertation Committee for Wen Tang certifies

that this is the approved version of the following dissertation:

BIOGENESIS OF TELOMERASE RNA IN FISSION YEAST

Peter Baumann (Advisor)

Steven M. LeVine (Co-chairperson)

Date approved: 8/10/2012

Abstract

Telomerase, a reverse transcriptase, counteracts the progressive loss of chromosome-terminal DNA sequences in most eukaryotes. Work from C. W. Greider and E. H. Blackburn first revealed telomerase activity in Tetrahymena extracts. Subsequent work demonstrated telomerase is a multisubunit ribonucleoprotein complex which uses part of an RNA subunit as a template to synthesize telomeric DNA. Most cancer cells show high telomerase activity, while telomerase insufficiency due to mutations in telomerase components leads to several degenerative syndromes including dyskeratosis congenita, aplastic anaemia, and idiopathic pulmonary fibrosis. Given the therapeutic value of modulating telomerase activity, it is important to study its assembly, regulation, and enzymatic action in human and model systems. We have identified the RNA subunit of telomerase (TER1) in fission yeast, and we showed that the mature 3' end of TER1 is generated by the spliceosome in a reaction ("slicing") akin to the first step of splicing. Through examining a putative Sm protein binding site that partially overlaps the 5' splicing site and thus is located at 3' end of mature TER1, we found that the canonical Sm complex and Lsm2-8 (Sm-like) complex sequentially bind to TER1 and play distinct roles on telomerase RNA biogenesis. Sm and Lsm proteins belong to an ancient family of RNA binding proteins represented in all three domains of life. They form multimeric complexes on specific sets of non-coding RNAs and play critical roles in their biogenesis, function and degradation. The Sm complex specifically

binds to the TER1 precursor, promotes spliceosomal cleavage, and facilitates trimethyguanosine (TMG) cap formation at its 5' end. At later stages, the Lsm2-8 complex replaces the Sm complex and binds to the majority of TER1. The Lsm complex protects mature TER1 from exonucleolytic degradation and promotes catalytic subunit to bind to TER1. Our findings provide new insights into telomerase biogenesis by defining roles for Sm and Lsm complexes as well as the TMG cap. Also our results constitute the first identification of RNA whose biogenesis requires both the Sm and Lsm2-8 complexes.

Acknowledgements

The preparation of my dissertation provides me an opportunity to recall and appreciate the contribution of many individuals to the completion of my PhD study. Although here I could not thank everyone made my past five years so rewarding and joyful, the following people deserve special recognition.

First of all, I must thank my wife Di. Your love, kindness, and unconditional support made my life so enjoyable and memorable. I only hope that I can return the favor during this lifetime that we share.

I am grateful for the insights and encouragement that I received from my parents. You have allowed these five years being about personal growth, rather than merely intellectual growth. Although you probably cannot understand a single word in this document, here I have to say I love you forever.

My advisors deserve special thanks. First, I would like to thank my mentor Dr. Peter Baumann. I appreciate your guidance and support on every aspect of my projects. I also wish to thank the following brilliant scientists for their advice and mentorship during the past four years: Joan W. Conaway, Steven M. LeVine, Christophe Nicot, Kausik Si, and Marco Blanchette. All I can say is that your efforts made me the best I can possibly be.

Thanks to every lab member in the Baumann lab, I could not have completed my PhD studies without your contributions. I must thank Ram Kannan for his direct contribution for my project, Jessica, Xiaorong and Jeremy for patiently teaching me key techniques when I joined the lab, Rachel for proofreading my

the comprehensive exam proposal and manuscript, and every other person in the lab for keeping such a great environment for doing science.

Finally, I would like to acknowledge my friends and colleagues both from University of Kansas Medical Center and Stowers Institute. It is nice to have you around and thanks for everything you have done.

Table of Contents

List of Figures

List of Appendices

Chapter I. Introduction

I.1: Introduction to telomere

I.1.a: The end protection problem

The main goal of cell division is to accurately transfer genomic information from a cell to its progenies. Cellular metabolites and exogenous DNA damaging agents including radiation and chemicals could damage DNA, thus threatening genomic integrity. One type of DNA damage particularly harmful to cells is DNA double strand breaks (DSB)s. DSBs are often recognized and dealt by the DNA damage response which triggers a signaling cascade activating cell cycle checkpoints to arrest cell cycle and repair the breaks (Sancar et al., 2004). The linear nature of chromosomes in eukaryotes poses a difficult situation for cells. How can the DNA damage response differentiate natural ends of chromosomes from DSBs? Studies of transposition events in maize and irradiated *Drosophila melanogaster*, pioneered by McClintock and Muller, respectively, showed that DSBs can be efficiently repaired while naturally occurring chromosome termini escape from this process (McClintock, 1941; Muller and Altenburg, 1930). This observation led them to postulate that chromosome ends have special qualities named as “telomeres” by Muller (Fig. 1.1). Subsequent work from Elizabeth Blackburn and Joseph Gall lab confirmed the existence of such special structures at the ends of chromosomes (Blackburn and Gall, 1978; Szostak and Blackburn, 1982)

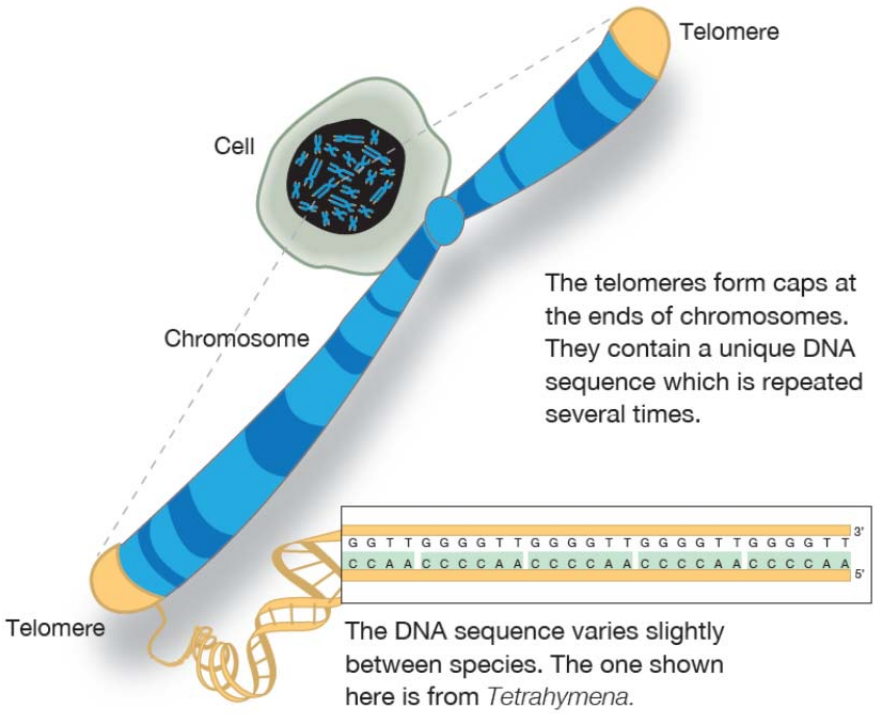

Figure 1.1: Schematic of chromosome and telomere in cells. Adapted from Press release for the Nobel Prize in Physiology or Medicine in 2009.

I.1.b: Telomere structure

Telomeres are conserved nucleo-protein complexes which consist of repeated DNA elements terminating in a single-stranded overhang, and specific DNA binding proteins (Henderson and Blackburn, 1989; Klobutcher et al., 1981; Palm and de Lange, 2008). The length of telomeric repeat tracts varies dramatically among different species. It ranges from less than 100 nucleotides in ciliates to tens of kilobases in mammals. The overhang extends from approximate 14 nucleotides in yeast (Larrivee et al., 2004) and ciliate

(Klobutcher et al., 1981) to over 100 nucleotides in human (Chai et al., 2006; Huffman et al., 2000; McElligott and Wellinger, 1997).

Human telomeric DNA is bound by the shelterin complex, which consists of Telomeric Repeat binding Factor 1 and 2 (TRF1 and TRF2) (de Lange, 2002, 2009), single strand telomere factor Protection of Telomere (POT1) (Baumann and Cech, 2001), TRF1-Interacting Nuclear protein 2 (TIN2) (Kim et al., 2004), Rap1(Li et al., 2000), and TPP1 (Liu et al., 2004a; Ye et al., 2004b) (Fig. 1.2). Shelterin proteins have distinct roles in maintaining telomeres. TRF1 promotes efficient replication of repeats and prevents fork stalling (Sfeir et al., 2009). TIN2 acts as a bridge to connect different shelterin proteins (Kim et al., 2004; Ye et al., 2004a). TRF2 and Rap1 efficiently inhibit non-homologous end joining at telomeres (Bae and Baumann, 2007; Sarthy et al., 2009; van Steensel et al., 1998). Pot1 and TPP1 stimulate processivity of telomerase in telomere extension (Wang et al., 2007; Xin et al., 2007). The proper maintenance of telomeres requires additional accessory factors including DNA-PKcs, Ku70, PARP-1, and WRN (Blasco, 2005).

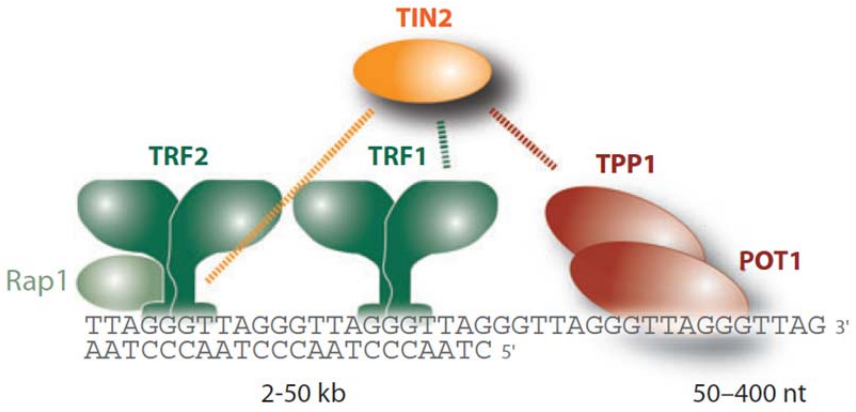

Figure 1.2: Schematic of human telomere. The figure listed the major telomere binding proteins and the size of telomeric DNA. Adapted from (Palm and de Lange, 2008).

I.1.c: The end replication problem

DNA polymerases use an RNA primer to synthesize DNA in the 5’ to 3’ direction. The RNA primer at the 5’ end is excised, generating a gap usually filled from adjacent Okazaki fragments. However, when the terminal 5’ RNA primer is removed, no upstream lagging strand DNA is available to initiate “fill in” synthesis. This leads to formation of a single-stranded DNA end in the lagging-strand (Fig. 1.3a). Thus, the lagging strand ends of each chromosome are shorter than their templates. The replication of the leading-strand is not expected to result in DNA loss, but the blunt end creates other problems for chromosome protection. Single-stranded overhangs are needed for recognition by single strand DNA-binding proteins that protect the ends from degradation and fusion. In addition, the presence of the 3' overhang is essential for telomere elongation mediated by telomerase. Neither single strand DNA binding proteins nor telomerase can act on blunt ends. In fact, blunt ends are further processed

to generate G overhangs (Fig. 1.3b). Due to incomplete replication of genomic DNA by DNA polymerase (Olovnikov, 1973; Watson, 1972) and nucleolytic processing (Lydall, 2003; Sfeir et al., 2005), telomeres are shortened during each round of cell division. This progressive deterioration of telomere is thought to act as a molecular clock that counts the number of times a cell has divided (Allsopp et al., 1995; Allsopp et al., 1992; Chang and Harley, 1995). Telomere shortening to a critical length leads to the loss of chromosome protection or loss of some essential genes. Under such circumstance, normal cells often enter into a nondividing state called replicative senescence (Bodnar et al., 1998). This “end replication problem” (Figure 1.3), as was independently hypothesized by Jim Watson and A. Olovnikov in the 1970s, is one of the most fundamental and interesting problems in biology.

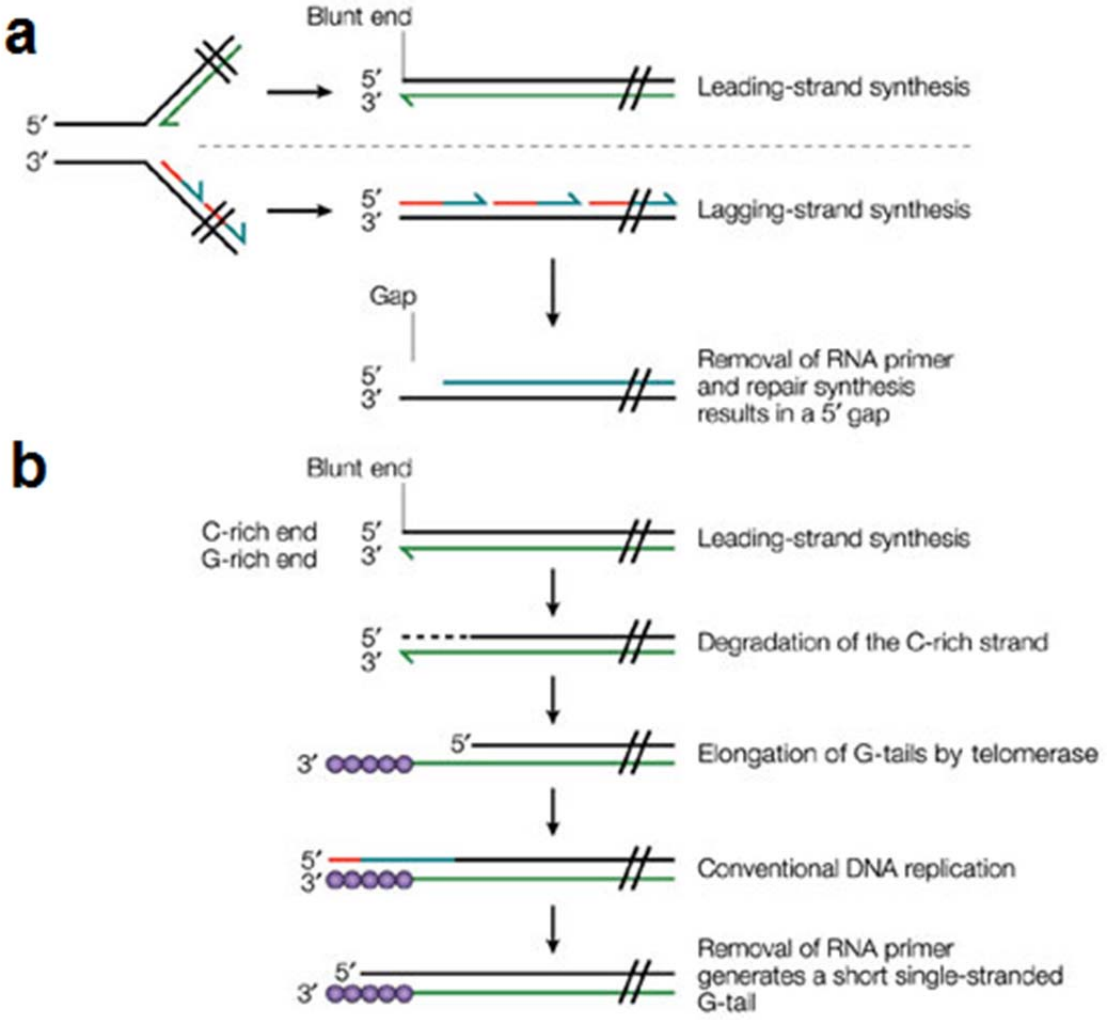

Figure 1.3: Semi-conservative DNA synthesis presents an end-replication problem. a, Schematic of the “end replication problem”. b, Generation of single-stranded overhangs and elongation by telomerase counteracts the “end replication problem”. Figures are adapted from (Vega et al., 2003).

I.2: Introduction to telomerase

Telomere sequence loss can be counteracted by de novo extension through a eukaryotic reverse transcriptase called telomerase (Greider and Blackburn, 1985, 1987, 1989). Core components of telomerase are the protein telomerase reverse transcriptase (TERT) and the telomerase RNA subunit (TER). TERT

harbors a central region that has homology with the active-site motifs of viral RT enzymes (Lingner et al., 1997b). The RNA component has a short template sequence that specifies telomere repeats and is important for the enzymatic activity of telomerase (Greider and Blackburn, 1989; Yu et al., 1990). Telomerase RNAs vary remarkably across eukaryotic species. Its size ranges from about 150 nt in ciliates to over 1000 nt in fungi (Chen et al., 2000; Lin et al., 2004). Comparison of telomerase RNAs from mammals, yeasts, and ciliates reveals no obvious sequence similarity (Chen and Greider, 2004). Instead, they share a common core structure that includes a long-range base pairing interaction enclosing the template and an adjacent pseudoknot structure (Lin et al., 2004). Additional telomerase components have been identified using biochemical and genetic approaches. For example, a La motif protein called p65 was identified by immune-purification (Witkin and Collins, 2004). Certain hnRNP proteins and ATPases are found to be present in active human telomerase (Fu and Collins, 2007; Venteicher et al., 2008). Genetic screens in budding yeast led to the identification of the EST (ever shorter telomeres) genes that, when mutated, display progressive telomere shortening phenotypes (Lendvay et al., 1996). Those proteins play various roles in telomerase biogenesis and affect stability, trafficking, and reverse transcriptase activity. Interestingly, during the time course of evolution different organisms seem to have acquired distinct accessory proteins besides the core components and

distinct maturation pathways. Here I will discuss the telomerase biogenesis pathway in three well-studied model organisms.

I.2.a: Telomerase biogenesis pathway in ciliates

Telomerase activity was first identified in *Tetrahymena* extracts (Greider and Blackburn, 1985). Subsequent biochemical characterization revealed that telomerase contains an essential RNA subunit, a portion of which serves as template to specify telomere sequences (Greider and Blackburn, 1987). The RNA component is an RNA polymerase III transcript terminating at a stretch of uridines (Greider and Blackburn, 1987; McCormick-Graham and Romero, 1995). Oligonucleotide-based telomerase purification from *Euplotes aediculatus* identified the catalytic subunit TERT (Lingner et al., 1997b) and a La motif protein named p43 (Lingner and Cech, 1996). Affinity purification of epitope-tagged TERT from *Tetrahymena* recovered a La motif protein called p65 which is most likely the ortholog of p43 in *Euplotes aediculatus* (Witkin and Collins, 2004). Both p43 and p65 are associated with active telomerase. More importantly, antibodies against p43/p65 can efficiently immunodeplete telomerase RNA and telomerase activity, suggesting the majority of telomerase contains those proteins (Aigner et al., 2000; Witkin et al., 2007). Additional proteins including p20, p45, and p75 are co-immunoprecipitated with the catalytic subunit, although their functions in telomerase are not well understood (Witkin and Collins, 2004).

Further characterization of p65 revealed its important roles in telomerase assembly. Recombinant p65 can directly bind to 3' stem of telomerase RNA and facilitate RNA association with TERT *in vitro* (Prathapam et al., 2005). Biochemical and single molecule studies demonstrated that p65 induces conformational change of telomerase RNA which in turn directs binding of TERT to form the ternary complex (Berman et al., 2010; Stone et al., 2007) (Fig. 1.4). It will be interesting to test whether p43 perform the same role in *Euplotes aediculatus* in the future.

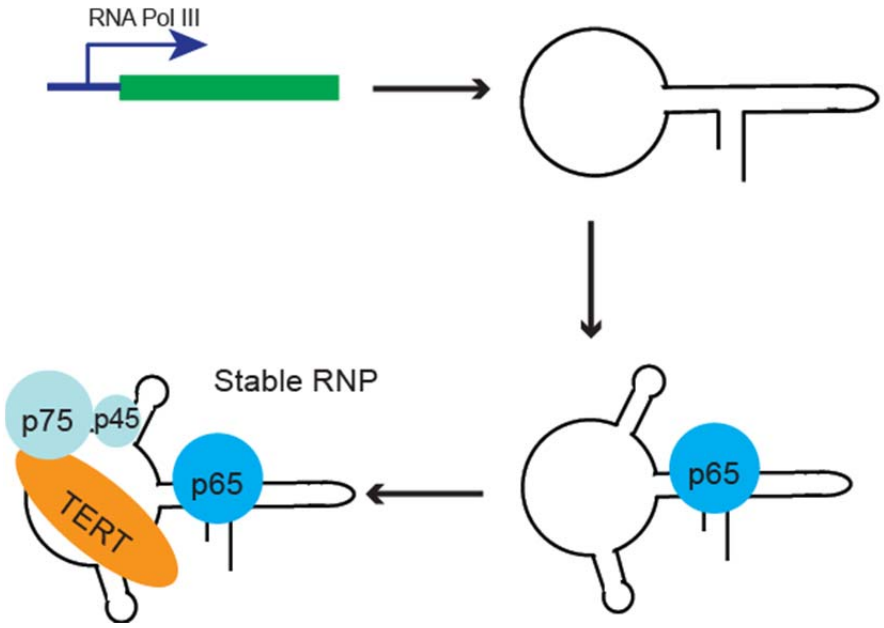

Figure 1.4: Telomerase biogenesis pathway in ciliates. Telomerase RNA is transcribed by RNA polymerase III and first bound by p65 which facilitates binding of the catalytic subunit.

I.2.b: Telomerase biogenesis pathway in yeast

Telomerase RNA in *Saccharomyces cerevisiae*, called TLC1, was first recovered from a screen to identify suppressors of telomeric silencing twenty years ago (Singer and Gottschling, 1994), while telomerase RNA from *Schizosaccharomyces pombe* (*S. pombe*) has been characterized using biochemical approaches only in recent years (Leonardi et al., 2008; Webb and Zakian, 2008). Both telomerase RNAs share some hallmarks of small nuclear RNA (snRNA), including a Sm protein binding site (Sm site) and a 5' TMG cap. Mutations in the Sm site result in reduced levels of telomerase RNA and telomere shortening (Box et al., 2008; Leonardi et al., 2008; Seto et al., 1999). In both yeast species, telomerase RNA is synthesized by RNA polymerase II, polyadenylated, but further processed to remove the poly (A) tail (Chapon et al., 1997; Leonardi et al., 2008). In *Saccharomyces cerevisiae* (*S. cerevisiae*), the removal of poly (A) tail is mediated by RNA-binding proteins Nrd1, Nab3, and the RNA helicase Sen1 (Noel et al., 2012). The same process is achieved by spliceosomal cleavage in *S. pombe* (Box et al., 2008). So far, the significance of generating polyadenylated precursors is not well understood.

Budding yeast telomerase RNA acts as a flexible scaffold for protein subunits (Zappulla and Cech, 2004). In fact, it retains its functions even when its size is reduced from 1157 nucleotides to 500 nucleotides, as long as essential protein binding domains remains intact (Zappulla et al., 2005).

Another aspect of telomerase RNA biogenesis is its trafficking. The proper localization of TLC1 depends on importin Mtr10 (Ferrezuelo et al., 2002). Using fluorescent in situ hybridization and living cell image technologies, it has been shown that TLC1 shuttles between the nucleus and the cytoplasm and that deletion of its binding parteners such as TERT or YKU70 results in cytoplasmic accumulation (Gallardo et al., 2011; Gallardo et al., 2008). While the biogenesis pathway of telomerase RNA is relatively well studied in budding yeast (Fig. 1.5), it remains largely unknown in fission yeast.

Several telomerase subunits have been identified based on mutant screens (Lendvay et al., 1996; Lundblad and Szostak, 1989) or homology search (Lingner et al., 1997b; Nakamura et al., 1997). Those are called Est1 (For Ever Shorter telomeres), Est2, Est3, and Est4/Cdc13. Among them, Est2 is the catalytic protein with RNA-directed DNA polymerase activity. Est4/Cdc13 is a single-strand telomeric DNA-binding protein while Est1 and Est3 are associated with telomerase (Lendvay et al., 1996; Lingner et al., 1997a; Tuzon et al., 2011). Interactions between Est4/cdc13 and Est1 have been suggested in recruiting telomerase to the telomere ends (Lendvay et al., 1996; Li et al., 2009). The actual function of Est3 in telomerase is currently unknown, instead it has been implicated in telomere replication (Lee et al., 2010).

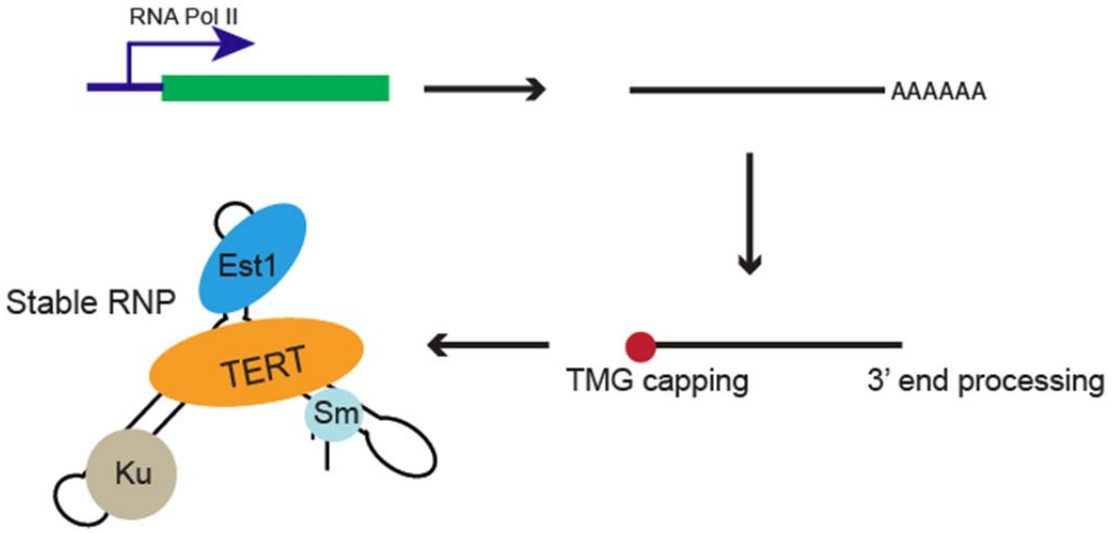

Figure 1.5: Telomerase biogenesis pathway in *S. cerevisiae*. The telomerase RNA precursor is synthesized by RNA polymerase II and has a poly (A) tail. The RNA loses the tail through 3' end processing and gains TMG cap at its 5' end. The active telomerase contains at Est1, TERT/Est2, Ku and Sm proteins.

I.2.c: Telomerase biogenesis pathway in human

The human TERT was first identified based on similarity to sequences of catalytic subunit from *Saccharomyces cerevisiae* and *Euplotes aediculatus* (Cech, 2004; Lingner et al., 1997b; Meyerson et al., 1997). The expression of TERT mRNA is mainly observed in primary tumors, cancer cell lines, and certain telomerase-positive tissues, but is undetectable in telomerase-negative cell lines and differentiated telomerase-negative tissues (Kolquist et al., 1998; Meyerson et al., 1997). The absence of TERT transcript is most frequently correlated with hypermethylation of its promoter (Liu et al., 2004b; Wang et al., 2009). However, reactivation of TERT can be directly induced by c-Myc transcription factor during tumorigenesis (Wu et al., 1999). Post transcriptional modifications of TERT, such as phosphorylation, have also been suggested to

affect catalytic activity (Haendeler et al., 2003). Recent studies stated that telomerase can promote the Wnt/beta-catenin signalling pathway as a transcriptional modulator independent of its telomerase activity (Park et al., 2009).

The RNA subunit of human telomerase was recovered using a PCR based approach (Feng et al., 1995). Its expression can be detected in most of tissues and cell lines, although germline tissues and tumor cell lines show the significant higher level than somatic tissues and cells. Human telomerase RNA is synthesized by RNA polymerase II, TMG capped on its 5' end post-transcriptionally, and processed at its 3' end to generated functional RNA (Feng et al., 1995; Jady et al., 2004) (Fig. 1.6). However, the existence of precursor and the mechanism of 3' end processing are currently unknown (Collins and Mitchell, 2002). A hairpin-Hinge-hairpin-ACA motif called H/ACA motif forms the 3' half of the telomerase RNA and is required for its accumulation (Mitchell et al., 1999; Mitchell and Collins, 2000). Like other snoRNAs with H/ACA motif, human telomerase RNA assembles with four H/ACA motif associated proteins dyskerin, NHP2, NOP10, and GAR1 (Dragon et al., 2000; Pogacic et al., 2000). A short sequence motif called CAB box has also been identified which is required for the telomere maintenance (Cristofari et al., 2007; Jady et al., 2004). Interestingly, the CAB box is not essential for telomerase RNA processing, accumulation, or telomerase activity *in vitro*. Instead it affects telomerase RNA location in Cajal bodies, which suggests subnuclear localization of this RNA is

an important regulatory mechanism for telomere length homeostasis. Indeed, a trans-acting factor called TCAB binds to telomerase RNA and promotes localization of telomerase RNA to Cajal bodies (Venteicher et al., 2009). Since the RNA subunit is ubiquitously expressed and TERT is not, the RNA might have other roles independent of its telomerase activity. For example, it may be involved in pseudouridylation of other RNAs as an H/ACA RNA or participate in nucleation of some nuclear bodies (Shevtsov and Dundr, 2011).

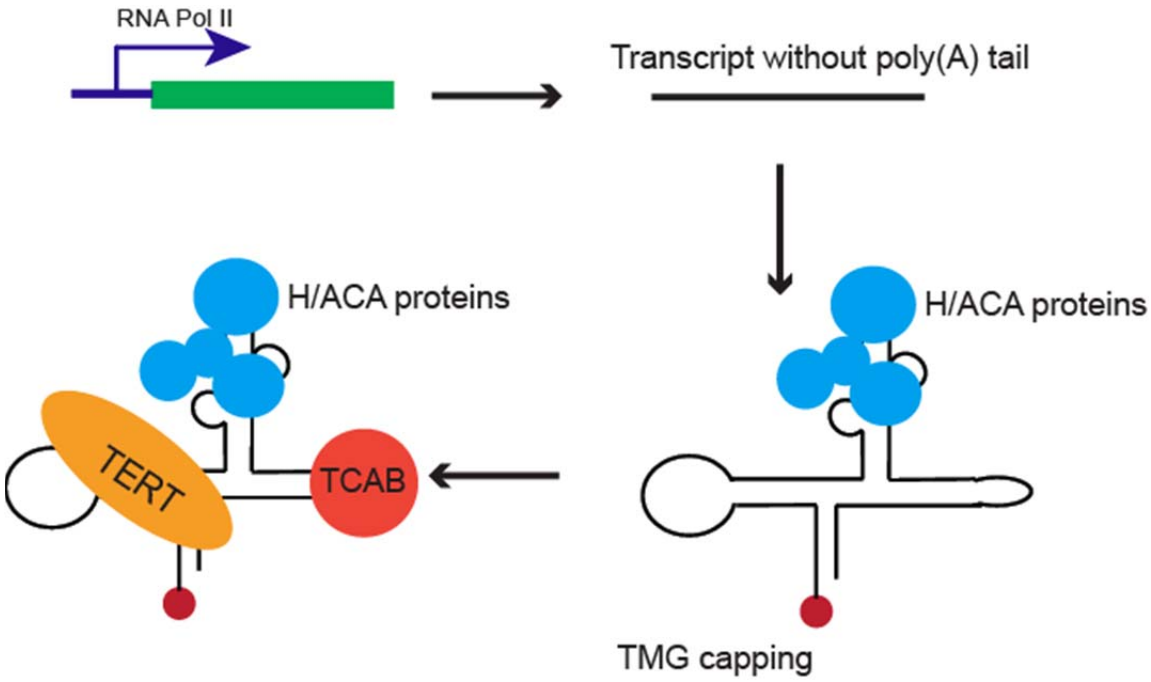

Figure 1.6: Telomerase biogenesis pathway in human. In human cells, telomerase RNA is transcribed by RNA polymerase II. The nascent transcript is modified with a 5′-TMG cap and associate with the H/ACA-motif proteins. The active complex includes TERT and TCAB.

I.3: Introduction to Sm and Lsm proteins

Pre-mRNA splicing is mediated by the spliceosome, a ribonucleoprotein complex composed of five snRNAs (U1, U2, U4, U5, and U6) and more than 150 proteins (Wahl et al., 2009). The protein subunits fall into two classes, the Sm proteins which are associated with U1, U2, U4 and U5 snRNAs, and specific proteins for each snRNP. Sm proteins were first identified from a patient named Stephanie Smith in which the systemic lupus erythematosus-associated anti-Sm autoimmune antibodies were found (Lerner and Steitz, 1979). Homology searches recovered several proteins termed Sm-like (Lsm) proteins. Both Sm and Lsm proteins have a conserved sequence motif that consists of seven amino acids embedded in a characteristic pattern of hydrophobic and hydrophilic residues (Achsel et al., 2001). In addition to the sequence similarity, they also share two other common features. The first is their ability to form heterogeneous heptamers. The second feature is their propensity to bind to RNA with specificity for oligo U stretches (Khusial et al., 2005). However, Sm and Lsm proteins have different modes of RNA binding, have non-overlapping sets of target RNAs, and perform distinct roles in RNA processing. Their different characteristics and functions will be discussed here.

I.3.a: Characteristics and functions of Sm proteins

Seven Sm proteins have been characterized. These are named SmB1, SmD1, SmD2, SmD3, SmE, SmF and SmG. They assemble in a stepwise manner onto

a single-stranded motif of the RNAs termed the Sm site (Kambach et al., 1999; Urlaub et al., 2001). The detailed hierarchical maturation pathway of the Sm complex has been recapitulated in vitro with recombinant Sm proteins (Raker et al., 1996). Sm proteins initially form three sub-complexes D1/D2, E/F/G and D3/B1, then binding of D1/D2 and E/F/G onto the RNA leads to the formation of a Sm subcore intermediate, which is finally bound by the D3/B heterodimer (Fig. 1.7). Even though mixing RNA and purified recombinant Sm proteins are sufficient for Sm core assembly in vitro, the essential Survival of Motor Neurons (SMN) complex which composed of the SMN protein and Gemins2–8 is required for this process in vivo in mammals (Chari et al., 2008; Meister and Fischer, 2002). Mutations in SMN cause spinal muscular atrophy featured with degeneration of motor neuron in the spinal cord (Lefebvre et al., 1995) presumably through affecting the snRNP assembly and thus RNA splicing (Zhang et al., 2008). Until now, only a few RNAs have been found to associate with Sm complex including U1, U2, U4, U5 snRNAs, (Lerner and Steitz, 1979), U11, U12, and U4atac minor snRNAs (Tarn and Steitz, 1996), trans-spliced leader RNA in C. elegans (Thomas et al., 1988), and telomerase RNA in budding yeast (Seto et al., 1999).

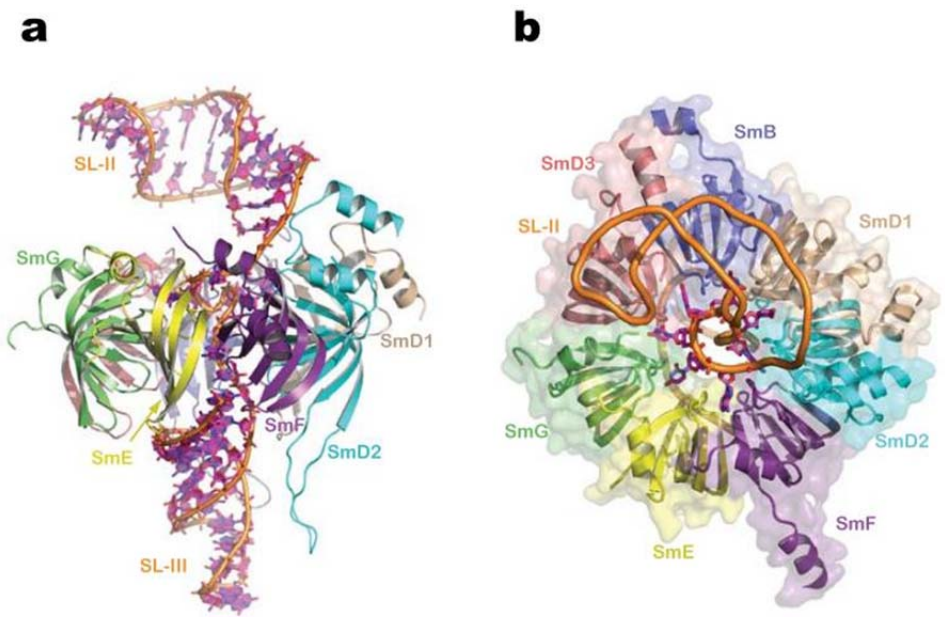

Figure 1.7: Overall structure of the Sm complex in U4 snRNP. a, Side view of the ring structure with its flat face up and tapered face down. b, View from the flat face of the ring. Adapted from (Leung et al., 2011).

Sm proteins are essential for the proper assembly of snRNP, and thus essential for cell viability. Two additional functions have been demonstrated. Firstly, they are involved in TMG cap formation in the 5' end of snRNAs. This function was first implicated when cap hypermethylation on U2 was studied in Xenopus extract (Mattaj, 1986), and it was later shown that TMG-capping of human U1 in vitro requires the presence of SmB/B'-SmD3 (Plessel et al., 1994; Raker et al., 1996). Physical association with Sm proteins led to the identification of the guanosine N2 methyltransferase Tgs1 (TrimethylGuanosine Synthase) (Mouaikel et al., 2002). However, whether this interaction is required for cap hypermethylation was not examined. The second function of Sm proteins is to facilitate pre-mRNA splicing. Three Sm proteins in U1 snRNP, SmB1, SmD1,

and SmD3, directly contact with pre-mRNA substrate through their long and positively charged C-terminal tails. This interaction may stabilize the association between U1 snRNP and pre-mRNA, and hence promote splicing (Zhang et al., 2001).

I.3.b: Characteristics and functions of Lsm proteins

As many as 16 different Lsm proteins have been identified in eukaryotes (Albrecht and Lengauer, 2004), and at least two different Lsm complexes have been well characterized (Mayes et al., 1999). One of them is called the Lsm1-7 complex which is composed of Lsm1, Lsm 2, Lsm 3, Lsm 4, Lsm 5, Lsm 6, and Lsm7. The Lsm1-7 complex is localized to the cytoplasm and functions in the decapping step of mRNA decay. Mutation in any of seven Lsm proteins causes inhibition of mRNA decapping and decay in budding yeast (Tharun et al., 2000). Co-immunoprecipitation shows the Lsm1-7 complex physically interacts with mRNA decapping enzyme (Dcp1) and a decapping activator called Pat1 (Bouveret et al., 2000; Tharun et al., 2000) (Fig. 1.8). However, whether the Lsm1-7 complex have the same function in other species or how many mRNAs are degradated through this pathway should be further examined.

Another interesting finding on the Lsm1-7 complex is that it interacts with poly(A) tracts at the 5' end of orthopoxvirus mRNAs and represses RNA decay by inhibiting exonucleases as well as decapping of RNA substrates (Bergman et al., 2007). Different from the situation discussed above, the Lsm1-7 complex

here simultaneously binds to the 5' and 3' end of orthopoxvirus mRNAs to form a loop structure which protects the 5' cap from decapping activity and the 3' end from exonucleolytic processing (Bergman et al., 2007).

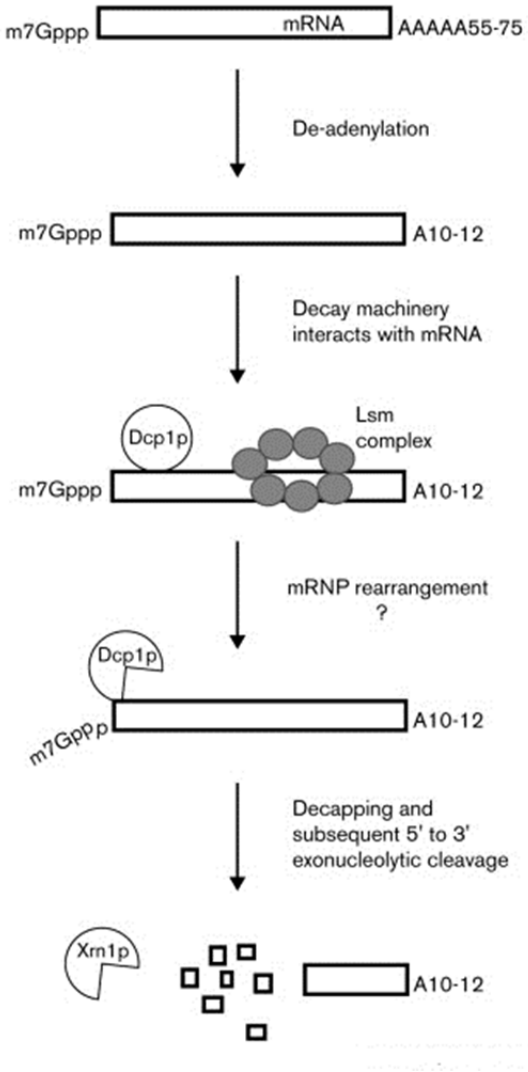

Figure 1.8: The Lsm1-7 complex functions in mRNA decapping. The figure illustrates one RNA degradation pathway involving Lsm proteins which recruits decaping enzyme Dcp1p onto mRNA. It should be noted that it is unclear when the Lsm complex and Dcp1p interact with the mRNA during the decay pathway. Adapted from (He and Parker, 2000)

The other well characterized complex is called the Lsm2-8 complex which consists of Lsm2, Lsm 3, Lsm 4, Lsm 5, Lsm 6, Lsm7, and Lsm8. Similar to the Sm complex, purified human Lsm2 to Lsm8 form heptameric ring structure (Achsel et al., 1999). Lsm2-8 is mainly localized in the nucleus where it is involved in the maturation of various RNAs (Wilusz and Wilusz, 2005). It binds to U6 snRNA through the 3' terminal oligo (U) stretch (Achsel et al., 1999). Since U6 snRNP is one of the main components of the spliceosome, it is not surprising that depletion of Lsm2 to Lsm8 proteins leads to splicing defect (Mayes et al., 1999). Further analysis demonstrates that Lsm2-8 is required for U6 snRNA stability *in vivo* and it has a chaperone-like function in remodeling RNP complexes in spliceosome (Verdone et al., 2004)(Fig. 1.9). Besides its essential role in U6 snRNA, the Lsm2-8 complex is required for the processing and stability of ribosome RNAs and tRNAs(Kufel et al., 2003; Kufel et al., 2002).

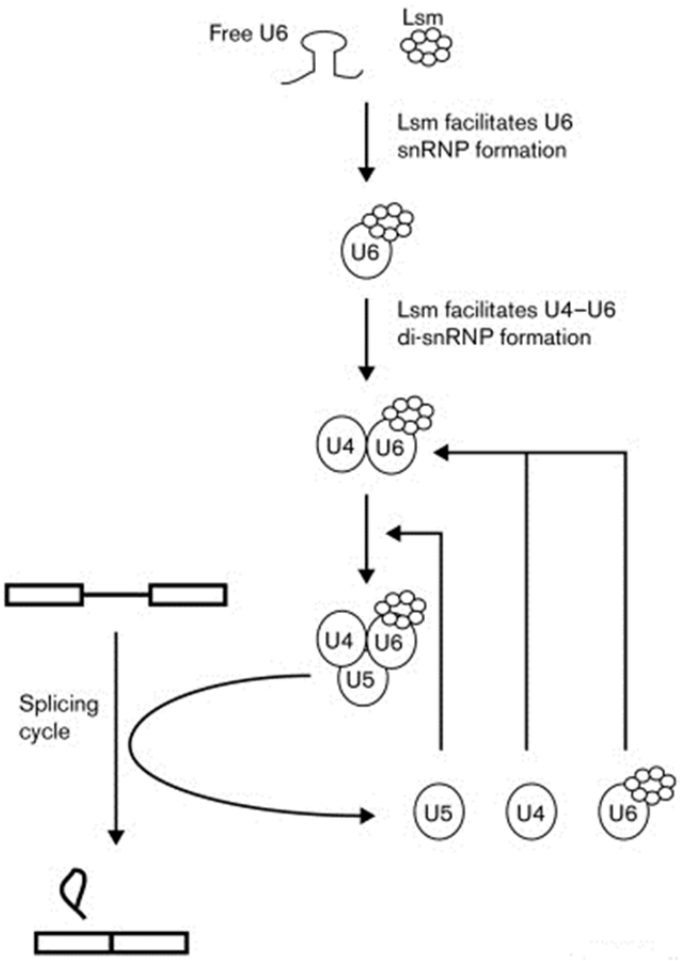

Figure 1.9: The Lsm2-8 complex involves U6 snRNP biogeneis. Lsm proteins is required for the stability of U6 snRNA and U4/U6 di-snRNP formation possibly by rearranging RNA-protein structures. Adapted from (He and Parker, 2000)

I.4: Scope of Dissertation

The primary goal of the research conducted during my tenure as a graduate student is to elucidate the biogenesis pathway of telomerase RNA in fission yeast from the transcription of the gene encoding the telomerase RNA subunit to the incorporation of a processed and functional form of this non-coding RNA into a complex with the catalytic subunit of telomerase. In 2008, our lab first reported the identification of telomerase RNA subunit, called TER1, in fission yeast (Leonardi et al., 2008). TER1 is first generated as a polyadenylated precursor with an intron at its 3' end by RNA polymerase II. Characterization of the maturation pathway uncovered an unexpected role for the spliceosome, which normally catalyses splicing of pre-messenger RNA. The first spliceosomal cleavage reaction generates the mature 3' end, releasing the active form of the RNA without exon ligation (Box et al., 2008). The functional mature TER1 ends at the putative Sm binding site which partially overlaps with the 5' splicing site.

My first project in the lab was to investigate whether the Sm site can affect spliceosomal cleavage. We found that Sm site mutations result in a nine fold reduction in the steady state level of mature TER1 RNA and in an increase of the TER1 precursor. Analysis of the mature 3' end of TER1 revealed that further exonucleolytic processing occurs after cleavage. In examining the proteins that associate with the mature 3' end of TER1 we found that the canonical Sm complex and the Lsm2-8 complex sequentially bind to TER1

during TER1 maturation. The Sm complex specifically binds to the precursor and mature TER1 following cleavage but prior to further exonucleolytic processing. At later stages the Lsm complex replaces the Sm complex and associated with majority of mature TER1. Immunoprecipitation of telomerase revealed that most of the telomerase activity is associated with the Lsm-containing complex. In addition, our data indicates the Sm complex and Lsm complex play important but different roles in TER1 maturation. Binding of the Sm complex might be required for trimethylguanosine (TMG) cap formation in TER, while binding of the Lsm complex protects the mature form of TER1 from exonucleolytic degradation. The results of this research are reported in Chapter Three of this thesis.

The comprehensive list of the materials and methods used to execute the studies is described in Chapter Two. Chapter Four provides a summary of the results presented in Chapter Three. In addition, this chapter describes numerous future directions for research that build upon the studies reported in Chapter Three.

Chapter II. Materials and Methods

II.1: Yeast strains and constructs

Strain Number	Genotype	Source
PP138	*h⁻ ade6-M216 leu1-32 ura4-D18 his3-D1*	Lab stock
PP298	*h⁻ ade6-M210 leu1-32 ura4-D18 his3-D1 trt1-myc$_9$*	(Haering et al., 2000)
PP399	*leu1-32 ura4-D18 his3-D1 trt1-Cmyc$_9$ ter1::kanMX6*	Lab stock
PP407	*h$^+$/h⁻ leu1-32/leu1-32 ura4-D18/ura4-D18 his3-D1/his3-D1 ade6-M210/ade6-M216 ter1$^+$/ter1::kanMX6*	(Box et al., 2008)
PP433	*h$^+$/h⁻ leu1-32/leu1-32 ura4-D18/ura4-D18 his3-D1/his3-D1 ade6-M210/ade6-M216 ter1$^+$/ ter1::ura4*	This study
PP574	*h⁻ ade6-M216 leu1-32 ura4-D18 his3-D1 lsm2-myc$_{13}$-nat*	This study
PP575	*h⁻ ade6-M216 leu1-32 ura4-D18 his3-D1 lsm1-myc$_{13}$-nat*	This study
PP576	*h⁻ ade6-M216 leu1-32 ura4-D18 his3-D1 lsm3-myc$_{13}$-nat*	This study
PP577	*h⁻ ade6-M216 leu1-32 ura4-D18 his3-D1 lsm4-myc$_{13}$-nat*	This study
PP578	*h⁻ ade6-M216 leu1-32 ura4-D18 his3-D1 lsm5-myc$_{13}$-nat*	This study
PP579	*h⁻ ade6-M216 leu1-32 ura4-D18 his3-D1 lsm6-myc$_{13}$-nat*	This study
PP580	*h⁻ ade6-M216 leu1-32 ura4-D18 his3-D1 smb1-myc$_{13}$-nat*	This study
PP582	*h⁻ ade6-M216 leu1-32 ura4-D18 his3-D1 sme1-myc$_{13}$-nat*	This study
PP583	*h⁻ ade6-M216 leu1-32 ura4-D18 his3-D1 lsm7-myc$_{13}$-nat*	This study
PP584	*h⁻ ade6-M216 leu1-32 ura4-D18 his3-D1 smd2-myc$_{13}$-nat*	This study
PP585	*h⁻ ade6-M216 leu1-32 ura4-D18 his3-D1 lsm8-myc$_{13}$-nat*	This study
PP588	*h⁻ ade6-M216 leu1-32 ura4-D18 his3-D1 smf1-*	This study

	myc$_{13}$-nat	
PP670	*h$^+$ ade6-M210 ura4-D18 his3-D1 tgs1::kanMX6*	This study, based on (Hausmann et al., 2007)
PP677	*h$^+$ ade6-M216 ura4-D18 leu1-32 lsm1::kanMX6*	This study, based on diploid strain from Bioneer
PP678	*h$^+$ ade6-M216 ura4-D18 leu1-32 lsm3::kanMX6*	This study, based on diploid strain from Bioneer
PP694	*ade6-M216 leu1-32 ura4-D18 his3-D1 ter1::kanMX6 lsm4-myc$_{13}$-nat*	This study
PP695	*ade6-M216 leu1-32 ura4-D18 his3-D1 ter1::kanMX6 smb1-myc$_{13}$-nat*	This study
PP758	*leu1-32 ura4-D18 his3-D1 ter1::kanMX6 lsm4-myc$_{13}$-nat*	This study
PP759	*leu1-32 ura4-D18 his3-D1 ter1::kanMX6 smb1-myc$_{13}$-nat*	This study
S. cryophilus		Lab stock
S. octosporus		Lab stock

II.2: Yeast two-hybrid

Yeast two-hybrid was conducted using the Matchmaker GAL4 Two Hybrid System 3 (Clontech). Briefly, *tgs1* cDNA was cloned into the vector pGBKT7, and each full length *lsm* and *sm* cDNA was cloned into pGADT7. Plasmids were cotransformed into the yeast strain AH109 and positive transformants were selected on SD-Leu-Trp plates. Interactions were analysed by plating 3-fold

serial dilutions of overnight cultures onto SD-Leu-Trp-His-Ade plates. Plates were incubated for three days at 30°C.

II.3: Genomic DNA extraction and telomere length analysis

The volume of 10 ml cells was harvested by centrifugation at a density of about $5x10^7$ cells ml^{-1}. Pellets in 10 ml Z buffer (50 mM sodium citrate, 50 mM sodium monohydrogen phosphate, 40 mM EDTA pH 7.8), and resuspended in 2 ml Z buffer containing 0.5 mg/ml Zymolyase (ICN) and 2 mM dithiothreitol. The cell suspension was incubated at 37°C for 1 h. Sodium dodecyl sulfate (SDS) was added to 4% (wt/vol), and incubation was continued at 65°C for 10 min. The volume was increased to 10 ml with 5 × TE (50 mM Tris-HCl, pH 8.0, 5 mM EDTA) and proteinase K (Sigma) was added to 50 g/ml. After 1 h of incubation at 50°C, 3 ml of potassium acetate solution (5 M) was added and samples were incubated on ice for 30 min. The precipitate was removed by centrifugation and the clarified supernatant was mixed with 1 volume of isopropanol to precipitate nucleic acids. After 20 min on ice, samples were subjected to centrifugation at 10,000 × g for 5 min. Nucleic acids were dried briefly and resuspended in 0.5 ml of 5× TE containing DNase-free RNase A. Following incubation at 37°C for 1 h, organic extraction, and ethanol precipitation, genomic DNA was resolubilized in 1× TE.

II.4: Telomerase activity assay

S. pombe were grown in YES (yeast extract supplements) and 6 litres of cell suspension were harvested by centrifugation at a density of $5x10^6$ cells ml^{-1}. Cells were washed in TMG(300) (10 mM Tris-HCl buffer, pH 8.0, 1 mM magnesium chloride, 10% (v/v) glycerol, 300 mM sodium acetate), the pellet was resuspended in two packed cell volumes TMG(300) plus supplements (5 µg ml^{-1} chymostatin, 5 µg ml^{-1} leupeptin, 1 µg ml^{-1} pepstatin, 1 mM benzamidine, 1 mM DTT, 1 mM EDTA and 0.5 mM PMSF) and the suspension was frozen in liquid nitrogen. Cells were lysed under liquid nitrogen using SPEX SamplePrep freezer mill. The lysed cell powder was transferred into a 50 ml tube and allowed to thaw on ice for 30 min. Cell extracts were cleared by two rounds of centrifugation at 14,000g for 7 min and frozen in liquid nitrogen for storage at -80 °C. The concentration of proteins in the whole cell extract was measured by Bradford protein assay. Telomerase was enriched on agarose beads coated with anti–c-Myc. Telomerase activity assays contained 20 µl of beads in 50 mM Tris-acetate, pH 8.0, 100 mM potassium acetate, 1 mM magnesium acetate, 5% (v/v) glycerol, 1 mM spermidine, 1 mM DTT, 0.2 mM dATP, dCTP and dTTP, 2 mM ^{32}P-α-dGTP and 5 mM PBoli14 (TGTGGTGTGTGGGTGTG). Reactions were incubated at 30 ºC for 90 min and then treated with proteinase K (2.0 mg ml^{-1} in 0.5% (w/v) SDS, 40 mM EDTA, 20 mM Tris-HCl, pH 7.5). ^{32}P-5'-end labeled 100-mer oligonucleotide was added as a precipitation control before phenol/chloroform extraction. The products were separated on a 10%

sequencing gel containing 8 M urea and were visualized by using a phosphorimager.

II.5: Total RNA isolation

Cultures (500 ml) were grown to a density of 5×10^6 cells ml^{-1} at 32 ºC and cells were harvested by centrifugation. After one wash with ddH_2O, cells were resuspended in ddH_2O and quick-frozen by dripping the cell suspension into liquid nitrogen. Cells were lysed in a 6850 Freezer mill (SPEX SamplePrep) using 8 cycles (2 min) at a rate of 10 per second with 2 min cooling time between cycles. The lysed cell powder was transferred directly into tubes containing 10 ml phenol/chloroform/isoamyl alcohol (25:24:1) and 10 ml sodium acetate (50mM), 1% (w/v) sodium dodecyl sulphate preheated to 65 ºC. RNA was extracted five times with phenol/chloroform/isoamyl alcohol and once with chloroform/ isoamyl alcohol. Total RNA was ethanol precipitated and resuspended in 50mM sodium acetate (pH 5.2).

II.6 Northern blotting analysis

Total and immunoprecipitated RNAs were separated on 4% polyacrylamide gels in TBE/7M urea and transferred onto Biodyne Nylon Transfer Membrane (Pall Corporation). RNA was ultraviolet cross-linked (254 nm, 120 mJ) in a Stratalinker (Stratagene). Hybridizations with radiolabelled probes were carried out in Church-Gilbert buffer at 65 ºC (*ter1* probe) or 42 ºC (DNA

oligonucleotides probe). The TER1 probe was generated by nick-translation of a PCR fragment (nucleotides 17 to 249) in the presence of ^{32}P-α-dCTP. DNA oligonucleotide probes were labelled with polynucleotide kinase (Nika et al.) in the presence of ^{32}P-α-ATP. The following oligonucleotides were used in this study: GCTGCAGAAACTCATGCCAGGTAAGT (snRNA U1), CGCTATTGTATGGGGCCTTTAGATTCTTA (snoRNA snR101), CTTCATCGATGCGAGAGCCAAGAGATCCGT(5.8S rRNA), and GCAGTGTCATCCTTGTGCAGGGGCCA (snRNA U6).

II.7: RNA immunoprecipitation

The cell lysate was produced in the same way for telomerase activity assay. For c-Myc immunoprecipitation, monoclonal anti-c-Myc antibody (20 µg, Sigma) was incubated with 150 µl protein A/G agarose slurry (Calbiochem) in phosphate buffered saline at room temperature for 30 min. Beads were washed three times with TMG(300) plus supplements and whole cell extract (1.2 ml) was added at a concentration of 5 mg ml^{-1} together with RNasin (1 µl, Promega), Tween 20 (0.1% final conc) and heparin (1mg ml^{-1}). For immunoprecipitation of TMG-capped RNA anti-TMG antibody (3 µg, Calbiochem) was bound to 50 µl proteinA/G agarose slurry (Calbiochem), washed with TMG (300) and 150 µg total *S. pombe* RNA was added in 0.7 ml TMG(300). Samples were incubated on a rotator at 4 ºC for 4 hours and then washed three times with TMG(300) plus supplements and 0.1% Tween 20 and once with TMG(50) (as TMG(300)

but only 50 mM sodium acetate). Protease inhibitors were omitted for TMG IPs. RNA was isolated by treatment with proteinase K (2.0 mg ml^{-1} in 0.5% (w/v) SDS, 40 mM EDTA, 20 mM Tris-HCl, pH 7.5) at 50 °C for 15 min, followed by extraction with acidic phenol and ethanol precipitation. RNA was then analysed by northern blotting, RT-PCR and 3' end sequencing.

II.8: Reverse Transcription-Polymerase Chain Reaction (RT-PCR)

Semi-quantitative RT-PCR was performed as described previously(Box et al., 2008). The reverse primer BLoli1275 (CGGAAACGGAATTCAGCATGT) and forward primer BLoli1020 (CAAACAATAATGAACGTCCTG) were used to amplify the intron-spanning fragment of TER1. The reverse primer PBoli918 (ACAACGGACGAGCTACACTC) and forward primer BLoli1006 (CATTTAAGTGCTTGTCAGATCACAACG) were used to amplify region in the first exon. The reverse primer BLoli2051 (GACCTTAGCCAGTCCACAGTTA) and forward primer Bloli2101 (ACCTGGCATGAGTTTCTGC) were used to amplify snRNA U1.

II.9: Determination of 3' end sequences using conventional sequencing

DNase-treated total RNA samples (2.5 mg) were incubated with poly(A) polymerase (US Biologicals), RNase inhibitor (RNasin, 40 U), and ATP (0.5mM) in the buffer provided by the manufacturer in 20 µl reactions at 30 °C for 30 min. The reaction volume was increased to 35.5 µl by the addition of the DNA

oligonucleotide PBoli560 (GCGGAATTCT18, 125 pmol) and dNTP mix (25 nmol), and the reactions were incubated at 65 °C for 3 min followed by slow cooling to room temperature (20 °C). The reaction volume was then adjusted to 50 ml with first strand buffer (Invitrogen), dithiothreitol (5 mM), RNasin (40 U) and Superscript III reverse transcriptase (200 U, Invitrogen), and reactions were incubated at 50 uC for 60 min. RNaseH (5 U, NEB) was added and incubation was continued at 37 °C for 20 min. Aliquots (2.5 ml) of this reaction were used for PCR amplification with Taq polymerase (5 U, NEB), and oligonucleotide specific for *S. cryophilus* and *S. octosporu* telomerase RNA and PBoli560 (200 nM each) under the following conditions: 3 min at 94 °C followed by 30 cycles of 30 s at 94 °C, 45 s at 55 °C and 120 s at 72 °C, followed by 7 min at 72 °C. PCR products were separated by electrophoresis on 0.8% agarose gels, and bands of the correct size were excised, purified and cloned into the TOPO TA cloning system (Invitrogen) for sequence analysis.

II.10: Determination of 3' end sequences using illumina sequencing

DNase-treated total RNA samples (2.5 µg) or immunoprecipitated and purified RNA was incubated with poly(A) polymerase (1 µl, US Biologicals), RNase inhibitor (RNasin, 40 U) and ATP (0.5 mM) in 20 µl reactions at 30 °C for 30 min. The reaction volume was increased to 35.5 µl by the addition of the oligonucleotide Bloli2327 (CAAGCAGAAGACGGCATACGA$(T)_{18}$, 125 pmol) and dNTP mix (25 nmol), and reactions were incubated at 65 °C for 3 min

followed by slow cooling to room temperature. The reaction volume was then adjusted to 50 µl with first strand buffer (Invitrogen), dithiothreitol (5 mM), RNasin (40 U) and Superscript III reverse transcriptase (200 U, Invitrogen), and reactions were incubated at 50 °C for 60 min. RNaseH (5 U, NEB) was added and incubation was continued at 37 °C for 20 min. Aliquots (3 µl) of this reaction were used in PCR with Taq polymerase (5 U, NEB) and primers (GTTCAGAGTTCTACAGTCCGACGATC##GCAAAATGTTAAAAGGAACG) and Bloli2330 (CAAGCAGAAGACGGCATACGA) (200 nM each, ## represents a two-nucleotide barcode used for multiplexing) under the following conditions: 3 min at 94 °C followed by 10 cycles of 30 sec at 94 °C, 45 sec at 55 °C and 60 sec at 72 °C, followed by 7 min at 72 °C. PCR products were purified using the QIAquick PCR Purification Kit (Qiagen) and eluted with 46 µl elution buffer. In the second round of PCR, 23 µl of the eluted product was amplified with Bloli2329 (AATGATACGGCGACCACCGACAGGTTCAGAGTTCTACAGTCCGA) and Bloli2330 (200 nM each) under the following conditions: 3 min at 94 °C followed by 29 cycles of 30 s at 94 °C, 45 s at 55 °C and 60 s at 72 °C, followed by 7 min at 72 °C. PCR products were separated by electrophoresis on 1.5 % agarose gels, and bands of the correct size were excised and purified. The concentration of the PCR products was measured using the Agilent 2100 Bioanalyzer (Agilent Technologies) and further adjusted to 10 nM for massively parallel sequencing using Illumina sequencing technology. Reads were

analysed using a custom script written in BioPerl to sort the reads into different bins based on the two nucleotide barcode and filter for those that contained the TER1 sequence (GCAAAAN$_{10}$AACG). The nucleotide sequence between GCAAAAN$_{10}$AACG and the oligo(A) sequence resulting from the poly(A) polymerase treatment represents the end of TER1 and was used to count reads to determine the 3' end sequence distribution at single nucleotide resolution. Further analysis and graphs were prepared in Excel.

II.11: Quantitative real-time RT-PCR and data analysis

Reverse transcription for input and immunoprecipitated RNA was performed with antisense primer Bloli2860 (TGCTCAGACCAAGTGAAAAA) or Bloli2051. Real-time PCR was performed in triplicate 12.5 µl reactions using Power SYBR Green PCR Master Mix (applied biosystems) according to the manufacturer's instructions. Bloli2860 and Bloli2859 (GGATCAAAGCTTTTGCTTGT) were used to amplify the first exon of TER1. Bloli2051 and Bloli2101 were used to amplify snRNA U1. The qRT-PCR results were imported into Excel and the average value and standard deviation of triplicate Ct values were calculated. Enrichment of immunoprecipitation is represented by ΔCt which is the Ct value from immunoprecipitation minus the Ct value from input. Error bars in the graph represent the positive and negative range of the standard error of the mean.

Chapter III: TER1 is sequentially bound by Sm and Lsm complexes

III.1: Abstract

In eukaryotes, the telomeric DNA is synthesized by telomerase through reverse transcription using part of an RNA subunit as a template. So far, the biogenesis of telomerase RNA is not well understood. Mapping the ends at single nucleotide resolution revealed exonucleolytic processing at the Sm site. In examining the proteins that associate with the mature 3' end of TER1 we find that the Sm complex and the Lsm2-8 complex sequentially bind to TER1. Specifically, the Sm complex binds to the TER1 precursor and the cleaved form prior to further exonucleolytic processing. At later stages, the Lsm2-8 complex replaces the Sm complex and binds to majority of mature TER1. Immunoprecipitation of telomerase from yeast extracts shows that most of the telomerase activity is associated with the Lsm-containing complex. Moreover, the Sm complex and Lsm complex have distinct functions in TER1 maturation. Binding of the Sm complex facilitates splicing and is required for trimethylguanosine (TMG) cap formation in TER, while binding of the Lsm2-8 complex protects the majority of mature TER1 from further exonucleolytic degradation and facilitates telomerase assembly. Altogether, our findings constitute the first identification of RNA whose biogenesis requires Sm and Lsm2-8 complexes.

III.2: Introduction

In eukaryotes telomeres protect the ends of chromosomes from degradation and fusion. Defects in telomere maintenance have been shown to contribute to several age-related diseases, as well as cancer (Blasco, 2005; Collins and Mitchell, 2002; Kim et al., 1994). Telomeric DNA is synthesized by the telomerase ribonucleoprotein complex which contains several protein subunits and an essential non-coding RNA. The RNA subunit functions as the scaffold for the assembly of protein subunits (Zappulla and Cech, 2004) and provides the template for reverse transcription.

The telomerase RNA subunit is surprisingly divergent among different species. First, the transcription machinery is different. Ciliate telomerase RNA is synthesized by RNA polymerase III, while RNA polymerase II transcribes the yeast and vertebrate telomerase RNA. Second, the size ranges from about 150 nt in ciliates, to 400–600 nt in vertebrates, and to over 1000nt in fungi (Chen et al., 2000; Lin et al., 2004). Third, comparison of telomerase RNAs from mammals, yeasts, and ciliates reveals no obvious sequence similarity. For those reasons, telomerase RNA is not discovered in lots of species and its biogenesis pathway remains largely unknown, although the first telomerase RNA was identified for over three decades ago (Greider and Blackburn, 1987, 1989).

Our lab uses *Schizosaccharomyces pombe* (*S. pombe*) as a model organism to study telomerase RNA biogenesis. We have identified the RNA subunit of

telomerase (TER1) (Leonardi et al., 2008) and we have shown that the mature 3' end of TER1 is generated by the spliceosome in a reaction akin to the first step of splicing (Box et al., 2008). The cleavage occurs at a 5' splice site that partially overlaps with a putative Sm site ($AU_{4-6}GR$) (Fig. 3.1). The presence of the Sm site is critical to maintain normal TER1 RNA levels in the cell (Box et al., 2008).

The Sm sites have been shown to be bound by Sm protein family members in other RNAs (Achsel et al., 2001). Seven Sm proteins (E, F, G, D1, D2, D3 and B) form a heteroheptameric ring at the Sm sites of U1, U2, U4 and U5 small nuclear RNAs (snRNAs) (Raker et al., 1996) which are involved in pre-mRNA splicing. The essential SMN (Survival of Motor Neurons protein) complex is required for Sm complex assembly in vivo (Chari et al., 2008). As many as 16 different Sm-like (Lsm) proteins have been identified in eukaryotes (Albrecht and Lengauer, 2004) and two different Lsm complexes have been characterized. The Lsm2-8 complex is localized in the nucleus and is known to bind to polymerase-III-transcribed U6 snRNA(Achsel et al., 1999), while the Lsm1-7 complex transiently associates with mRNAs and promotes mRNA decay by recruiting the decapping enzyme (Dcp1) in the cytoplasm(Bouveret et al., 2000; Tharun et al., 2000).

My results show that further exonucleolytic processing occurs at the Sm site after spliceosomal cleavage of the TER1 RNA. More interestingly, my results suggest that TER1 is sequentially bound by the Sm and Lsm2-8 complexes.

The Sm complex specifically binds to the TER1 precursor and the cleaved form prior to further exonucleolytic processing. At later stages, the Lsm2-8 complex replaces the Sm complex and is present on the majority of mature TER1. The Sm and Lsm complex play distinct roles in TER1 biogenesis. The Sm complex stimulates spliceosomal cleavage and subsequently promotes the hypermethylation of the 5' cap. The Lsm complex protects the mature 3' end of TER1 from exonucleolytic degradation and remains associated with the majority of active telomerase in the cell. These observations make TER1 the first RNA whose biogenesis involves the binding of Sm and Lsm complexes.

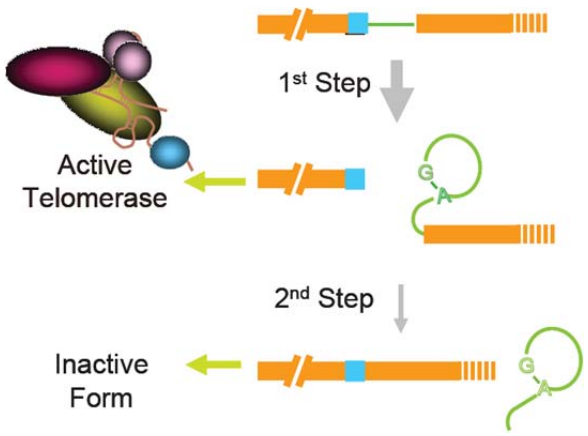

Figure 3.1: The mature 3' end of TER1 is generated by the first step of splicing. The second reaction rarely happens and when it does an inactive form is generated. The following features of TER1 are depicted: Sm site, light blue box; intron, green bar; poly (A) tail, dashed line.

III.3: Results

III.3a Exonucleolytic processing occurs at the Sm site after spliceosomal cleavage of the TER1 RNA

Previous work has established that mature TER1 is generated through the first transesterification reaction of splicing occurring right after a Sm site (AUUUUUUG) (Box et al., 2008). This conserved Sm site has been identified at the 3' end of telomerase RNA among yeast species (Dandjinou et al., 2004; Gunisova et al., 2009; Leonardi et al., 2008) and has been suggested to be bound by the heteroheptameric Sm complex in *Saccharomyces cerevisiae* (Seto et al., 1999). To accurately measure where mature TER1 terminates, a strategy was developed using deep sequencing to analyze its 3' ends at single nucleotide resolution and to identify the most abundant terminal sequences (Fig. 3.2a). The analysis revealed that, after spliceosomal cleavage, over 60% of TER1 molecules lost additional nucleotides at the 3' end and terminated in a stretch of three to six uridines (Fig. 3.2b). The discovery of this shortened Sm site and the notion that a conserved stem-loop downstream of the Sm site is required for stable binding of the Sm complex raised the question whether Sm can still bind to mature TER1. The 3' end of mature TER1 with the run of uridines resembles the end of Lsm-class snRNA (U6 and $U6_{atac}$) which are known to be bound by the Lsm2-8 complex (Shibata et al., 1974).

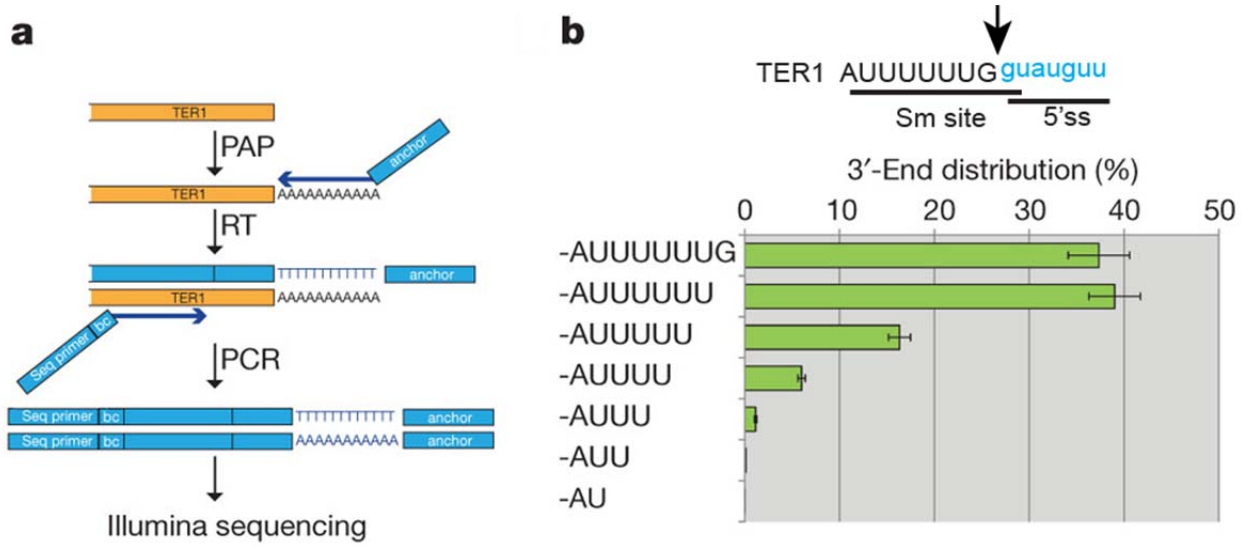

Figure 3.2: Exonucleolytic processing occurs at the Sm site. a, Method used to map the 39-end distribution of TER1 post spliceosomal cleavage. RNA is depicted in orange, DNA in blue. PAP, poly(A) polymerase; RT, reverse transcription; PCR, polymerase chain reaction; bc, barcode. b, The upper panel is the sequences of the Sm site and 5' splicing site (5'ss) in TER1. The arrow points the spliceosomal cleavage site. The low panel is the distribution of 3' end positions in mature TER1 from wild-type cells. The average of four experiments is shown; error bars, standard deviation.

III.3b. TER1 RNA associates with Sm and Lsm proteins

To test whether Sm or Lsm proteins associate with TER1, carboxy-terminal c-Myc epitope tags were inserted at the genomic loci of all Sm and Lsm proteins. Immunoprecipitations were performed with a subset of strains that did not show overt growth defects, expressed tagged proteins and maintained telomeres (data not shown). The snRNA U1 control specifically co-precipitated with Sm proteins, confirming that the epitope tags did not interfere with immunoprecipitation of RNP complexes (Fig. 3.3a and Fig. 3.3b). TER1 co-immunoprecipitated with Smd2, Sme1 and Smb1 which represent members of each of three sub-complexes D1/D2, E/F/G and D3/B (Raker et al., 1996) (Fig.

3.3a, lanes 2–4). Interestingly, several-fold more TER1 was recovered from Lsm immunoprecipitates resulting in an approximately 80% depletion of TER1 from the immunoprecipitation supernatant (Fig. 3.3a, lanes 5–7). TER1 precipitated with all subunits of the Lsm2–8 complex (Fig. 3.3a), but not with Lsm1 (Fig. 3.3a, lane 8), the subunit specific to the Lsm1–7 complex. In addition, the catalytic subunit of telomerase Trt1 co-immunoprecipitated similar amount of TER1 as Lsm proteins suggesting that telomerase is an Lsm-containing complex (Fig. 3.3b).

TER1 precursor presented by the slower migrating band on TER1 northern blot is only co-immunoprecipitated with Sm proteins (Fig. 3.3a and Fig. 3.3b). To gain further insights into the functions of Sm and Lsm binding to telomerase we initially focused on the Sm association. For most characterized snRNAs, sequences downstream of the Sm-binding site are critical for Sm loading (Yong et al., 2010). As the mature form of TER1 lacks such sequences, we tested whether the Sm complex was loaded onto the TER1 precursor before spliceosomal cleavage. Reverse transcription PCR (RT–PCR) confirmed that the precursor is indeed specifically associated with the Sm complex, but is undetectable in Lsm immunoprecipitations (Fig. 3.3c).

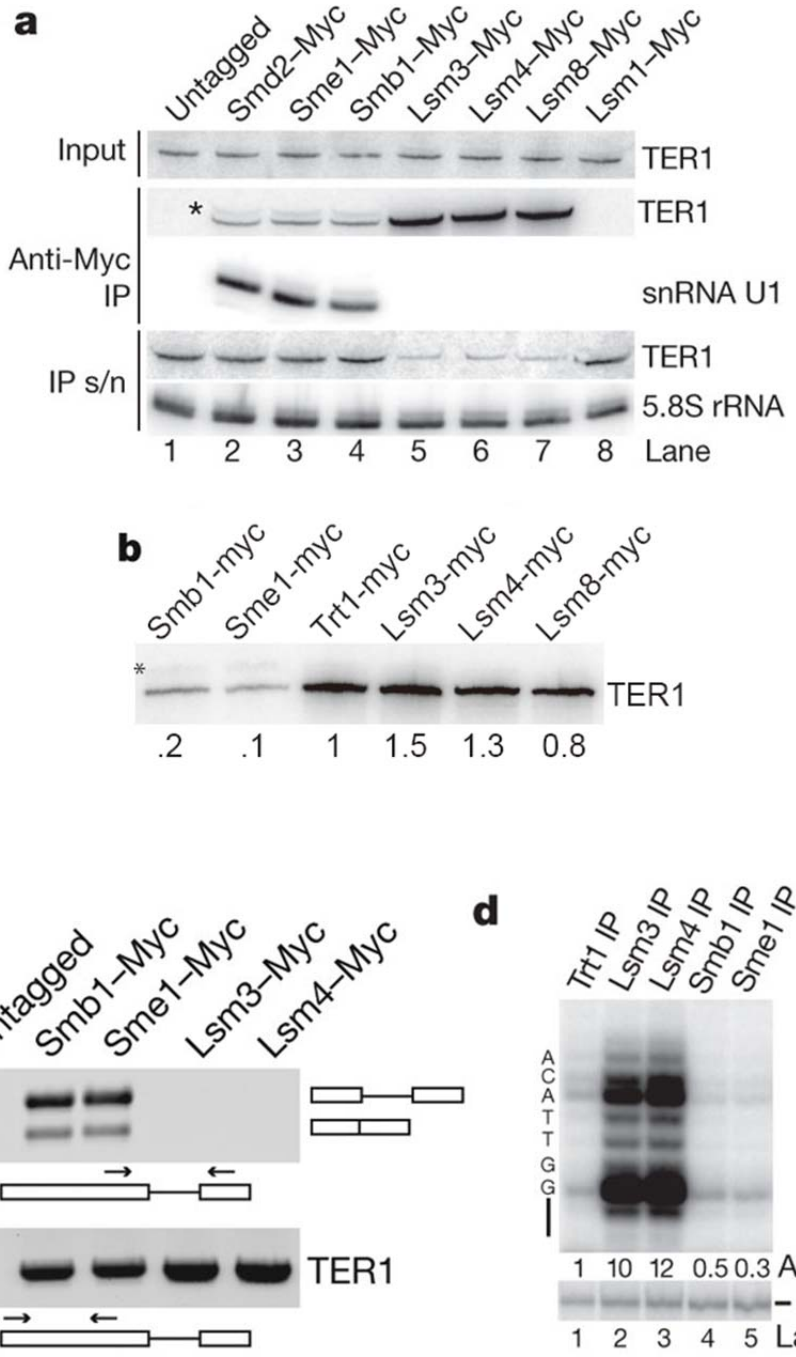
a
Untagged
Smd2–Myc
Sme1–Myc
Smb1–Myc
Lsm3–Myc
Lsm4–Myc
Lsm8–Myc
Lsm1–Myc
Input
TER1
*
TER1
Anti-Myc IP
snRNA U1
IP s/n
TER1
5.8S rRNA
1 2 3 4 5 6 7 8 Lane
b
Smb1-myc
Sme1-myc
Trt1-myc
Lsm3-myc
Lsm4-myc
Lsm8-myc
*
TER1
.2 .1 1 1.5 1.3 0.8
c
Untagged
Smb1–Myc
Sme1–Myc
Lsm3–Myc
Lsm4–Myc
TER1
d
Trt1 IP
Lsm3 IP
Lsm4 IP
Smb1 IP
Sme1 IP
ACATTGG
1 10 12 0.5 0.3 Activity
LC
1 2 3 4 5 Lane

Figure 3.3: Sm and Lsm proteins bind to TER1 RNA. a, Northern blot of RNA isolated from immunoprecipitations with anti-c-Myc antibodies. Input and immunoprecipitation (IP) supernatant (s/n) represent 10% of the sample. An asterisk marks the position of the TER1 precursor. The lower band corresponds to the mature form of TER1. b, Northern blot analysis for TER1 from the Trt1, Sm and Lsm immunoprecipitation used in the telomerase assay shown in Figure 3.3d. c, RNA from anti-c-Myc immunoprecipitates was analysed by RT–PCR using primers in the first and second exon (primers represented by arrows in the schematic below the gel) to amplify the precursor form(upper panel).The primer pair alsoamplify the spliced form(lower band in Sm immunoprecipitates). Aprimer pair in the first exon was used to visualize all forms of TER1 combined (lower panel). d. Telomerase activity assay performed on beads after c-Myc immunoprecipitation of tagged proteins as indicated above each lane. Activity was quantified relative to the Trt1 immunoprecipitate sample. A 100-mer [^{32}P] oligonucleotide was used as recovery and loading control (LC).

To determine whether Sm and/or Lsm are associated with active telomerase, direct in vitro activity assays were performed. Telomerase activity was detected in all samples, but was 20-fold higher in Lsm3 and 4 than Smb1 and Sme1 immunoprecipitates (Fig. 3.3d). In part, this can be explained by the lower recovery of telomerase with Sm proteins, as judged by quantification of telomerase RNA on northern blots (Fig. 3.3b). However, even after normalization to the amount of TER1 in each immunoprecipitate, Lsm-associated telomerase activity was still 2.8-fold higher than that associated with Sm proteins. The simplest explanation for this observation is that a fraction of Sm-associated TER1 is not yet associated with the catalytic subunit of telomerase. Indeed, further experiments confirmed that Sm binding precedes Trt1 binding to TER1.

III.3c. The Sm and Lsm2-8 complexes directly bind to the Sm site

As the spliceosome snRNP contains Sm proteins, the TER1–Sm interaction may reflect binding of the spliceosome to the TER1 pre-mRNA. To test whether Sm proteins bind TER1 directly, we generated constructs with either a mutant 5'ss which eliminates the association of spliceosome with pre-mRNA (Bahler et al., 1998) or a deletion of the intron. In both mutants, RNAs are co-immunoprecipitated with Smb1 (Fig. 3.4a). In contrast, replacing the Sm-binding sequence with a random sequence (ter1-sm6 mutant) reduced Sm association by 22-fold based on quantitative RT-PCR analysis (Fig. 3.4b). Similarly, Lsm association was undetectable for ter1–sm6 by northern blot analysis (Fig. 3.4c). We therefore surmised that Sm and Lsm proteins directly bind to the previously identified site in TER1.

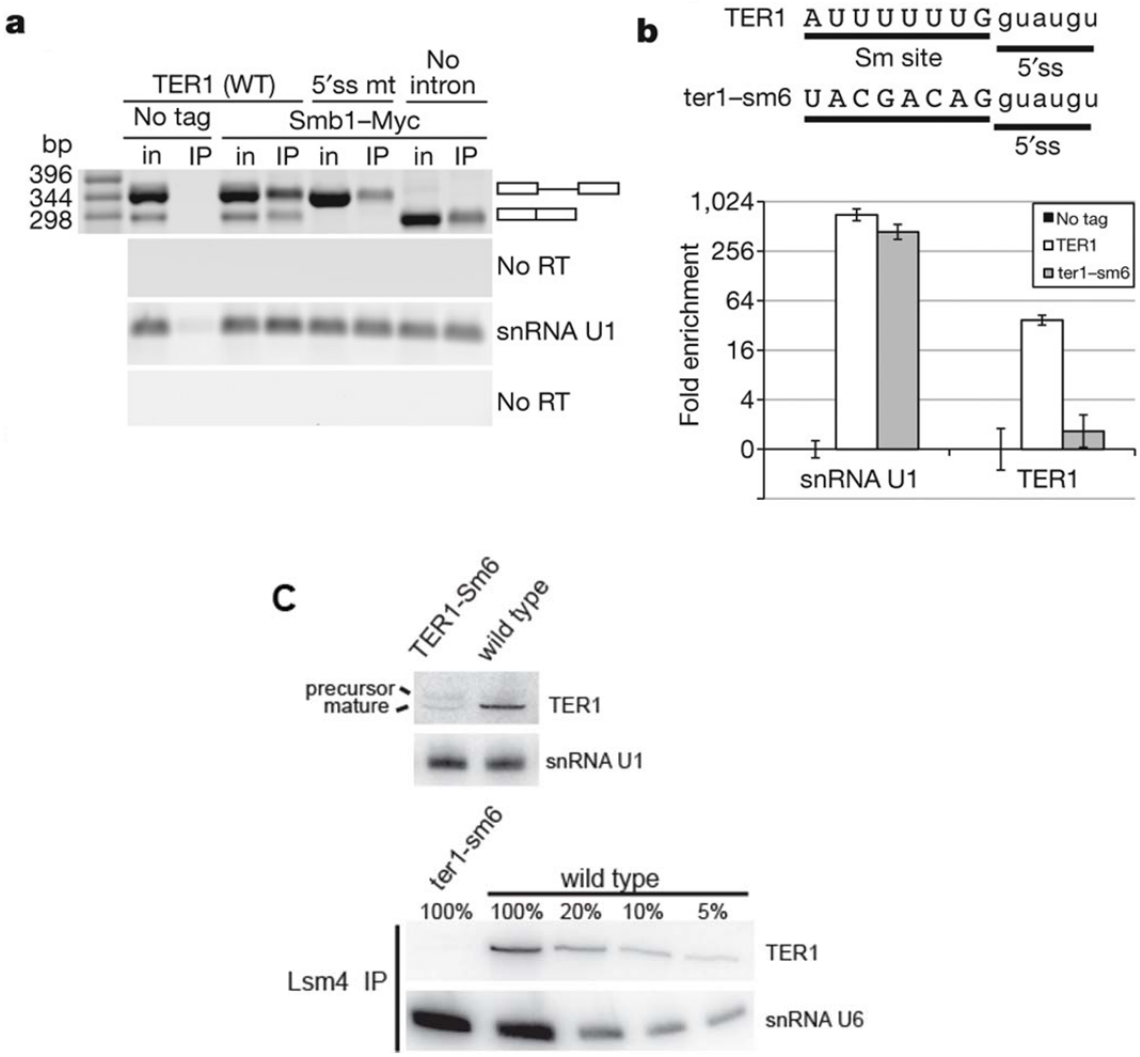

Figure 3.4: Sm and Lsm proteins directly bind to TER1 through the Sm site. a. Sm association does not require spliceosome assembly on TER1. RT–PCR was performed on RNA purified from input (in) and anti-c-Myc immunoprecipitate beads (IP). Primers amplifying snRNA U1 were used as a positive control. b, The Sm-binding site (upper case) and 5'ss for wild-type TER1 and the ter1–sm6 mutant. Replacing the Sm-binding site on TER1 (ter1–sm6 mutant) compromises Sm association. RNA recovered from anti-c-Myc immunoprecipitates from untagged control and Smb1–Myc strains was quantified by real-time PCR. Data are plotted as enrichment over the untagged control. Error bars, standard error of triplicate experiments. c. Loss of the Sm site compromises Lsm association. Total RNA samples were analysed by northern blot for TER1 and snRNA U1as the loading control (Top). Lsm4 immunoprecipitations were carried out in parallel from cell extracts containing wild type or the ter1-sm6 mutant followed by northern analysis. Whereas 5% of wild type TER1 was readily detected, ter1-sm6 was undetectable. U6 snRNA served as a control (Bottom).

III.3d. The Sm complex stimulates splicesomal cleavage

The physical overlap between the Sm site and 5'ss raised the question of whether Sm binding is involved in 3' end processing by the spliceosome. We have already noticed that loss of Sm binding in the ter1-sm6 mutant resulted in a sevenfold reduction in the processed form and an increase of TER1 precursor (Fig. 3.4c). Furthermore, a series of deletion mutants within the Sm site caused progressive inhibition of TER1 cleavage (Fig. 3.5a). Finally, introducing an eight nucleotide spacer between the Sm site and 5'ss also impaired processing (Fig. 2e). In summary, weakening or abolishing Sm association with the TER1 precursor reduces spliceosomal cleavage, indicating that Sm proteins promote 3' end processing of TER1.

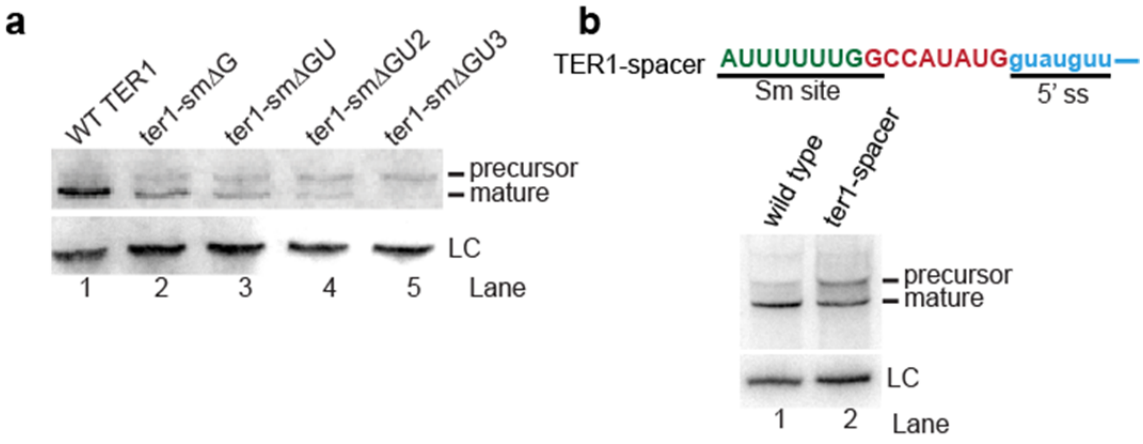

Figure 3.5: The Sm binding affects 3' end processing by the spliceosome. a, Shortening the Sm site compromises TER1 processing. Deleted nucleotides are indicated in the mutant names. Northern blot for TER1 and snoRNA snR101 as loading control. b, Increasing the distance between the Sm site and 5'ss in the ter1–spacer mutant impairs TER1 processing. Northern blot for TER1 and snoRNA snR101 as loading control.

III.3e. Binding of the Sm complex is required for TMG cap formation in TER1

A conserved feature among yeast and mammalian telomerase RNAs is the post-transcriptional hypermethylation of the 5' cap into a 2,2,7- trimethyl guanosine (TMG) form (Cech and Lingner, 1997; Leonardi et al., 2008; Seto et al., 1999). Sm proteins were first implicated in promoting cap hypermethylation on U2 snRNA in Xenopus extract (Hansson et al., 1974). Additional analysis showed *that in vitro* TMG-capping of human U1 requires the presence of SmB-SmD3 (Raker et al., 1996). A yeast two-hybrid screen for physical association with Sm proteins led to the identification of the methylase Tgs1 in budding yeast (Mouaikel et al., 2002). To elucidate the roles of Sm and/or Lsm in the hypermethylation of the 5'cap on TER1, we tested which, if any, of these proteins interact with *S. pombe* Tgs1 by two-hybrid analysis. Three Sm proteins scored positive, with Smd2 displaying the strongest interaction, and the other Sm proteins and all Lsm proteins showing weak or no interaction (Fig. 3.6a).

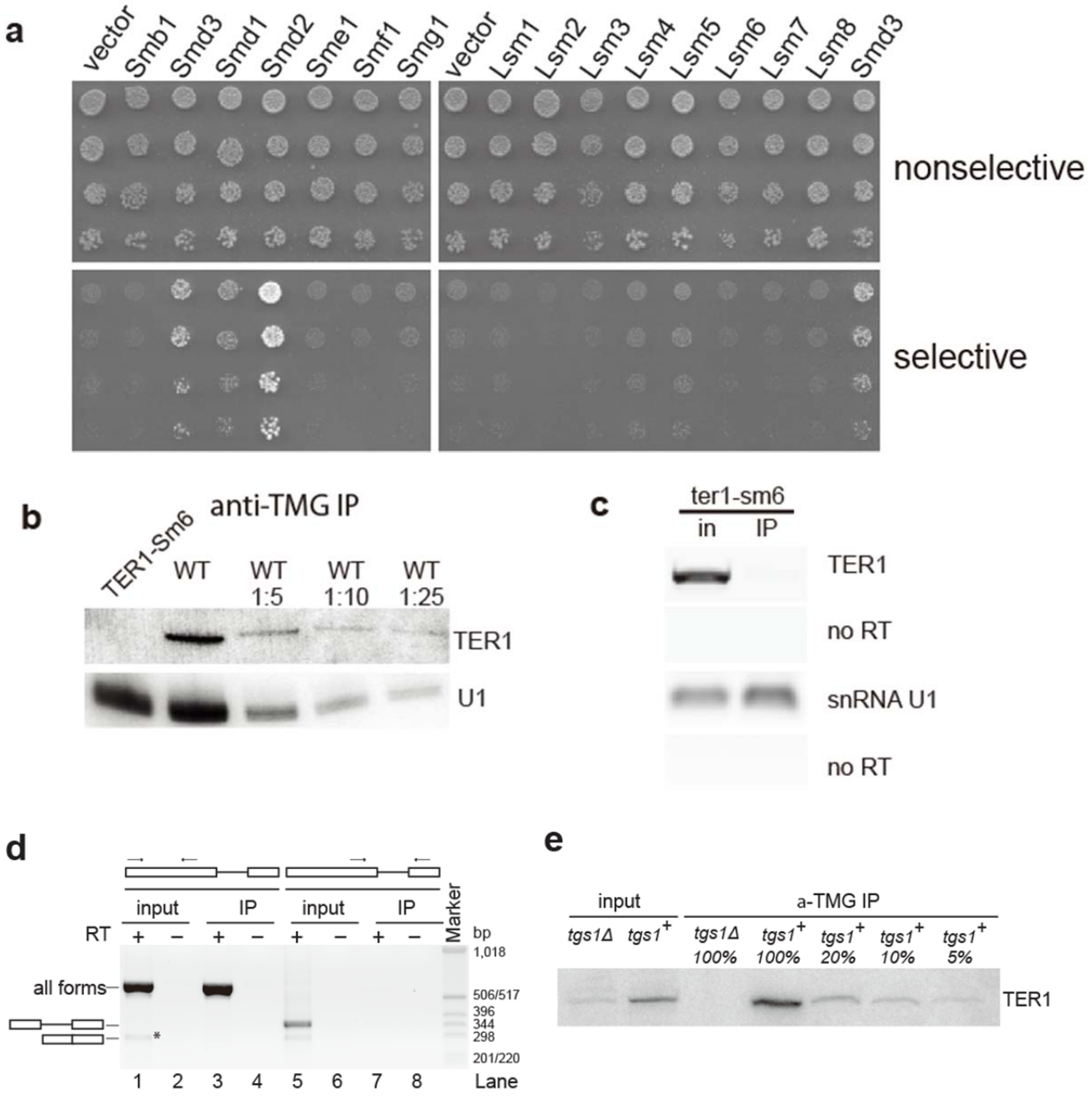
a
vector
Smb1
Smd3
Smd1
Smd2
Sme1
Smf1
Smg1
vector
Lsm1
Lsm2
Lsm3
Lsm4
Lsm5
Lsm6
Lsm7
Lsm8
Smd3
nonselective
selective
b
anti-TMG IP
TER1-Sm6
WT
WT 1:5
WT 1:10
WT 1:25
TER1
U1
c
ter1-sm6
in
IP
TER1
no RT
snRNA U1
no RT
d
input
IP
input
IP
Marker
RT
+ − + − + − + −
bp
1,018
all forms
506/517
396
344
298
201/220
*
1 2 3 4 5 6 7 8
Lane
e
input
a-TMG IP
tgs1Δ
tgs1+
tgs1Δ 100%
tgs1+ 100%
tgs1+ 20%
tgs1+ 10%
tgs1+ 5%
TER1

Figure 3.6: Sm binding is required for cap hypermethylation on TER1 and Tgs1modifies TER1. Yeast two-hybrid assays using Tgs1 as bait with Sm and Lsm proteins as indicated. Columns represent 3-fold dilutions on non-selective medium (top) and selective medium (bottom). b, Loss of Sm site compromises TMG cap formation. Northern blot of TER1 and ter1-sm6 from α-TMG IP samples. A dilution series for the wild type sample was included to show that 4% of wild type levels are detectable by this method. c, Loss of Sm site compromises TMG cap formation. RT–PCR amplifying all forms of TER1 and ter1–sm6 mutant from anti-TMG immunoprecipitation (IP) and input (in) samples; snRNAU1served as control. d, Precursor and spliced form are not TMG-capped. RNA isolated from input and α-TMG IP samples was subjected to RT-PCR analysis to detect the presence of TER1 (all forms, lanes 1 to 4) or specifically the precursor and spliced forms (lanes 5 to 8). An asterisk marks a non-specific and not reproducible band observed in lane 1. e, Tgs1 is responsible for 5' cap hypermethylation on TER1. Northern analysis of TER1 from α-TMG IP samples from wild type and *tgs1Δ* strains.

We next examined whether loss of Sm binding to TER1 affects TMG cap formation. Whereas wild-type TER1 was readily precipitated with a monoclonal antibody against the TMG cap, ter1-sm6 recovery was at least 25-fold reduced based on the results from northern blot analysis (Fig. 3.6b) and RT-PCR (Fig. 3.6c). Interestingly, only the cleaved form of TER1 was recovered in TMG immunoprecipitations from wild-type cells, suggesting that spliceosomal cleavage precedes hypermethylation (Fig 3.6d). Further we showed that TER1 was not TMG-capped in a *tgs1Δ* strain, confirming that Tgs1 is the sole enzyme responsible for TER1 cap hypermethylation (Fig. 3.6e).

III.3f. Tgs1 is required for TER1 biogenesis and telomere maintenance

In light of the reported increase in telomerase RNA and longer telomere in *tgs1Δ* budding yeast (Franke et al., 2008), we were surprised to observe a fivefold reduction in mature TER1 RNA in *tgs1Δ* compared with wild type in *S.*

pombe (Fig. 3.7a). In addition, an increase in the precursor indicated a 3' end processing defect. The viability of *tgs1Δ* cells ruled out a major splicing defect, but we consistently noted a small reduction in spliceosomal snRNAs isolated from *tgs1Δ* cells (Fig. 3.7a). To differentiate between a processing defect and a direct effect of the TMG cap on TER1 stability, we mutated the spliceosomal cleavage site and inserted a hammerhead ribozyme sequence to generate the mutant ter1-5'ssmut-HH (Fig. 3.7c). In this construct, processing of TER1 occurs independently of the spliceosome by ribozyme cleavage. When comparing ter1–5'ssmut-HH levels between wild-type and *tgs1Δ* cells, a twofold reduction was observed (Fig. 3.7a). Taken together, these results show that *tgs1Δ* affectsTER1 processing by the spliceosome as well as TER1 stability. Consequentially, the reduction of TER1 in *tgs1Δ* leads to shortened telomeres (Fig. 3.7b).

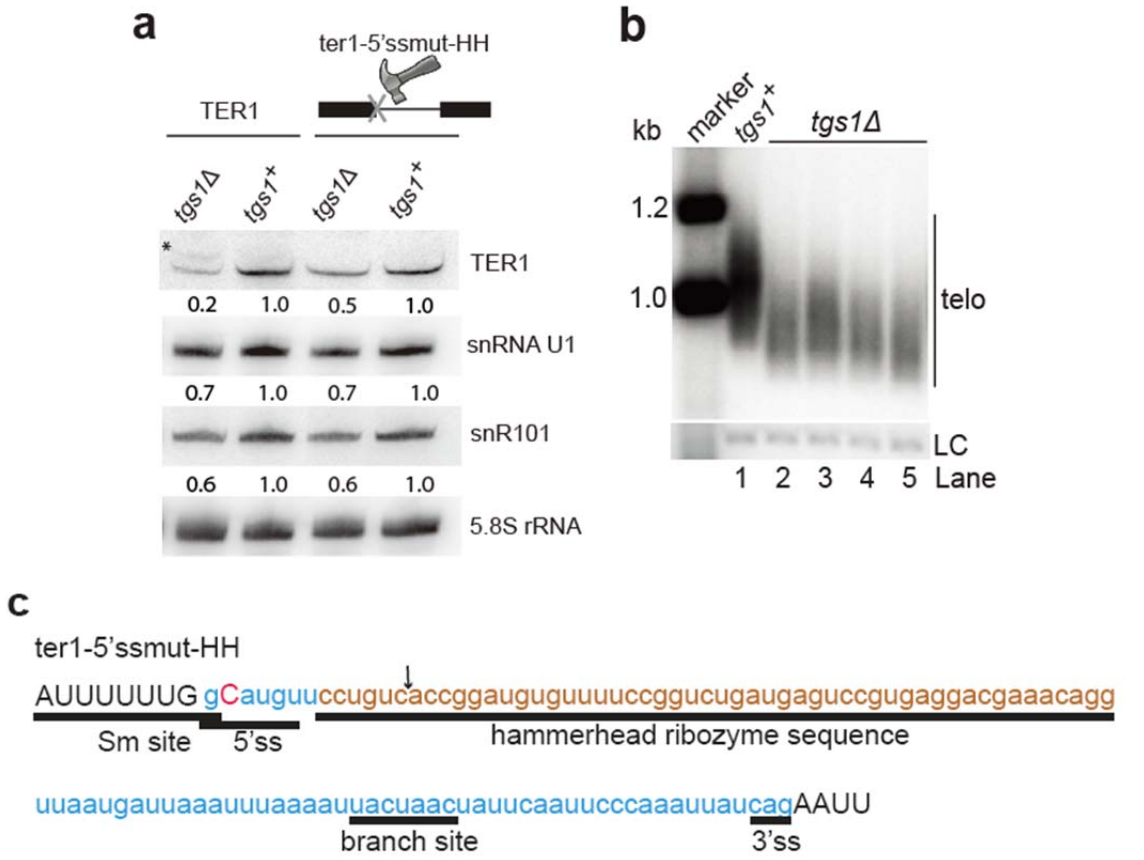

Figure 3.7: Both TER1 level and telomere length decrease in *tgs1Δ* strain. a, Bypass of spliceosomal cleavage reveals functions of Tgs1 in TER1 processing and stability. Northern blot analysis of TER1, snRNA U1, snR101 and 5.8 s rRNA from total RNA prepared from wild-type and *tgs1Δ* strains having either TER1 or the ter-5'ssmut-HH mutant. An asterisk marks the position of the TER1 precursor. b, Deletion of tgs11 causes telomere shortening. Telomere length was analyzed by Southern blotting of EcoRI-digested genomic DNA from four independent *tgs1Δ* isolates and an otherwise isogenic tgs11 strain. A probe for the rad16 gene was used as a loading control (LC). c, Schematic of ter1-5'ssmut-HH mutant. The 5'ss was mutated and a hammerhead ribozyme sequence was inserted downstream. The hammerhead cleavage site is indicated with a vertical arrow.

III.3g. TMG cap on TER1 does not stimulate telomerase activity or the binding of Sm and Lsm proteins

TMG cap is bound by specific nuclear-import receptor called Snurportin1 which promotes reimporting RNA back to nucleus in vertebrates (Huber et al., 1998).

The ortholog of Snurportin1 in yeast is absent, so actual function of TMG cap on TER1 is not known. Considering the interplays of Sm, Lsm and TMG cap, we next examined whether the TMG cap is important for telomerase activity or the binding of Sm/Lsm proteins. *In vitro* telomerase activity assay and Sm/Lsm immuneprecipitation were perform in the *tgs1Δ* strain. Neither telomerase activity nor Lsm association was reduced beyond the effects expected from the reduced steady-state level of TER1 (Fig. 3.8a and Fig. 3.8b).

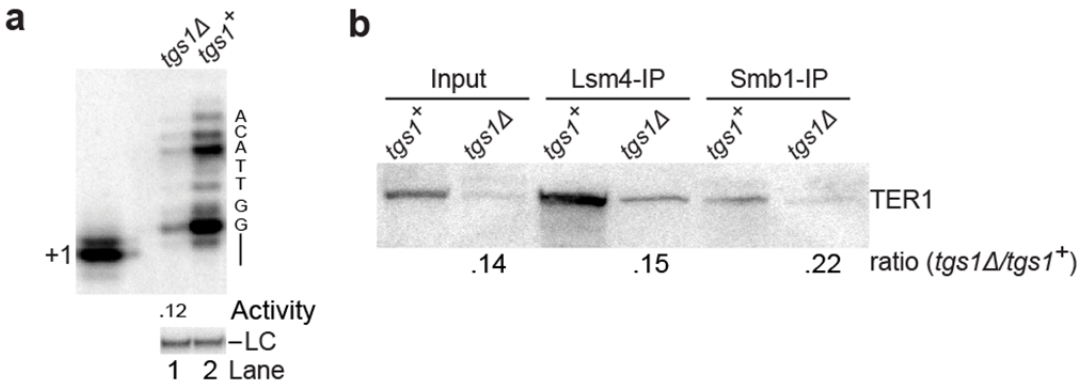

Figure 3.8: Deletion of tgs1^{+} does not impair telomerase activity or the association of Sm and Lsm proteins. a, Deletion of tgs1^{+} does not impair telomerase activity recovered from Lsm4 IPs beyond the effect expected from the reduced TER1 level in *tgs1Δ* cells. Activity of the tgs1Δ sample is indicated relative to wild type. A ^{32}P-labelled 100mer oligonucleotide was used as recovery and loading control. b, Northern blot for TER1 using RNA isolated from Lsm4 and Smb1 immunoprecipitations. For input and each IP, the ratio between the tgs1 deletion and wild type is shown below the *tgs1Δ* lane.

III.3h. Lsm proteins replace Sm proteins at 3' end of TER1

Most TER1 post-spliceosomal cleavage was bound by Lsm2–8, but a small fraction was associated with Sm proteins (Fig. 3.3a). To investigate whether this was indicative of a switch from Sm to Lsm binding, we examined the distribution

of 3’ ends in each immunoprecipitation by massively parallel sequencing (Fig. 3.2a). Around 70% of Sm-bound TER1 post cleavage terminated precisely at the spliceosomal cleavage site (Fig. 3.9a). Enrichment of this form in the Sm-bound fraction is consistent with Sm proteins binding the TER1 precursor and remaining associated with TER1 until after cleavage and cap hypermethylation have occurred. In contrast, Lsm-associated TER1 predominantly terminated in $U_{3–6}$, indicating that a switch between Sm and Lsm binding occurs after spliceosomal cleavage and is associated with exonucleolytic processing (Fig. 3.9b). Consistent with most telomerase activity being associated with Lsm2–8, the TER1 3'end distribution from Trt1 immunoprecipitates was indistinguishable from that of Lsm-bound TER1.

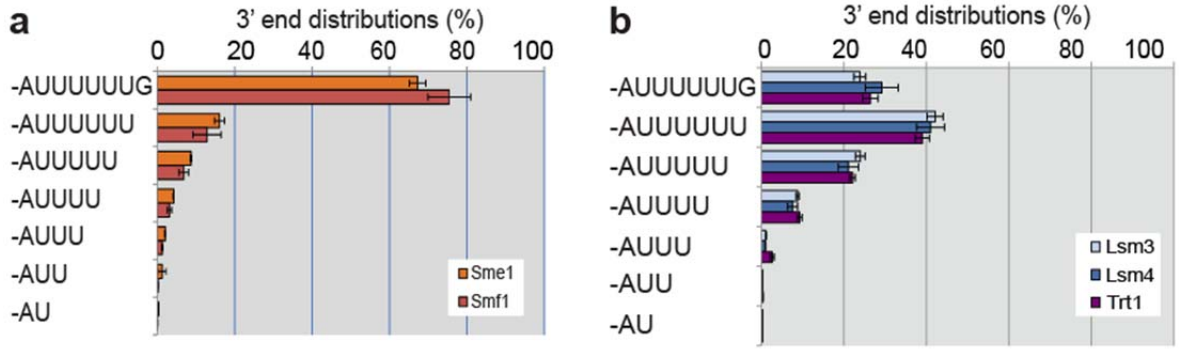

Figure 3.9: The Sm complex and Lsm2-8 complex sequentially bind to TER1. a, 3’ end sequence analysis of TER1 from Sme1 and Smf1 IPs. Average and standard deviation from two experiments, number of sequences analyzed: 3.9 x 10^6 (Sme1) and 7.9 x 10^6 (Smf1). b, 3’ end sequence analysis of TER1 Lsm3 (four experiments, 21.2 x 10^6 sequence reads), Lsm4 (three experiments, 11.4 x 10^6 sequence reads) and Trt1 (three experiments, 15.2 x 10^6 sequence reads).

III.3i. Lsm proteins protect TER1 from degradation

The observation that loss of Sm binding coincided with the loss of terminal nucleotides led us to speculate that Lsm2–8 may function in protecting the 3' end of TER1 against further exonucleolytic degradation. To test this, we attempted to generate Lsm deletion strains. Whereas most Lsm proteins are essential, *lsm1 Δ* and *lsm3Δ* cells were viable. Consistent with a protective function for Lsm2–8, the levels of TER1 and U6 snRNA were reduced approximately five fold in *lsm3Δ* cells, while the level of U1 snRNA remains the same (Fig. 3.10b). No such effect was seen when deleting lsm1. The 3' end sequence distribution for TER1 from total RNA of *lsm3Δ* cells closely resembled the Sm-bound fraction in wild type, whereas the Lsm-bound fraction was selectively lost in the mutant (Fig. 3.10a). The viability of *lsm3Δ* cells further allowed us to confirm that cap hypermethylation is unaffected by the absence of Lsm consistent with Tgs1 acting on TER1 before Lsm binding (Fig. 3.10c).

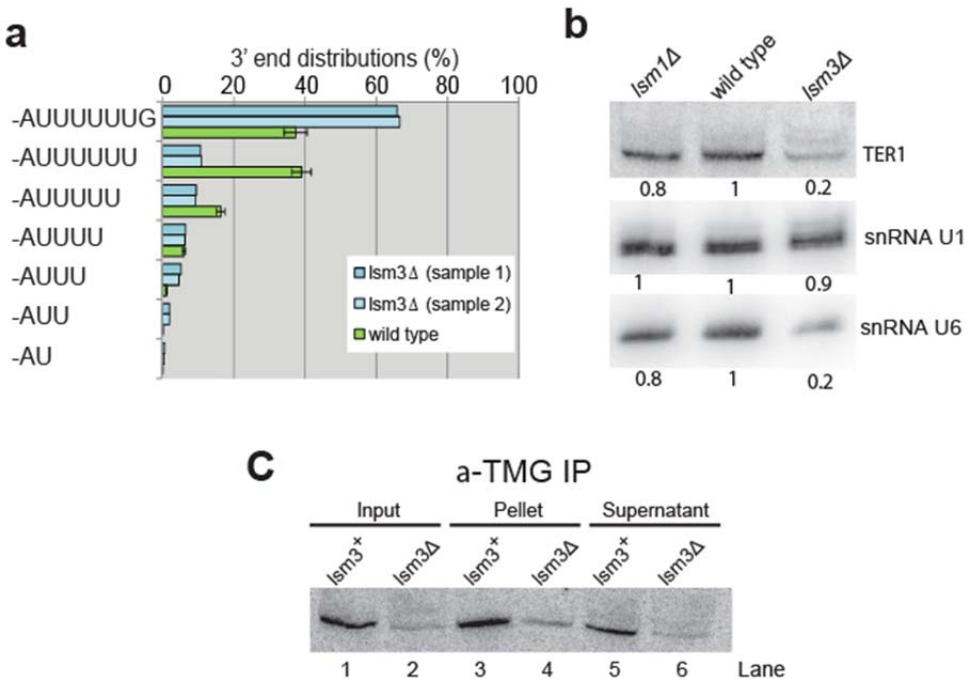

Figure 3.10: Lsm protects the 3' end of the mature form of TER1. a, Specific loss of Lsm2-8 bound fraction of TER1 in *lsm3Δ* cells based on 3' end sequence analysis from two total RNA samples (1.7×10^6 and 3.3×10^6 sequence reads were analyzed). b, Northern blot analysis from total RNA prepared from wild-type, *lsm1Δ* and *lsm3Δ* strains, quantified relative to wild type for each RNA. c, Deletion of $lsm3^+$ affects TER1 processing and stability but not 5' guanosine cap hypermethylation. Northern blot for TER1 with RNA isolated from α-TMG IP samples from wild type and *lsm3Δ* strains.

III.3j. The Lsm2-8 complex facilitates telomerase assembly

To verify independently a role for Lsm proteins in stabilizing TER1, we took advantage of the observation that Lsm binding requires a stretch of consecutive uridines(Achsel et al., 1999). In contrast, Sm binding tolerates other nucleotides in certain positions of the binding motif, as exemplified by the Sm-binding site in human U1 snRNA (AAUUUGUG). When the TER1 Sm site was mutated to reduce the number of consecutive uridines, the level of mature TER1 was decreased (Fig. 3.11a). We next precipitated Smb1, Lsm4 and Trt1 from wild

type and strains containing the ter1-SmU1 mutant. As expected, the mutation had little effect on the binding of Sm proteins (Fig. 3.11b). In fact, when normalized for the lower level of ter1-SmU1 compared with wild type, recovery of ter1–SmU1 with Smb1 was increased 1.6-fold. In contrast, Lsm binding was diminished by more than 20-fold. Most surprisingly, the interaction between the catalytic subunit Trt1 and telomerase RNA was also compromised in the ter1-SmU1 mutant (Fig. 3.11b). The normalized recovery of ter1-SmU1 with Trt1 was 15-fold lower than wild type, indicating that Lsm binding facilitates Trt1-TER1 association, possibly by inducing a conformational change in the RNA analogous to how binding of the p65 protein facilitates telomerase assembly in Tetrahymena (Stone et al., 2007). Consistent with the poor recovery of ter1-SmU1 in Trt1 immunoprecipitations, in vitro telomerase activity was below the level of detection (Fig. 3.11c).

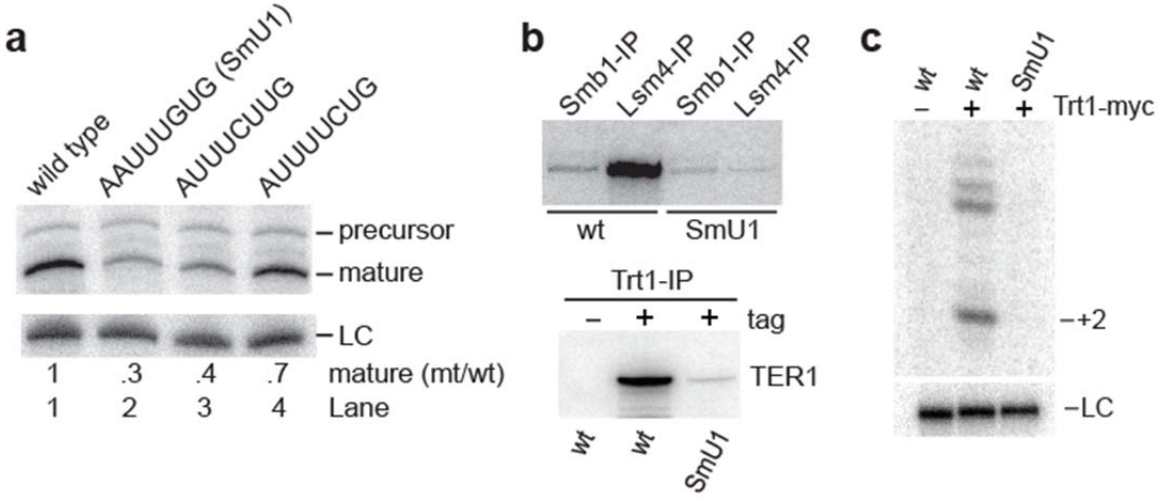

Figure 3.11: Lsm binding toTER1 promotes telomerase assembly and protects TER1 from degradation. a, Northern blot for TER1. The indicated ratios of mutant to wild type (WT) are normalized to the loading control snR101. b, Northern blot for TER1 and the ter1–SmU1 mutant using RNA isolated from anti-c-Myc immunoprecipitations performed on extract from strains harboring Smb1-Myc, Lsm4-Myc or Trt1-Myc as indicated. c, Telomerase activity assay performed on Trt1 immunoprecipitates from strains harboring either wild type or ter1-SmU1. An untagged Trt1 strain was used as negative control.

III.3K. 3' end processing pathway might be conserved among fission yeast species

To test whether this unique 3' end processing pathway of telomerase RNA is conserved in other fission yeast species, we cloned telomerase RNA from *Schizosaccharomyces cryophilus* (*S. cryophilus*) (Helston et al., 2010) and *Schizosaccharomyces octosporus* (*S. octosporus*). Sequences matching for the 5'ss, branch site and 3'ss (Zhang and Marr, 1994) are found in both telomerase RNAs (Fig. 3.12c and d). We reasoned that TER1 intron is necessary and sufficient to induce spliceosomal cleavage (Box et al., 2008 and data not shown) and speculated that introns of telomerase RNAs from *S. cryophilus* and *S.*

octosporus can also under the similar process of cleavage. Such a model predicts that introns of telomerase RNA from other two fission yeast species would substitute for the TER1 intron. We replaced TER1 intron with the heterologous intron from *S. cryophilus* and *S. octosporus* telomerase RNA respectively. RT–PCR across the intron revealed that those two mutants had similar splicing efficiency as wild type TER1 (Fig. 3.12a). When total RNA was examined by northern blotting, both of the mutants could efficiently produce the cleaved form (Fig. 3.12b). The success of heterologous telomerase RNA introns to substitute for the TER1 intron suggests that the spliceosomal cleavage occurs in other yeast species.

Next, we examined the mature 3' end of telomerase RNAs from *S. cryophilus* and *S. octosporus.* Different from TER1 where 5'ss and Sm site are overlapped, telomerase RNAs in those two species have extra nucleotides between the Sm site and the 5'ss (Fig. 3.12c and d). Cloning and sequencing of 3' end of both telomerase RNAs revealed that further exonucleolytic processing occurs after cleavage (Fig. 3.12c and d). Similar to TER1, most telomerase RNAs from the other two yeast species terminated at a stretch of uridines. We speculate that the truncated Sm site could serve as target for Lsm binding. The association of those telomerase RNAs with Lsm proteins needs to be further verified experimentally.

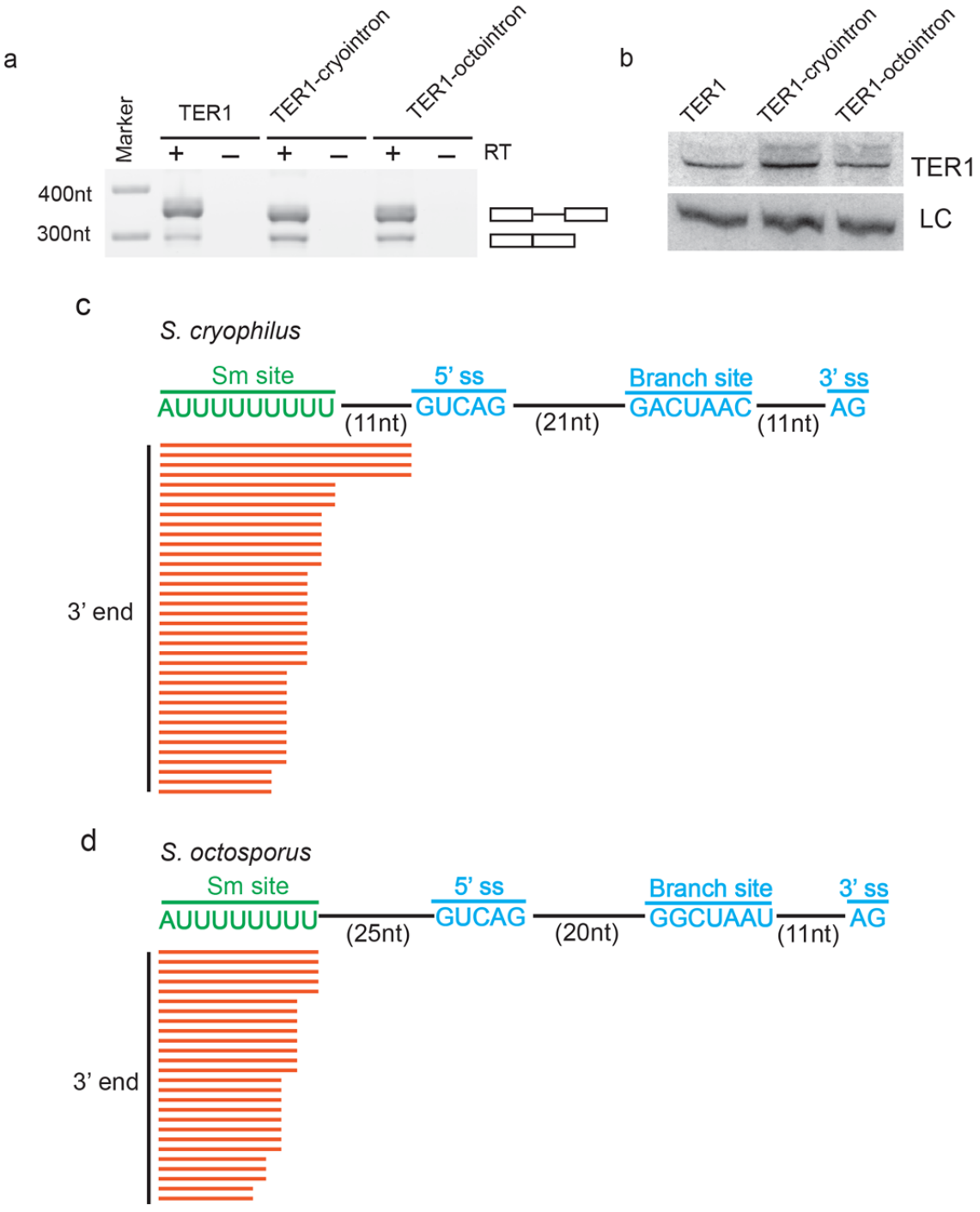

Figure 3.12: 3' end processing of telomerase RNA is conserved among fission yeast species. a, RT-PCR across the intron to assess the abundance of the spliced form. b, Northern blot for TER1 and the loading control (LC) c, Schematic of the Sm site and intron sequences. 3' end sequences of mature telomerase RNA (based on 36 clones) in *S. cryophilus.* d, Schematic of the Sm site and intron sequences and 3' end sequences of mature telomerase RNA (based on 26 clones) in *S. octosporus.*

III.4: Discussion

Unlike the evolutionarily conserved telomerase catalytic subunit, telomerase RNAs are distinct in length and sequences among species (Chen et al., 2000). Using biochemical, genetic and bioinformatic approaches, telomerase RNA genes have been identified in several model organisms. However, their biogenesis pathways have so far remained elusive. One important aspect of biogenesis is the 3' end processing of RNA. Until recently, the mechanisms for generation of mature 3' end of telomerase RNA were unknown. In *S. pombe*, telomerase RNA is first transcribed as a polyadenylated precursor, then spliceosomal cleavage generates the 3' end of telomerase RNA (Box et al., 2008). In *Saccharomyces cerevisiae*, the process is directed by the RNA-binding proteins Nrd1, Nab3, and the RNA helicase Sen1 (Jamonnak et al., 2011; Noel et al., 2012). In vertebrates, the presence of precursor and the termination mechanism so far have not been discovered. Interestingly, telomerase RNAs in different yeast species all terminate at a uridine-rich motif termed Sm binding site which is known to be bound by Sm proteins.

Sm proteins have been studied extensively in the context of their association with the spliceosomal snRNAs U1, U2, U4 and U5, where they play critical roles in snRNP assembly and function (Patel and Bellini, 2008). They have also been suggested to be involved in telomerase RNA biogenesis in several yeast species (Box et al., 2008; Gunisova et al., 2009; Seto et al., 1999). Lsm proteins associate with a wide spectrum of RNAs, including tRNA and

polymerase III-transcribed U6 snRNA (Khusial et al., 2005). Despite their structural similarity and related binding motifs, Sm and Lsm complexes have different modes of RNA binding and were thought to have distinct and non-overlapping sets of target RNAs.

Our finding that the TER1 precursor is exclusively associated with the Sm complex, whereas most mature TER1 is bound by Lsm2–8, revealed that biogenesis of telomerase RNA involves both Sm and Lsm complexes (Fig. 4.1). Both complexes play different roles during TER1 maturation. The Sm complex promotes spliceosomal cleavage and 5’ guanosine cap hypermethylation, while the Lsm2-8 complex prevents 3’-5’ degradation and facilitates the loading of telomerase catalytic subunit. Considering the central roles that Sm and Lsm proteins play in RNA metabolism, it will be important to determine whether biogenesis of other non-coding RNAs also involves Sm- and Lsm2–8-bound stages. Furthermore, it is interesting to note that several human Sm/Lsm proteins have been reported to co-purify with telomerase (Fu and Collins, 2007) raising the possibility that these proteins also function in TMG cap formation and telomerase assembly in metazoans.

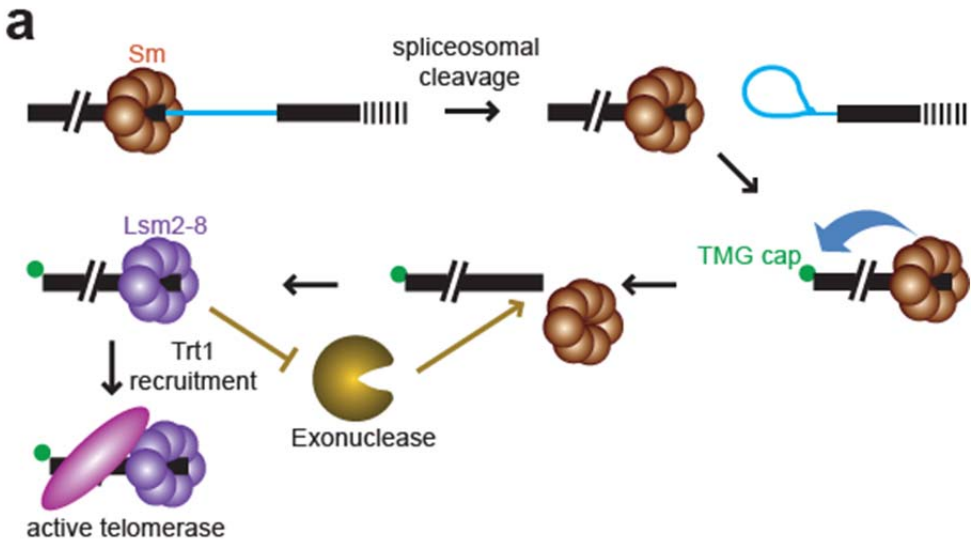

Figure 4.1: Sequence of events during telomerase RNA biogenesis in fission yeast

Chapter IV: Conclusions and Future Directions

IV.1: Future directions for chapter III

The telomerase RNA plays essential roles in telomerase activity. A portion of it serves as a template for DNA synthesis. It also functions as a scaffold for the protein subunits (Zappulla and Cech, 2004). Despite its importance, telomerase RNA has not been identified in numerous species due to its wide variance in length and sequences. Even in the organisms whose telomerase RNAs are known, the biogenesis pathway is poorly studied. The identification of telomerase RNA in fission yeast from our lab (Leonardi et al., 2008) provided us an opportunity to study the sequence of events that lead from the transcription of the gene encoding the telomerase RNA subunit TER1 to the incorporation of

a processed and functional form of this non-coding RNA into a complex with the catalytic subunit of telomerase.

The observation that TER1 could sequentially be bound by Sm and Lsm complexes constitutes the first identification of RNA whose biogenesis involves both closely related proteins. Additionally, we demonstrated that these two complexes have distinct and critical roles in TER1 maturation. The Sm complex stimulates spliceosomal processing and TMG capping. The Lsm complex protects 3' end of TER1 from exonucleolytic degradation and promotes the binding of the catalytic subunit. While I believe that our work sets a paradigm to test similar functions of Sm and Lsm proteins in telomerase RNAs from other species and other noncoding RNAs, the more important contribution of the research presented in Chapter Three are the new questions raised by these results.

The first and most pressing question put forth pertains to the mechanism by which the Lsm2-8 complex facilitates the association of TERT with telomerase RNA. One possibility is that the Lsm complex could directly bind to and recruit TERT onto the RNA. Yeast two hybrid system and in vitro protein binding analysis are good approaches to test this hypothesis. Alternatively, the binding of Lsm proteins could induce conformational changes of TER1 which favors TERT association. This hypothesis can be tested *in vitro* by many straightforward strategies, including X-ray crystallography and structure probing with specialized chemicals or ribonucleases.

The second question put forth by Chapter Three is to determine mechanism by which Sm promotes spliceosomal cleavage. Sm proteins are essential and ubiquitously expressed in eukaryotic cells, but their exact functions are largely unknown (Seraphin, 1995). Sm proteins have been suggested to function in TMG capping of U2 snRNA in Xenopus extract (Mattaj, 1986). Here we showed they perform the same function in telomerase RNA by directly interacting with the capping enzyme Tgs1. More surprisingly, we found that Sm has additional roles in promoting spliceosomal cleavage. Specific sequence elements termed exonic splicing enhancers (ESEs) have been shown to promote splicing (Blencowe, 2000). Those ESEs are bound by SR-family proteins, which function early in spliceosome formation to promote the formation of complexes containing U1 snRNP bound to the 5' splice site and U2 snRNP bound to the pre-mRNA branch site (Zahler and Roth, 1995). It will be interesting to define whether the Sm site act as an ESE and binding of Sm proteins can facilitate spliceosome assembly in pre-mRNA.

The third question pertains to functions of Lsm and Sm proteins in biogenesis of human telomerase RNA. Several Sm/Lsm proteins have been reported to co-purify with human telomerase (Fu and Collins, 2006, 2007). However, their roles have not been analyzed. Observations in our model system lead us to hypothesize that Sm/Lsm binding might be important for the maturation of human telomerase RNA. Another conserved feature among mammalian and yeast telomerase RNAs is the presence of the TMG cap (Jady et al., 2004). We

can analyze the function of the TMG cap in TER1 processing by knocking down Tgs1 in cell lines or using *tgs1* knock-out (KO) mice. In about half of patients with dyskeratosis congenita, the disorder is caused by mutations in the TERT, telomerase RNA, DKC1, or TIN2 genes which are known to involve in telomerase maintenance. However, the cause of the disorder in other affected individuals is unknown. If we prove Sm/Lsm and Tgs1 have direct roles in telomerase biogenesis, we can further map mutations in those genes from patients with dyskeratosis congenita.

The last question is to define which RNAs are bound by Sm and Lsm proteins on a genome-wide scale. Only a few RNAs have been shown to be bound by Sm or Lsm complexes so far (Khusial et al., 2005). Our finding that both Sm and Lsm associate with telomerase RNA raises the question whether they bind to other RNAs. Utilizing second generation sequencing technology (Schuster, 2008) will allow us to discover RNAs associated with each Sm/Lsm protein on a genome-wide scale. We are particularly interested in three classes of RNAs: (i), RNAs bound by both Sm and Lsm complexes. The finding of those RNAs would be an indication that the unique maturation pathway involving Sm and Lsm is not limited in telomerase RNA, but apply to other RNAs. (ii), RNAs bound to certain Sm proteins and certain Lsm proteins indicative of the existence of previously unknown “hybrid” rings. So far only one hybrid ring has been reported which plays the essential roles in histone mRNA processing in metazoans (Pillai et al., 2003). We expect to find evidences for other hybrid

complexes through this analysis. (iii), Uncharacterized RNAs associated with Sm or Lsm proteins. High-throughput sequencing technology revealed a large and complex transcriptome in *S. pombe* (Wilhelm et al., 2008). The recent finding of long non-coding RNAs adds another layer for the transcriptome (Guttman and Rinn, 2012). Considering the central roles of Sm and Lsm proteins in RNA processing, it would not be surprising that they also are involved in the maturation of other unidentified RNAs.

References:

Achsel, T., Brahms, H., Kastner, B., Bachi, A., Wilm, M., and Luhrmann, R. (1999). A doughnut-shaped heteromer of human Sm-like proteins binds to the 3'-end of U6 snRNA, thereby facilitating U4/U6 duplex formation in vitro. EMBO J *18*, 5789-5802.

Achsel, T., Stark, H., and Luhrmann, R. (2001). The Sm domain is an ancient RNA-binding motif with oligo(U) specificity. Proc Natl Acad Sci U S A *98*, 3685-3689.

Aigner, S., Lingner, J., Goodrich, K.J., Grosshans, C.A., Shevchenko, A., Mann, M., and Cech, T.R. (2000). Euplotes telomerase contains an La motif protein produced by apparent translational frameshifting. EMBO J *19*, 6230-6239.

Albrecht, M., and Lengauer, T. (2004). Novel Sm-like proteins with long C-terminal tails and associated methyltransferases. FEBS Lett *569*, 18-26.

Allsopp, R.C., Chang, E., Kashefi-Aazam, M., Rogaev, E.I., Piatyszek, M.A., Shay, J.W., and Harley, C.B. (1995). Telomere shortening is associated with cell division in vitro and in vivo. Exp Cell Res *220*, 194-200.

Allsopp, R.C., Vaziri, H., Patterson, C., Goldstein, S., Younglai, E.V., Futcher, A.B., Greider, C.W., and Harley, C.B. (1992). Telomere length predicts replicative capacity of human fibroblasts. Proc Natl Acad Sci U S A *89*, 10114-10118.

Bae, N.S., and Baumann, P. (2007). A RAP1/TRF2 complex inhibits nonhomologous end-joining at human telomeric DNA ends. Mol Cell *26*, 323-334.

Bahler, J., Wu, J.Q., Longtine, M.S., Shah, N.G., McKenzie, A., 3rd, Steever, A.B., Wach, A., Philippsen, P., and Pringle, J.R. (1998). Heterologous modules for efficient and versatile PCR-based gene targeting in Schizosaccharomyces pombe. Yeast *14*, 943-951.

Baumann, P., and Cech, T.R. (2001). Pot1, the putative telomere end-binding protein in fission yeast and humans. Science *292*, 1171-1175.

Bergman, N., Moraes, K.C., Anderson, J.R., Zaric, B., Kambach, C., Schneider, R.J., Wilusz, C.J., and Wilusz, J. (2007). Lsm proteins bind and stabilize RNAs containing 5' poly(A) tracts. Nat Struct Mol Biol *14*, 824-831.

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

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

Blasco, M.A. (2005). Telomeres and human disease: ageing, cancer and beyond. Nat Rev Genet *6*, 611-622.

Blencowe, B.J. (2000). Exonic splicing enhancers: mechanism of action, diversity and role in human genetic diseases. Trends Biochem Sci *25*, 106-110.

Bodnar, A.G., Ouellette, M., Frolkis, M., Holt, S.E., Chiu, C.P., Morin, G.B., Harley, C.B., Shay, J.W., Lichtsteiner, S., and Wright, W.E. (1998). Extension of

life-span by introduction of telomerase into normal human cells. Science *279*, 349-352.

Bouveret, E., Rigaut, G., Shevchenko, A., Wilm, M., and Seraphin, B. (2000). A Sm-like protein complex that participates in mRNA degradation. EMBO J *19*, 1661-1671.

Box, J.A., Bunch, J.T., Tang, W., and Baumann, P. (2008). Spliceosomal cleavage generates the 3' end of telomerase RNA. Nature *456*, 910-914.
Cech, T.R. (2004). Beginning to understand the end of the chromosome. Cell *116*, 273-279.

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

Chai, W., Du, Q., Shay, J.W., and Wright, W.E. (2006). Human telomeres have different overhang sizes at leading versus lagging strands. Mol Cell *21*, 427-435.

Chang, E., and Harley, C.B. (1995). Telomere length and replicative aging in human vascular tissues. Proc Natl Acad Sci U S A *92*, 11190-11194.

Chapon, C., Cech, T.R., and Zaug, A.J. (1997). Polyadenylation of telomerase RNA in budding yeast. RNA *3*, 1337-1351.

Chari, A., Golas, M.M., Klingenhager, M., Neuenkirchen, N., Sander, B., Englbrecht, C., Sickmann, A., Stark, H., and Fischer, U. (2008). An assembly chaperone collaborates with the SMN complex to generate spliceosomal SnRNPs. Cell *135*, 497-509.

Chen, J.L., Blasco, M.A., and Greider, C.W. (2000). Secondary structure of vertebrate telomerase RNA. Cell *100*, 503-514.

Chen, J.L., and Greider, C.W. (2004). An emerging consensus for telomerase RNA structure. Proc Natl Acad Sci U S A *101*, 14683-14684.

Collins, K., and Mitchell, J.R. (2002). Telomerase in the human organism. Oncogene *21*, 564-579.

Cristofari, G., Adolf, E., Reichenbach, P., Sikora, K., Terns, R.M., Terns, M.P., and Lingner, J. (2007). Human telomerase RNA accumulation in Cajal bodies facilitates telomerase recruitment to telomeres and telomere elongation. Mol Cell *27*, 882-889.

Dandjinou, A.T., Levesque, N., Larose, S., Lucier, J.F., Abou Elela, S., and Wellinger, R.J. (2004). A phylogenetically based secondary structure for the yeast telomerase RNA. Curr Biol *14*, 1148-1158.

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

de Lange, T. (2009). How telomeres solve the end-protection problem. Science *326*, 948-952.

Dragon, F., Pogacic, V., and Filipowicz, W. (2000). In vitro assembly of human H/ACA small nucleolar RNPs reveals unique features of U17 and telomerase RNAs. Mol Cell Biol *20*, 3037-3048.

Feng, J., Funk, W.D., Wang, S.S., Weinrich, S.L., Avilion, A.A., Chiu, C.P., Adams, R.R., Chang, E., Allsopp, R.C., Yu, J.*, et al.* (1995). The RNA component of human telomerase. Science *269*, 1236-1241.

Ferrezuelo, F., Steiner, B., Aldea, M., and Futcher, B. (2002). Biogenesis of yeast telomerase depends on the importin mtr10. Mol Cell Biol *22*, 6046-6055.

Franke, J., Gehlen, J., and Ehrenhofer-Murray, A.E. (2008). Hypermethylation of yeast telomerase RNA by the snRNA and snoRNA methyltransferase Tgs1. J Cell Sci *121*, 3553-3560.

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

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

Gallardo, F., Laterreur, N., Cusanelli, E., Ouenzar, F., Querido, E., Wellinger, R.J., and Chartrand, P. (2011). Live cell imaging of telomerase RNA dynamics reveals cell cycle-dependent clustering of telomerase at elongating telomeres. Mol Cell *44*, 819-827.

Gallardo, F., Olivier, C., Dandjinou, A.T., Wellinger, R.J., and Chartrand, P. (2008). TLC1 RNA nucleo-cytoplasmic trafficking links telomerase biogenesis to its recruitment to telomeres. EMBO J *27*, 748-757.

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

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

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

Gunisova, S., Elboher, E., Nosek, J., Gorkovoy, V., Brown, Y., Lucier, J.F., Laterreur, N., Wellinger, R.J., Tzfati, Y., and Tomaska, L. (2009). Identification and comparative analysis of telomerase RNAs from Candida species reveal conservation of functional elements. RNA *15*, 546-559.

Guttman, M., and Rinn, J.L. (2012). Modular regulatory principles of large non-coding RNAs. Nature *482*, 339-346.

Haendeler, J., Hoffmann, J., Rahman, S., Zeiher, A.M., and Dimmeler, S. (2003). Regulation of telomerase activity and anti-apoptotic function by protein-protein interaction and phosphorylation. FEBS Lett *536*, 180-186.

Haering, C.H., Nakamura, T.M., Baumann, P., and Cech, T.R. (2000). Analysis of telomerase catalytic subunit mutants in vivo and in vitro in Schizosaccharomycespombe. Proc Natl Acad Sci U S A *97*, 6367-6372.

Hansson, L., Henningsen, N.C., Karlberg, B., Aberg, H., Jameson, S., Malmcrona, R., and Hersvall, C. (1974). Hypotensive action of ICI66.082, a new beta-adrenergic blocking agent. Int J Clin Pharmacol *10*, 206-211.

Hausmann, S., Ramirez, A., Schneider, S., Schwer, B., and Shuman, S. (2007). Biochemical and genetic analysis of RNA cap guanine-N2 methyltransferases

from Giardia lamblia and Schizosaccharomyces pombe. Nucleic Acids Res *35*, 1411-1420.

He, W., and Parker, R. (2000). Functions of Lsm proteins in mRNA degradation and splicing. Curr Opin Cell Biol *12*, 346-350.

Helston, R.M., Box, J.A., Tang, W., and Baumann, P. (2010). Schizosaccharomyces cryophilus sp. nov., a new species of fission yeast. FEMS Yeast Res *10*, 779-786.

Henderson, E.R., and Blackburn, E.H. (1989). An overhanging 3' terminus is a conserved feature of telomeres. Mol Cell Biol *9*, 345-348.

Huber, J., Cronshagen, U., Kadokura, M., Marshallsay, C., Wada, T., Sekine, M., and Luhrmann, R. (1998). Snurportin1, an m3G-cap-specific nuclear import receptor with a novel domain structure. EMBO J *17*, 4114-4126.

Huffman, K.E., Levene, S.D., Tesmer, V.M., Shay, J.W., and Wright, W.E. (2000). Telomere shortening is proportional to the size of the G-rich telomeric 3'-overhang. J Biol Chem *275*, 19719-19722.

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

Jamonnak, N., Creamer, T.J., Darby, M.M., Schaughency, P., Wheelan, S.J., and Corden, J.L. (2011). Yeast Nrd1, Nab3, and Sen1 transcriptome-wide binding maps suggest multiple roles in post-transcriptional RNA processing. RNA *17*, 2011-2025.

Kambach, C., Walke, S., Young, R., Avis, J.M., de la Fortelle, E., Raker, V.A., Luhrmann, R., Li, J., and Nagai, K. (1999). Crystal structures of two Sm protein complexes and their implications for the assembly of the spliceosomal snRNPs. Cell *96*, 375-387.

Khusial, P., Plaag, R., and Zieve, G.W. (2005). LSm proteins form heptameric rings that bind to RNA via repeating motifs. Trends Biochem Sci *30*, 522-528.

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

Kim, S.H., Beausejour, C., Davalos, A.R., Kaminker, P., Heo, S.J., and Campisi, J. (2004). TIN2 mediates functions of TRF2 at human telomeres. J Biol Chem *279*, 43799-43804.

Klobutcher, L.A., Swanton, M.T., Donini, P., and Prescott, D.M. (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 U S A *78*, 3015-3019.

Kolquist, K.A., Ellisen, L.W., Counter, C.M., Meyerson, M., Tan, L.K., Weinberg, R.A., Haber, D.A., and Gerald, W.L. (1998). Expression of TERT in early premalignant lesions and a subset of cells in normal tissues. Nat Genet *19*, 182-186.

Kufel, J., Allmang, C., Petfalski, E., Beggs, J., and Tollervey, D. (2003). Lsm Proteins are required for normal processing and stability of ribosomal RNAs. J Biol Chem *278*, 2147-2156.

Kufel, J., Allmang, C., Verdone, L., Beggs, J.D., and Tollervey, D. (2002). Lsm proteins are required for normal processing of pre-tRNAs and their efficient association with La-homologous protein Lhp1p. Mol Cell Biol *22*, 5248-5256.

Larrivee, M., LeBel, C., and Wellinger, R.J. (2004). The generation of proper constitutive G-tails on yeast telomeres is dependent on the MRX complex. Genes Dev *18*, 1391-1396.

Lee, J., Mandell, E.K., Rao, T., Wuttke, D.S., and Lundblad, V. (2010). Investigating the role of the Est3 protein in yeast telomere replication. Nucleic Acids Res *38*, 2279-2290.

Lefebvre, S., Burglen, L., Reboullet, S., Clermont, O., Burlet, P., Viollet, L., Benichou, B., Cruaud, C., Millasseau, P., Zeviani, M.*, et al.* (1995). Identification and characterization of a spinal muscular atrophy-determining gene. Cell *80*, 155-165.

Lendvay, T.S., Morris, D.K., Sah, J., Balasubramanian, B., and Lundblad, V. (1996). Senescence mutants of Saccharomyces cerevisiae with a defect in telomere replication identify three additional EST genes. Genetics *144*, 1399-1412.

Leonardi, J., Box, J.A., Bunch, J.T., and Baumann, P. (2008). TER1, the RNA subunit of fission yeast telomerase. Nat Struct Mol Biol *15*, 26-33.

Lerner, M.R., and Steitz, J.A. (1979). Antibodies to small nuclear RNAs complexed with proteins are produced by patients with systemic lupus erythematosus. Proc Natl Acad Sci U S A *76*, 5495-5499.

Leung, A.K., Nagai, K., and Li, J. (2011). Structure of the spliceosomal U4 snRNP core domain and its implication for snRNP biogenesis. Nature *473*, 536-539.

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

Li, S., Makovets, S., Matsuguchi, T., Blethrow, J.D., Shokat, K.M., and Blackburn, E.H. (2009). Cdk1-dependent phosphorylation of Cdc13 coordinates telomere elongation during cell-cycle progression. Cell *136*, 50-61.

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

Lingner, J., and Cech, T.R. (1996). Purification of telomerase from Euplotes aediculatus: requirement of a primer 3' overhang. Proc Natl Acad Sci U S A *93*, 10712-10717.

Lingner, J., Cech, T.R., Hughes, T.R., and Lundblad, V. (1997a). Three Ever Shorter Telomere (EST) genes are dispensable for in vitro yeast telomerase activity. Proc Natl Acad Sci U S A *94*, 11190-11195.

Lingner, J., Hughes, T.R., Shevchenko, A., Mann, M., Lundblad, V., and Cech, T.R. (1997b). Reverse transcriptase motifs in the catalytic subunit of telomerase. Science *276*, 561-567.

Liu, D., Safari, A., O'Connor, M.S., Chan, D.W., Laegeler, A., Qin, J., and Songyang, Z. (2004a). PTOP interacts with POT1 and regulates its localization to telomeres. Nat Cell Biol *6*, 673-680.

Liu, L., Saldanha, S.N., Pate, M.S., Andrews, L.G., and Tollefsbol, T.O. (2004b). Epigenetic regulation of human telomerase reverse transcriptase promoter activity during cellular differentiation. Genes Chromosomes Cancer *41*, 26-37.

Lundblad, V., and Szostak, J.W. (1989). A mutant with a defect in telomere elongation leads to senescence in yeast. Cell *57*, 633-643.

Lydall, D. (2003). Hiding at the ends of yeast chromosomes: telomeres, nucleases and checkpoint pathways. J Cell Sci *116*, 4057-4065.

Mattaj, I.W. (1986). Cap trimethylation of U snRNA is cytoplasmic and dependent on U snRNP protein binding. Cell *46*, 905-911.

Mayes, A.E., Verdone, L., Legrain, P., and Beggs, J.D. (1999). Characterization of Sm-like proteins in yeast and their association with U6 snRNA. EMBO J *18*, 4321-4331.

McClintock, B. (1941). The Stability of Broken Ends of Chromosomes in Zea Mays. Genetics *26*, 234-282.

McCormick-Graham, M., and Romero, D.P. (1995). Ciliate telomerase RNA structural features. Nucleic Acids Res *23*, 1091-1097.

McElligott, R., and Wellinger, R.J. (1997). The terminal DNA structure of mammalian chromosomes. EMBO J *16*, 3705-3714.

Meister, G., and Fischer, U. (2002). Assisted RNP assembly: SMN and PRMT5 complexes cooperate in the formation of spliceosomal UsnRNPs. EMBO J *21*, 5853-5863.

Meyerson, M., Counter, C.M., Eaton, E.N., Ellisen, L.W., Steiner, P., Caddle, S.D., Ziaugra, L., Beijersbergen, R.L., Davidoff, M.J., Liu, Q.*, et al.* (1997). hEST2, the putative human telomerase catalytic subunit gene, is up-regulated in tumor cells and during immortalization. Cell *90*, 785-795.

Mitchell, J.R., Cheng, J., and Collins, K. (1999). A box H/ACA small nucleolar RNA-like domain at the human telomerase RNA 3' end. Mol Cell Biol *19*, 567-576.

Mitchell, J.R., and Collins, K. (2000). Human telomerase activation requires two independent interactions between telomerase RNA and telomerase reverse transcriptase. Mol Cell *6*, 361-371.

Mouaikel, J., Verheggen, C., Bertrand, E., Tazi, J., and Bordonne, R. (2002). Hypermethylation of the cap structure of both yeast snRNAs and snoRNAs requires a conserved methyltransferase that is localized to the nucleolus. Mol Cell *9*, 891-901.

Muller, H.J., and Altenburg, E. (1930). The Frequency of Translocations Produced by X-Rays in Drosophila. Genetics *15*, 283-311.

Nakamura, T.M., Morin, G.B., Chapman, K.B., Weinrich, S.L., Andrews, W.H., Lingner, J., Harley, C.B., and Cech, T.R. (1997). Telomerase catalytic subunit homologs from fission yeast and human. Science *277*, 955-959.

Nika, J., Yang, W., Pavitt, G.D., Hinnebusch, A.G., and Hannig, E.M. (2000). Purification and kinetic analysis of eIF2B from Saccharomyces cerevisiae. J Biol Chem *275*, 26011-26017.

Noel, J.F., Larose, S., Abou Elela, S., and Wellinger, R.J. (2012). Budding yeast telomerase RNA transcription termination is dictated by the Nrd1/Nab3 non-coding RNA termination pathway. Nucleic Acids Res.

Olovnikov, A.M. (1973). A theory of marginotomy. The incomplete copying of template margin in enzymic synthesis of polynucleotides and biological significance of the phenomenon. Journal of theoretical biology *41*, 181-190.

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

Park, J.I., Venteicher, A.S., Hong, J.Y., Choi, J., Jun, S., Shkreli, M., Chang, W., Meng, Z., Cheung, P., Ji, H.*, et al.* (2009). Telomerase modulates Wnt signalling by association with target gene chromatin. Nature *460*, 66-72.

Patel, S.B., and Bellini, M. (2008). The assembly of a spliceosomal small nuclear ribonucleoprotein particle. Nucleic Acids Res *36*, 6482-6493.

Pillai, R.S., Grimmler, M., Meister, G., Will, C.L., Luhrmann, R., Fischer, U., and Schumperli, D. (2003). Unique Sm core structure of U7 snRNPs: assembly by a specialized SMN complex and the role of a new component, Lsm11, in histone RNA processing. Genes Dev *17*, 2321-2333.

Plessel, G., Fischer, U., and Luhrmann, R. (1994). m3G cap hypermethylation of U1 small nuclear ribonucleoprotein (snRNP) in vitro: evidence that the U1 small nuclear RNA-(guanosine-N2)-methyltransferase is a non-snRNP

cytoplasmic protein that requires a binding site on the Sm core domain. Mol Cell Biol *14*, 4160-4172.

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

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

Raker, V.A., Plessel, G., and Luhrmann, R. (1996). The snRNP core assembly pathway: identification of stable core protein heteromeric complexes and an snRNP subcore particle in vitro. EMBO J *15*, 2256-2269.

Sancar, A., Lindsey-Boltz, L.A., Unsal-Kacmaz, K., and Linn, S. (2004). Molecular mechanisms of mammalian DNA repair and the DNA damage checkpoints. Annu Rev Biochem *73*, 39-85.

Sarthy, J., Bae, N.S., Scrafford, J., and Baumann, P. (2009). Human RAP1 inhibits non-homologous end joining at telomeres. EMBO J *28*, 3390-3399.
Schuster, S.C. (2008). Next-generation sequencing transforms today's biology. Nat Methods *5*, 16-18.

Seraphin, B. (1995). Sm and Sm-like proteins belong to a large family: identification of proteins of the U6 as well as the U1, U2, U4 and U5 snRNPs. EMBO J *14*, 2089-2098.

Seto, A.G., Zaug, A.J., Sobel, S.G., Wolin, S.L., and Cech, T.R. (1999). Saccharomyces cerevisiae telomerase is an Sm small nuclear ribonucleoprotein particle. Nature *401*, 177-180.

Sfeir, A., Kosiyatrakul, S.T., Hockemeyer, D., MacRae, S.L., Karlseder, J., Schildkraut, C.L., and de Lange, T. (2009). Mammalian telomeres resemble fragile sites and require TRF1 for efficient replication. Cell *138*, 90-103.

Sfeir, A.J., Chai, W., Shay, J.W., and Wright, W.E. (2005). Telomere-end processing the terminal nucleotides of human chromosomes. Mol Cell *18*, 131-138.

Shevtsov, S.P., and Dundr, M. (2011). Nucleation of nuclear bodies by RNA. Nat Cell Biol *13*, 167-173.

Shibata, H., Reddy, R., Henning, D., Ro-Choi, T.S., and Busch, H. (1974). Low molecular weight nuclear RNA. The 3'-terminal sequence of the U2 RNA. Mol Cell Biochem *4*, 3-19.

Singer, M.S., and Gottschling, D.E. (1994). TLC1: template RNA component of Saccharomyces cerevisiae telomerase. Science *266*, 404-409.

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

Szostak, J.W., and Blackburn, E.H. (1982). Cloning yeast telomeres on linear plasmid vectors. Cell *29*, 245-255.

Tarn, W.Y., and Steitz, J.A. (1996). Highly diverged U4 and U6 small nuclear RNAs required for splicing rare AT-AC introns. Science *273*, 1824-1832.

Tharun, S., He, W., Mayes, A.E., Lennertz, P., Beggs, J.D., and Parker, R. (2000). Yeast Sm-like proteins function in mRNA decapping and decay. Nature *404*, 515-518.

Thomas, J.D., Conrad, R.C., and Blumenthal, T. (1988). The C. elegans trans-spliced leader RNA is bound to Sm and has a trimethylguanosine cap. Cell *54*, 533-539.

Tuzon, C.T., Wu, Y., Chan, A., and Zakian, V.A. (2011). The Saccharomyces cerevisiae telomerase subunit Est3 binds telomeres in a cell cycle- and Est1-dependent manner and interacts directly with Est1 in vitro. PLoS Genet *7*, e1002060.

Urlaub, H., Raker, V.A., Kostka, S., and Luhrmann, R. (2001). Sm protein-Sm site RNA interactions within the inner ring of the spliceosomal snRNP core structure. EMBO J *20*, 187-196.

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

Vega, L.R., Mateyak, M.K., and Zakian, V.A. (2003). Getting to the end: telomerase access in yeast and humans. Nat Rev Mol Cell Biol *4*, 948-959.

Venteicher, A.S., Abreu, E.B., Meng, Z., McCann, K.E., Terns, R.M., Veenstra, T.D., Terns, M.P., and Artandi, S.E. (2009). A human telomerase holoenzyme protein required for Cajal body localization and telomere synthesis. Science *323*, 644-648.

Venteicher, A.S., Meng, Z., Mason, P.J., Veenstra, T.D., and Artandi, S.E. (2008). Identification of ATPases pontin and reptin as telomerase components essential for holoenzyme assembly. Cell *132*, 945-957.

Verdone, L., Galardi, S., Page, D., and Beggs, J.D. (2004). Lsm proteins promote regeneration of pre-mRNA splicing activity. Curr Biol *14*, 1487-1491.
Wahl, M.C., Will, C.L., and Luhrmann, R. (2009). The spliceosome: design principles of a dynamic RNP machine. Cell *136*, 701-718.

Wang, F., 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.

Wang, S., Zhao, Y., Hu, C., and Zhu, J. (2009). Differential repression of human and mouse TERT genes during cell differentiation. Nucleic Acids Res *37*, 2618-2629.

Watson, J.D. (1972). Origin of concatemeric T7 DNA. Nature: New biology *239*, 197-201.

Webb, C.J., and Zakian, V.A. (2008). Identification and characterization of the Schizosaccharomyces pombe TER1 telomerase RNA. Nat Struct Mol Biol *15*, 34-42.

Wilhelm, B.T., Marguerat, S., Watt, S., Schubert, F., Wood, V., Goodhead, I., Penkett, C.J., Rogers, J., and Bahler, J. (2008). Dynamic repertoire of a eukaryotic transcriptome surveyed at single-nucleotide resolution. Nature *453*, 1239-1243.

Wilusz, C.J., and Wilusz, J. (2005). Eukaryotic Lsm proteins: lessons from bacteria. Nat Struct Mol Biol *12*, 1031-1036.

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

Witkin, K.L., Prathapam, R., and Collins, K. (2007). Positive and negative regulation of Tetrahymena telomerase holoenzyme. Mol Cell Biol *27*, 2074-2083.

Wu, K.J., Grandori, C., Amacker, M., Simon-Vermot, N., Polack, A., Lingner, J., and Dalla-Favera, R. (1999). Direct activation of TERT transcription by c-MYC. Nat Genet *21*, 220-224.

Xin, H., Liu, D., Wan, M., Safari, A., Kim, H., Sun, W., O'Connor, M.S., and Songyang, Z. (2007). TPP1 is a homologue of ciliate TEBP-beta and interacts with POT1 to recruit telomerase. Nature *445*, 559-562.

Ye, J.Z., Donigian, J.R., van Overbeek, M., Loayza, D., Luo, Y., Krutchinsky, A.N., Chait, B.T., and de Lange, T. (2004a). TIN2 binds TRF1 and TRF2 simultaneously and stabilizes the TRF2 complex on telomeres. J Biol Chem *279*, 47264-47271.

Ye, J.Z., Hockemeyer, D., Krutchinsky, A.N., Loayza, D., Hooper, S.M., Chait, B.T., and de Lange, T. (2004b). POT1-interacting protein PIP1: a telomere length regulator that recruits POT1 to the TIN2/TRF1 complex. Genes Dev *18*, 1649-1654.

Yong, J., Kasim, M., Bachorik, J.L., Wan, L., and Dreyfuss, G. (2010). Gemin5 delivers snRNA precursors to the SMN complex for snRNP biogenesis. Mol Cell *38*, 551-562.

Yu, G.L., Bradley, J.D., Attardi, L.D., and Blackburn, E.H. (1990). In vivo alteration of telomere sequences and senescence caused by mutated Tetrahymena telomerase RNAs. Nature *344*, 126-132.

Zahler, A.M., and Roth, M.B. (1995). Distinct functions of SR proteins in recruitment of U1 small nuclear ribonucleoprotein to alternative 5' splice sites. Proc Natl Acad Sci U S A *92*, 2642-2646.

Zappulla, D.C., and Cech, T.R. (2004). Yeast telomerase RNA: a flexible scaffold for protein subunits. Proc Natl Acad Sci U S A *101*, 10024-10029.
Zappulla, D.C., Goodrich, K., and Cech, T.R. (2005). A miniature yeast telomerase RNA functions in vivo and reconstitutes activity in vitro. Nat Struct Mol Biol *12*, 1072-1077.

Zhang, D., Abovich, N., and Rosbash, M. (2001). A biochemical function for the Sm complex. Mol Cell *7*, 319-329.

Zhang, M.Q., and Marr, T.G. (1994). Fission yeast gene structure and recognition. Nucleic Acids Res *22*, 1750-1759.

Zhang, Z., Lotti, F., Dittmar, K., Younis, I., Wan, L., Kasim, M., and Dreyfuss, G. (2008). SMN deficiency causes tissue-specific perturbations in the repertoire of snRNAs and widespread defects in splicing. Cell *133*, 585-600.

Appendix 1:

1. Publications

Box, J.A., Bunch, J.T., Tang, W. and Baumann, P. (2008). Spliceosomal cleavage generates the 3' end of telomerase RNA. Nature. 456 (7224), 910-4.

Helston, R.M., Box, J.A., Tang, W. and Baumann, P. (2010). *Schizosaccharomyces cryophilus* sp. nov., a new species of fission yeast. FEMS Yeast Research. 10 (6), 779-86.

Brown, W. R., Liti, G., Rosa, C., James, S., Roberts, I., Robert, V., Jolly, N., Tang, W., Baumann, P., Green, C., Schlegel, C., Young, J., Hirchaud, F., Leek, S., Thomas, G., Blomberg, A. and Warringer, J. (2011). A Geographically Diverse Collection of *Schizosaccharomyces pombe* Isolates Shows Limited Phenotypic Variation but Extensive Karyotypic Diversity. Genes, Genomes, Genetics. 1 (7), 615-26.

Tang, W., Kannan, R., Blanchette, M., and Baumann, P. (2012). Telomerase RNA biogenesis involves sequential binding by Sm and Lsm complexes. Nature. 484 (7393), 260-4

Appendix 2:

2. Meetings Attended

April 2009, Sixth Cold Spring Harbor meeting on Telomeres & Telomerase. Poster presented. Title: The Biogenesis Pathway of Telomerase RNA

March 2010, AACR meeting: The Role of Telomeres and Telomerase in Cancer Research. Talk presented. Title: 3'end processing of telomerase RNA in fission yeast

June 2011, The HHMI Scientific Meeting. Poster presented. Title: Sequential binding of Sm and Lsm complexes during telomerase RNA biogenesis in fission yeast

June 2012, The 17th Annual Meeting of the RNA Society. Title: Telomerase RNA biogenesis involves sequential binding by Sm and Lsm complexes

Appendix 3:

3. Collaborates' Contributions

Ram Kannan contributed to the characterization some of Sm mutants and analyzed telomere length of Myc-tagged strains. Ram did northern blot analysis in Figure 3.11a.

Marco Blanchette wrote the script for analysis of illumina sequencing data and provided valuable advice for the experiments.