BIOGENESIS OF TELOMERASE RNA SUBUNIT IN

SCHIZOSACCHAROMYCES POMBE

Dissertation
submitted to attain the academic degree
"Doctor
of Natural Sciences"
at the Department of Biology
of the Johannes Gutenberg University
Mainz

Diego J. Páez-Moscoso

M.Sc., Johannes Gutenberg University

B.S., Pontificia Universidad Católica del Ecuador

born in Quito, Ecuador

Mainz, 2021

Dean: Prof.

1st Rapporteur:

2nd Rapporteur:

Day of oral exam:

Abstract

Telomerase is a multi-subunit ribonucleoprotein complex responsible to replenish the missing telomeric DNA sequence during each round of DNA replication. It has two constitutive members: the protein catalytic subunit (TERT) and the telomerase RNA (TER). The TERT subunit is a reverse transcriptase dependent on TER which has the template region for DNA synthesis and the scaffold sequence for its assembly. Dysregulation of telomerase causes human disorders like dyskeratosis congenital, aplastic anemia among others; and reactivation is observed in ~85% of cancers. Since the clinical relevance of telomerase, it is important to develop strategies to manipulate its function. Here, I use the eukaryote *Schizosaccharomyces pombe* to understand the biogenesis of telomerase. Previous works have identified the *S. pombe* telomerase RNA (TER1) precursor is bound by the Sm complex. The Sm complex promotes spliceosomal cleavage (splicing reaction albeit skipping the second transesterification step) and trimethyguanosine (TMG) cap formation. TER1 mature form then transitions from Sm to Lsm2-8 (Sm-like) complex, where the latter stabilize the RNA and promotes binding to the catalytic subunit Trt1. I now demonstrate that an efficient sequential binding from Sm to Lsm requires Pof8, a homolog of ciliate telomeres subunit p65/43. Deletion of *pof8* results in decreased Lsm2-8 complex loading onto TER1 mature form which causes an important assembly defect and critically short telomeres. The binding of Pof8 to TER1 is enhanced by at least two other members Bmc1, a methyltransferase, and a Thc1, a previously uncharacterized protein. Deletion of Bmc1 and Thc1 resulted in reduce TER1 levels and shorter telomeres. Overall, the findings in this work provide new insights into the stepwise telomerase biogenesis and suggest a possible regulatory mechanism to coordinate the biogenesis of the 5’ and 3’ of TER1 molecules.

Acknowledgments

I would like to dedicate my dissertation to my parents. Your love, sacrifice, and dedication provided me with a privileged position in the world making this day possible. Although I am on the other side of the “puddle”, part of my heart and mind are always at home with you. I also want to dedicate my dissertation to my lovely daughter that showed me the meaning of unconditional love. Your dad will always be proud of you no matter what. You are already achieving amazing things in your scarce years of life. Just be happy and healthy always.

I am in complete debt to my wife and co-worker. She immensely helps me in the lab and in life. I cannot imagine how I could have made it to this point without her. She was not only my guide in the scientific maze, but she held and holds my hand while we walk together in life. We don’t have a paved road; we make the road as we walk.

I cannot complete a list of the people that directly or indirectly help me during my studies. However, I will give my best shot.

I would like to thank my adviser. He allowed me to experience first-hand how top-tier science is made. Especial thanks for supporting me throughout my time in the lab and when we needed to pack months of work in few weeks (twice). I definitely would not have achieved key parts of my Ph.D. without your support.

I would like to express my gratitude to my collaborators for their contribution to this work.

I would like to express my sincere gratitude to the Stowers Institute for Medical Research, Institute for Molecular Biology, and Johannes Gutenberg University. They provided incredible resources and environment.

To my friends I met during the Ph.D. for sharing their time and advise during the dark and light sections of this adventure.

To past and present lab members for teaching me the dark ways of RNA biology, for shearing their friendship, always positive commentaries and smiles, for helping me during the earlier stages of my Ph.D. studies and also for her dark humor, for their friendship and discussions, to the lab manager for always keep me on track, to making the lab environment more enjoyable. In general, to all the people that in so many different ways help me achieving my goals.

Table of Contents

List of Tables

List of Figures

Chapter I: Introduction

I.1: Introduction to telomeres

During the evolution of life on Earth, the deoxyribonucleic acid (DNA) polymers were selected to store genetic information due to their high stability in all cell-based organisms. These DNA molecules exist in a circular or linear form present in the majority of prokaryotes, and eukaryotes, respectively. The widespread prevalence of linear DNA in the latter group might lay on its ability to maintain heterozygosity and accommodate ever-growing genetic complexity in a highly compact manner [1,2]. However, the cell faces two inherent problems of linear chromosomes. The end-protection problem where the natural ends that resemble double-strand breaks (DSBs) can be recognized by the repair machinery. The second problem is the end-replication problem where the DNA polymerases cannot fully replicate the ends, losing stretches of DNA in each round of replication. The innovation of the highly repetitive nature of the chromosome ends and a specialized reverse transcriptase provide solutions to these problems. In the next section, I will talk more about these two problems.

I.1.1: The end-protection problem

The DNA molecules storing the genetic information are constantly subject to endogenous and exogenous insults which compromise their integrity. For example, radiation, reactive oxygen species, chemicals, and abortive topoisomerase actions can cause changes in the DNA structure and affect the survival of the organism. One such particularly harmful change in DNA structure is the DSBs which disrupts the continuity of the sugar-phosphate backbone in both DNA strands. The DSBs can lead to chromosome rearrangements, disruption of gene function, and cell senescence and death[3]. Diverse cellular pathways have evolved to recognize and repair these DNA breaks the checkpoints activating kinases, ataxia telangiectasia mutated (ATM) and ATM and Rad3-related (ATR) kinases [4], and the repair

pathways, homolog recombination (HR), and classical non-homologous end-joining (c-NHEJ). The kinases ATM and ATR recognize DSBs and single-stranded DNA, respectively, resulting in cell cycle arrest [5]. The c-NHEJ repairs DNA by ligating the two broken ends and is error-prone while the HR pathway repairs DNA based on homology [3]. Together they are the major DSBs pathways to maintain cell viability.

Eukaryotes distinguish the natural DNA ends of linear chromosomes from being recognized by the DNA damage response machinery. Independent observations made by Müller and McClintock in *Drosophila* and *Zea mays*, respectively, in the 1930s-40s showed that DSBs in chromosomes caused by irradiation are efficiently repaired. McClintock further described the "breakage-fusion-bridge" model where misrepaired DSB can result in dicentric chromosomes which through rounds of cell cycle causes deleterious cellular events [6]. However, the natural chromosome ends were resistant to the fusion events [7,8]. Müller coined the term "telomeres" for the natural chromosome termini with special properties that confer protection from DNA damage response and repair [8]. Seminal contribution from Blackburn and Szostak demonstrated the special protective properties of telomeres by stabilizing artificial linear chromosomes in both ciliates and yeast [9,10]. These pioneer works were the foundations of the telomere field.

I.1.3: Telomeric DNA sequence and structure

Telomeres are composed of repetitive DNA sequences, DNA binding protein complexes, and non-coding RNAs [11]. Elizabeth Blackburn sequenced the first telomeric DNA from *Tetrahymena thermophile,* finding simple tandem repeats of GGGGTT composition [9]. Now we know that the telomeric sequences are diverse among species. Some lineages have perfect repeat sequences, such as GGTTAG invertebrates, some plants, and yeasts [12,13]; while others have degenerative repeat sequences. Two prominent examples are budding yeast *Saccharomyces cerevisiae* and fission yeast *Schizosaccharomyces pombe* which have $TG_{1\text{-}3}$

and $G_{2-8}TTACA_{0-1}C_{0-1}$, respectively [14,15]. The number of repeats per telomere is also regulated in a species-specific manner and varies greatly among organisms. The ciliate *Oxytricha* has 36 base-long telomeres while some breeds of mice can reach 150 kilobases in length [16–18]. In *S. pombe,* the telomere length is ~300 bp [15,19]. Although the telomeric repeat does not fit in a consensus sequence or length, most repeats are G-rich with an orientation specificity concerning the end of the chromosome [12,20].

The telomeric DNA in most species has two regions, a double-stranded, and a single-stranded 3' end -an overhang region on the G-rich strand. Each region recruits specific proteins to maintain the chromosome ends. In mammals, six proteins bind telomeres and are collectively referred to as the shelterin: TRF1, TRF2, Rap1, TIN2, TPP1, and POT1 [11]. They not only protect the telomeres from DNA damage responses but also regulate telomere length. In fission yeast, a similar complex is present at telomeres (Figure 1.1). The protein binding to the double-stranded region is Taz1, the homolog of the human TRF1 and TRF2 [21]. Taz1 at the telomeres represses chromosome end-to-end fusion by inhibiting c-NHEJ and negatively regulates telomere elongation [21,22]. Additionally, Taz1 associates with Rif1, the primary replication timing regulator, to maintain telomere replication [23–25]. Similar functions have been reported for TRF2 and TRF1 in mammals where TRF2 deletion result in chromosome end-to-end fusion in the somatic and carcinogenic cell but not in embryonic stem cells [26–28]. TRF1 facilitates the replication fork progression at telomers, and jointly with TRF2 regulates telomere length by mediating telomerase recruitment [29–31]. TRF2 can also promote the invasion of the 3' -overhand into the double-stranded region of telomeres (T-loop) hiding the ends from DDR and DNA damage checkpoint ATM [32,33].

Pot1 (protection of telomeres) is the single-strand telomeric DNA binding protein, in fission yeast and humans, identified through a sequences homology with the single-stranded binder from ciliates protozoa [34] (Figure 1.1). Deletion of Pot1 causes rapid telomere loss and

chromosome circularization mediated by single-strand annealing in a microhomology site on the sub-telomeric region, in fission yeast [34,35]. In humans and mice, Pot1 also prevents the activation of DNA damage checkpoint ATR and inhibits HR at telomeres [36–39].

The double and single-stranded binders are connected through protein-protein interactions, forming a bridge between them. The three proteins Rap1 (homolog of human Rap1), Poz1 (functional homolog of the human TIN2), and Tpz1 (homolog of human TPP1) form this molecular bridge between Taz1 and Pot1. The bridge is formed by Rap1 interacting with Taz1 and Tpz1 which the last ones connect to Pot1 [23,40] (Figure 1.1b). The proteins forming the molecular bridge not only help to regulate the homeostasis of telomeres, but Rap1 protects telomeres from end-to-end fusions by NHEJ and telomere clustering during meiosis [21,40–42]. Meanwhile, deletion of *tpz1* causes rapid telomere loss and chromosome circularization similar to *pot1Δ* [40]. Tpz1 also interacts with Ccq1, a protein involves in telomere length regulation and heterochromatin maintenance [43]. Similarly in humans, TRF1-TRF2 interacts with TIN2 connecting the double-stranded binder to the single-stranded binders Pot1-TPP1 [44–47]. Rap1 in humans interacts only TRF2 [11] and together with TRF2 inhibit NHEJ; while TPP1 regulates telomere extension and end protection [48–52]. In summary, by recruiting the shelterin complex to the telomeres, the cells can mask the DNA ends from DNA repair machinery solving the end-protection problem and safeguard the genome [11].

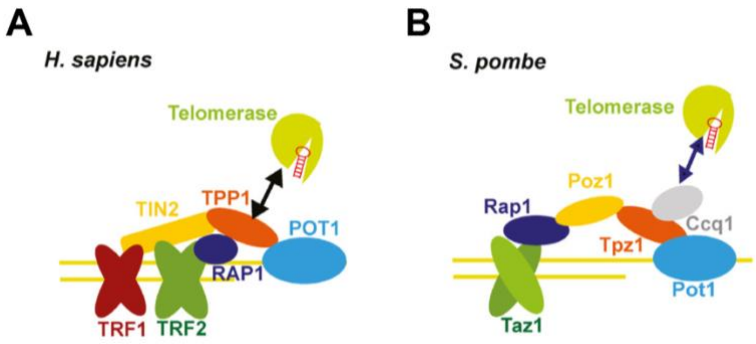

Figure 1.1: Shelterin complex in humans and fission yeast. **a**) The human shelterin complex is composed of six telomeric proteins that maintain telomere homeostasis. **b**) Shelterin complex in fission yeast, *S. pombe*. Modified from [53]

I.1.2: The end-replication problem

The goal of an organism is to propagate and expand; they do so by faithfully transferring the genetic information through cell divisions and DNA replication. The replication of DNA is a semiconservative process in which the sister strand served as a template for the synthesis of a new DNA strand. James Watson and Francis Crick described the double anti-parallel helix structure of the DNA [53]. Based on the proposed structure, they further hypothesized that the DNA can propagate through cell division. A model where each DNA strand, called parental, is the template for the synthesis of the daughter anti-parallel complementary strand [54]. Furthermore, the enzyme catalyzing the synthesis of DNA, DNA polymerase I described by Kornberg and colleagues, can only add nucleotides to the 3' hydroxyl (3' -OH) group [55,56]. This property of DNA polymerases condemns them to a 5' to 3' directionality and the requirement of a primer to initiate the synthesis. These requirements result in a differential synthesis between the two DNA parental strands as they follow the replication fork (Figure 1.2). The lagging strand travels opposite of the replication fork and synthesizes discontinuously. The discontinuous regions named Okazaki fragments are composed of a short oligonucleotide of ribonucleic acid (RNA) and a larger fragment of DNA. The RNA primer is then removed and replaced by the extension of the 3' end of the previous Okazaki fragment. At the ends of the chromosomes, the filling of the RNA primer gap is not possible, leading to loss of DNA sequence in each replication cycle [57]. This attrition of the linear chromosome ends is known as the end-replication problem and was first realized by James Watson and Alexey Olonikov [58,59].

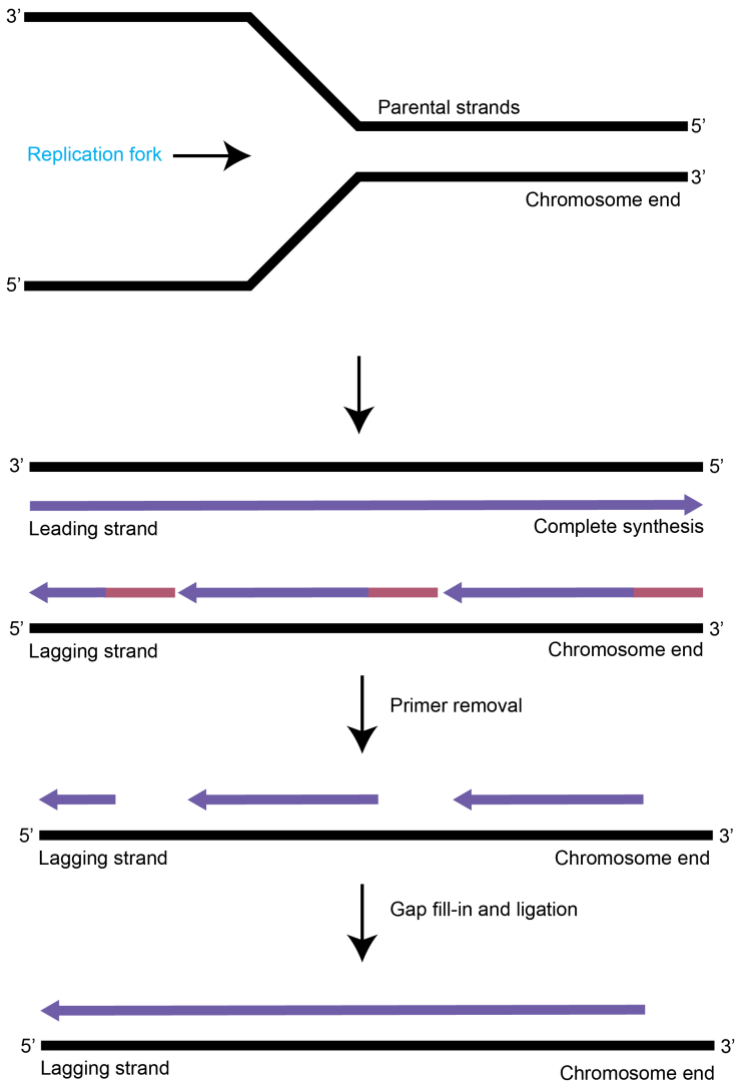

Figure 1.2: The end-replication problem.The semiconservative DNA replication model describes that each of the parental strands is used as a template and requires a polynucleotide primer (mostly RNA) to initiate synthesis by the DNA polymerases. The primer is then removed and the resulting gap is filled by extending the preceding DNA fragment. As the DNA synthesis follows the replication fork, the leading strand is completed continuously to the end of the chromosome, but the very 5'-terminal gap of the lagging strand cannot be filled due to the lack of an upstream fragment, resulting in progressive shortening of DNA at each round of replication.

In 1881, August Weissman, a German scientist, speculated that death is the result of limited tissue renovation [60]. Leonard Hayflick, in the 60s, further described this limited growth potential in primary human cell lines and the positive correlation with the number of passages and cells derived from younger individuals [61,62]. He postulated that the observed phenomenon is an expression of cellular aging or senescence, nowadays known as the

"Hayflick limit". An explanation for the cellular senescence was given by Alexey Olovnikov and Watson who rightly suggested that telogenes (telomeres) confers a buffer zone for chromosome attrition. Loss of this buffer zone would cause chromosomal aberration or senescence/apoptosis. Even more, they proposed that organisms must have mechanisms to overcome telomere shortening. [58,59]. The proposed model was confirmed with several observations made in the early 1990s showing unstable telomere length in somatic human cells [63,64]. Since unicellular eukaryotes organisms or germline cells divide unrestrictive, there must be a mechanism to overcome the end-replication problem.

I.2: Introduction to telomerase

Most eukaryotes resolve the end-replication problem utilizing a specialized enzyme named telomerase. In the middle 1980s, Carol Greider and Elizabeth Blackburn showed a terminal transferase activity present in the *Tetrahymena* native protein extract when incubated with an oligonucleotide primer resembling the 3' end of ciliates and yeast telomeres [65]. The fraction with the activity was isolated and shown to be a ribonucleoprotein complex named telomerase [66]. They further characterized that the RNA component contains the reverse complement sequence of the G-rich repeats of telomeres [67]. Since then, telomerase has been found in a plethora of organisms to be the major pathway for replenishing telomere sequences.

Telomerase is a multi-subunit reverse transcriptase which at its core is composed of the catalytic protein subunit (TERT) and the RNA subunit (TER). The catalytic subunit was first identified in *S. cerevisiae* (EST2) in a genetic screen for phenotypes that resemble TLC1 deletion, the telomerase RNA subunit [68]. Then a further comparison of the EST2 active site with the ciliate *Euplotes aediculatus* confirms the presence of a viral-like RT motif conserved among these proteins [69]. Fission yeast and human telomerase were then cloned using these conserved sites [70,71] making clear that TERT is highly conserved among species [72]. Besides

the central RT catalytic domain, almost all TERTs have N-Terminal (TEN) domain, TR-binding domains (TRBD), and the C-terminal extension (CTE) that the three together form a contiguous ring shape with the active site central to the ring and regulate telomerase function [73,74]. The TEN domain regulates the recruitment to telomeres and telomeric repeat synthesis processivity. The TRBD, as the name suggests, closely interacts with the RNA subunit together with the RT domain. The CTE domains may help either in the recruitment or processivity of telomerase at the telomeres [74,75].

Telomerase RNA subunit was first isolated from *Tetrahymena* where it co-precipitated with telomerase [66]. TER, in contrast with TERT, is highly divergent in sequence and structure between species. For instance, TER length varies from ~145 nucleotides in ciliates to more than 1200 bases in fungi [72]. Despite this variability, there are several common features in all this: 1) a template for repeat synthesis, 2) a pseudoknot (PK) contributing to TERT binding, 3) the boundary element determining the length of each repeat [74]. Additionally, the majority of TRs go through the biogenesis process to ensure the proper function of telomerase. In the next section, I will describe prominent examples of TR biogenesis.

I.2.2: Telomerase RNA biogenesis and assembly in ciliates

In Ciliate *T. thermophila* and *Euplotesaediculatus,* TER is transcribed by RNA polymerase III, and it contains a stretch of uridines at the 3' end [67,76]. After transcription, p65 and p45, a La-related protein, binds to TER and forms part of the telomerase holoenzyme in *T. thermophila* (Figure 1.3) and *Euplotesaediculatus*, respectively [77–79]. p65 is the homolog of human LARP7 and it harbors three domains, La-motif–RRM1 (together known as the La module) and xRRM, an atypical RNA recognition motif (RRM) with extended alpha-helix at the C-terminus [80]. The atypical xRRM domain and the La-module cooperatively confers a high specificity and affinity for the binding to TER. The La-module binds to the terminal

uridine-rich sequence, and the xRRM motif binds to the stem IV, including a GA bulge nucleotides [80,81]. The p65 binding confers TER protection from degradation and conformational changes [81–85]. The conformational changes occurs in the stem IV enhancing the affinity for the catalytic subunit to assemble telomerase [86]. Additional subunits form the telomerase RNP. For instance, p19, p45, p50, and p75 serving as a telomerase adapter complex connecting it to the single-stranded binder of telomeres, Teb1 [87–89]. In ciliate Euplotes, p43 functions similar to its orthologue p65 [90].

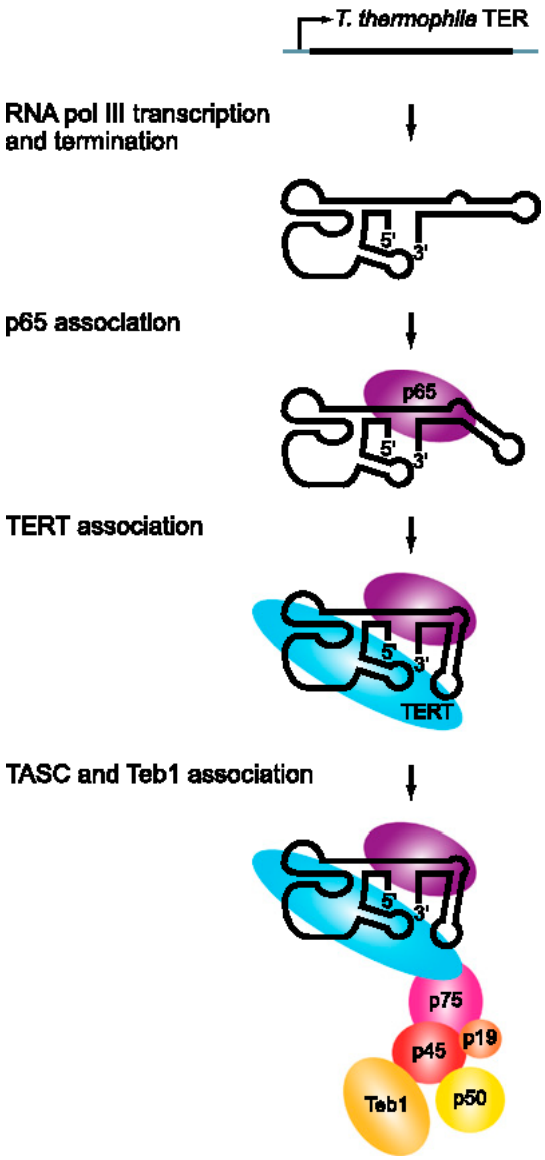

Figure 1.3: Schematic of telomerase RNA (TER) biogenesis in *T. thermophila*.TER is transcribed by RNA polymerase III. The binding of p65 alters confirmation of stem IV that enhances TERT binding. The p65-TER-TERT ternary complex then interacts with a complex of p75, p50, p45, and p19 that recruit the single-stranded telomeric DNA-binding protein Teb1, which provides a stable grip on DNA and promotes telomerase activity. Adapted from [76].

I.2.1: Telomerase RNA biogenesis in humans

The human telomerase RNA subunit, hTR, was identified through PCR-based cyclic selection resulting in a characterization of a 451 nts long ncRNA [91]. hTR transcription was sensitive to RNA polymerase (pol) II inhibition but does not contain a polyA termination signal. hTR also contains a consensus H/ACA three nucleotides upstream of the 451 nts mature 3' end, a motif commonly found on the snoRNAs [92]. Here dyskerin, NOP10, NHP2, and GAR1 binding stabilizes and promotes nuclear retention [92–94] (Figure 1.4). Even the transcription termination is unknown, several studies have shown that hTR molecules extended beyond the 451 nts of the mature form [81,95,96]. An oligo-adenylated form generated by PAPD5, a non-canonical poly-A-polymerase and the catalytic subunit of the TRAMP complex, promotes its degradation [96,97]. While PARN (poly(A)-specific ribonuclease), a counteracting protein, promotes processing by removing the post-transcriptionally added oligo A tails [96,98,99]. PARN function is favored by the binding of the H/ACA complex which helps to unwind a triple helix that attenuates PARN activity [100]. hTR contains another short motif named CAB box which is recognized by TCAB [101,102]. TCAB here localizes hTR to the Cajal bodies and promotes proper folding for assembly with the protein catalytic subunit, TERT [103,104]. The normal RNA pol II 7-methyl-guanosine (m7G) cap of hTR is also further methylated to a 2,2,7-trimethylguanosine (TMG) cap by Tgs1 which negatively regulates telomerase by promoting 5' to 3' degradation in the cytoplasm [102,103]. Recent Cryo-EM structure of human telomerase shows all these trans-acting elements described above forming part of the telomerase holoenzyme plus the surprising addition of histone H2A-H2B bound to as an essential telomerase RNA motif, CR4/5 [105,106].

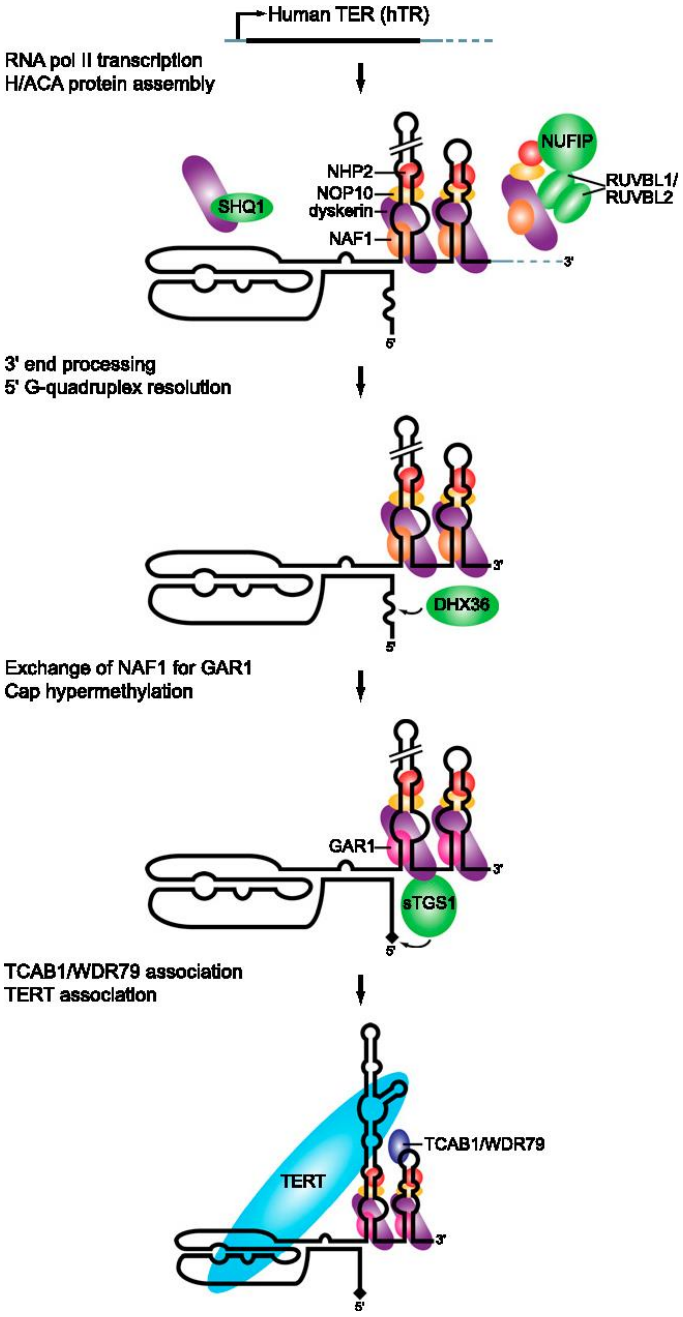

Figure 1.4: Schematic of human telomerase RNA biogenesis.Human telomerase RNA (hTR) is transcribed by RNA polymerase II as a longer polyadenylated form that is inactive and is processed to mature hTR via a PARN (not included in the figure). Dyskerin, NAF1, NHP2, NOP10 bind hTR co-transcriptionally promoting its accumulation *in vivo*. Histone H2A-H2B is not present in the diagram. Adapted from [76].

I.2.3: Telomerase RNA biogenesis in yeast

In budding yeast, the telomerase RNA subunit TLC1 biogenesis has characteristics of the small nucleolar RNAs (snRNAs). TLC1 is transcribed by RNA polymerase II [107,108] and has two major forms in the cell, a polyA+, and a polyA-. TLC1 polyA+ represents only 5-10% in the cell and appears not to be part of the telomerase RNP while the polyA- has been co-precipitated with the active telomerase [107,109]. Even though the relationship between the two forms is unclear [110], the consensus model is that TLC1 is processed by two parallel pathways. The polyA+ form is terminated by the canonical polyA termination pathway [107]. The polyA- transcription termination is dependent on the Nrd1/Nab3/Sen1 small non-coding RNAs pathway. TLC1 harbors the Nrd/Nab binding sites between the mature 3' end and the polyA termination signal, and probably it's the pathway contributing to the most to the TLC1 mature form [111]. During or after transcription termination, the Sm complex binds to the TLC1 Sm binding site located 15 nts upstream of the mature 3' -end. The Sm complex then recruits Tgs1 to convert the m7G cap to TMG cap in the nucleoli [112]. The TMG capped TLC1 is then exported to the cytoplasm where the telomerase is assembled with the Pop1/6/7, Est1, Est2 (protein catalytic subunit), Est3, and Ku70/80 heterodimer and recruited back to the nucleus [113–117] where it can replenish telomere sequences (Figure 1.5).

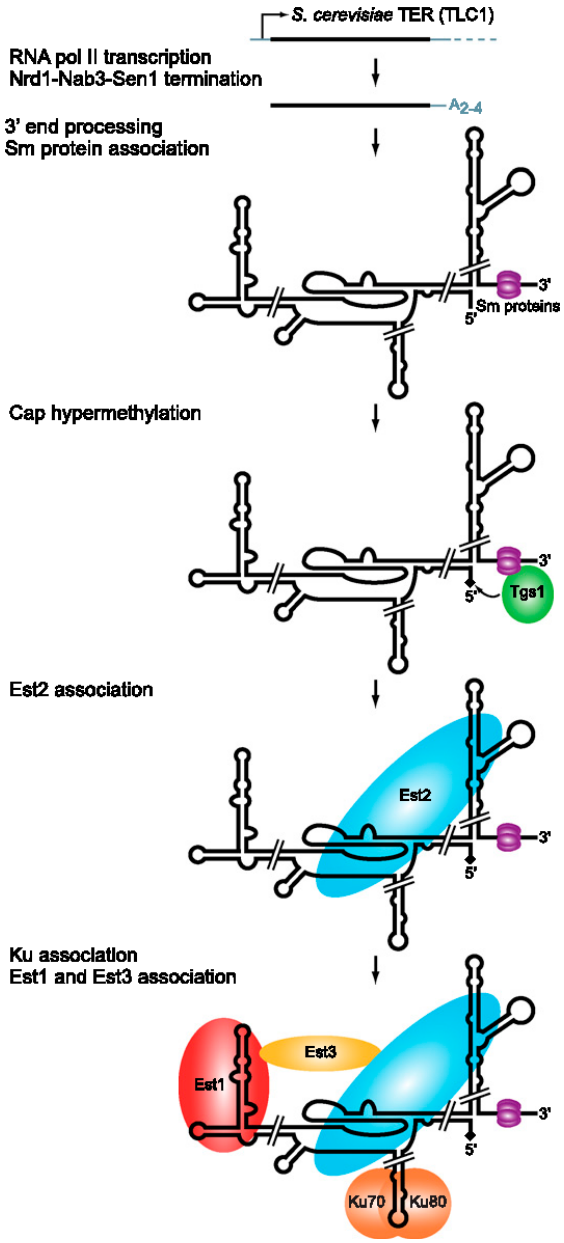

Figure 1.5: Schematic of budding yeast telomerase RNA biogenesis. Budding yeast telomerase RNA (TLC1) is transcribed by RNA polymerase II and terminated by the Nrd1-Nab3-Sen1 pathway. The binding of Sm proteins stabilize the 3'end and recruits Tgs1 which hypermethylated the cap structure. In addition, Ku heterodimer is required for TLC1 stability and nuclear import. Adapted from [76].

TER1, the RNA subunit in *S. pombe*, goes through a stepwise assembly pathway to ensure the quality of the RNA and binding to the catalytic subunit. TER1 is transcribed as a precursor by the RNA polymerase II, and as such, it contains an m7G cap and a poly(A) tail [118–120]. TER1 precursor contains important *cis-regulatory* elements like a 5' splice site (5'SS), a branch site (BS) in its intron, and a 3' splice site (3'SS) [121]. The spliceosome machinery

recognizes the sites and splices the transcript. Curiously, only the first transesterification step occurs discarding the reaction intermediates. The decoupling of the two transesterification steps is due to strong BS-U2 interaction, a long distance to the 3'SS, and a weak polypyrimidine tract in *S. pombe* [121,122]. This reaction named spliceosomal cleavage generates the mature 3' end of TER1. TER1 contains a U-rich binding site resembling the canonical Sm binding site (RAU_5GR) overlapping with the 5'SS [120,121]. It is shown that the Sm complex binds to the precursor in the Sm site promoting splicesome recruitment to TER1. After spliceosomal cleavage, the Sm complex recruits Tgs1 which in turn hypermethylated the m7G cap to TMG cap [119]. TMG cap not only further stabilizes TER1 but also promotes further processing [119]. The Sm ring also interacts with Ctr1-Mtl1 promoting TER1 maturation with unknown mechanism [123]. Finally, the Sm ring is replaced by a paralog complex, the Lsm2-8 complex (Figure 1.6). The transition between Sm and Lsm2-8 complex on TER1 is hypothesized to be a passive mechanism where the Sm complex binding is disfavored by the change in structure/sequence after the spliceosomal cleavage. The free 3' end allows access to a nuclear preformed Lsm2-8 ring complex with a high affinity for terminal polyuridines [119,124]. The 3' end sequencing profile confirmed that the major species pulled-down by the Sm complex contains a terminal U_6G-3' in contrast to the Lsm2-8 complex species containing a $U_{4\text{-}6}$-3'. These data suggest further processing events between the Sm and Lsm2-8 complex transition. For instance, snRNA U6, the canonical Lsm2-8 complex partner, precursor first binds to a La protein with affinity to uridine stretches. Then the exonuclease Mpn1 process U6 3' end leaving a 2'3' cyclic phosphate enhancing the loading of Lsm2-8 complex [125]. It is appealing that there are other factors involved in the transit from the Sm to Lsm2-8 complex. Finally, the loading of the Lsm2-8 complex is crucial to prevent exonucleolytic degradation and recruitment to the catalytic subunit, Trt1.

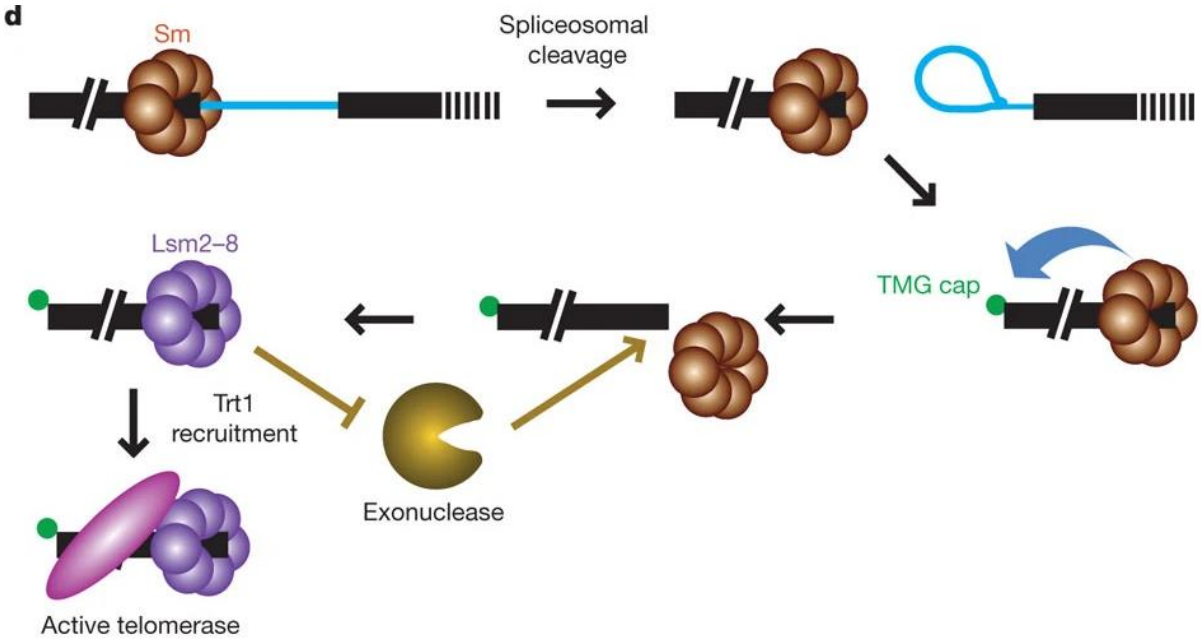

Figure 1.6: Schematic of fission yeast telomerase RNA biogenesis. Fission yeast telomerase RNA (TER1) is transcribed by RNA polymerase II and bound by Sm complex. The binding of Sm proteins recruits the spliceosome for spliceosomal cleavage and Tgs1 which hypermethylates the cap structure. Then the Sm is replaced by the Lsm2-8 complex which protects from 3'-5' degradation and recruits TER1 to the catalytic subunit. Adapted from [119].

I.3: Introduction to the Sm protein family

The Sm proteins are widespread in all domains of life and play a critical role in the life of RNAs. The members in this family were first identified by reacting to autoantibodies from a patient with systemic lupus erythematosus and named in honor of Stephanie Smith, one of the patients [126]. Nowadays, more than 20 members in the family were identified and they all share the Sm fold domain, a domain with an α-helix, and five anti-parallel β-strands [127,128]. The Sm family can be further divided into two sub-families the Sm and Sm-like (Lsm) in eukaryotes (Figure 1.7). Recognized first, the Sm sub-family contains the founder members that assembled into a hetero-heptameric ring associated with small nucleolar RNAs (snRNAs) in eukaryotes [129]. The Lsm sub-family, discovered by homology searches, also forms multimeric rings with a variety of functions depending on the subunit composition. For

instance, the Lsm2-8 complex associates with U6 snRNA and together with the Sm complex form the core subunits of the spliceosome [129]. Their characteristics and functions will be discussed further in the next sections.

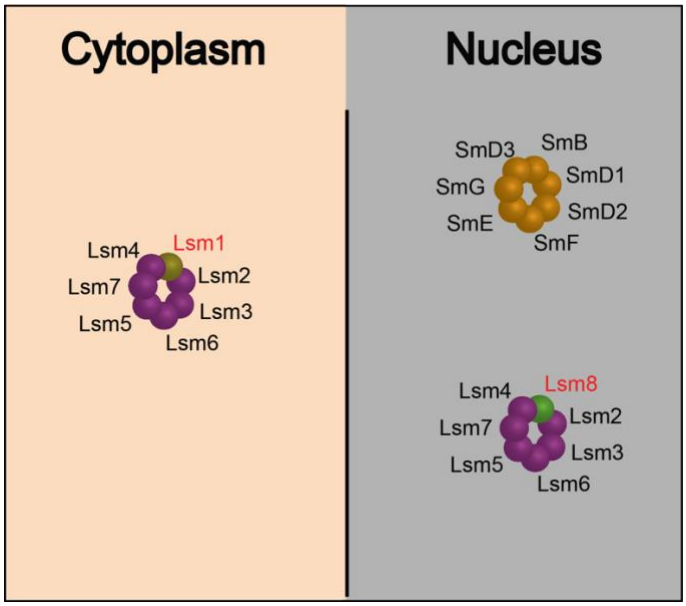

Figure 1.7: The overall composition of the Sm and Lsm complexes.The Sm and Lsm complex form a ring with their seven members. Sm and Lsm2-8 complex are mostly nuclear while the Lsm1-7 complex is located in the cytoplasm.

I.3.1: Characteristics and functions of Sm sub-family

The Sm sub-family has seven members SmB1, Sm D1, SmD2, SmD3, SmE, SmF, and SmG which are mostly known to be part of the spliceosome. These members were first identified from Hela cell nuclear protein extracts using the anti-serum from the patient Stephanie Smith and found to share a common Sm fold domain. They form a complex that associate with a set of U snRNAs, U1, U2, U4, and U5, forming part of the spliceosome and the U11, U12, and U4atac, forming part of the minor spliceosome [126,130]. Additionally, Sm complex associates with the U7 snRNA, functioning on the 3' end formation of histone mRNAs, and with the telomerase RNA subunits of budding, fission yeasts, and humans [103,121,131,132]. Early *in vitro* studies using HeLa native protein extracts and recombinant proteins identified that the Sm complex exists as three sub-complexes in suspension, D1/D2,

E/F/G, and D3/B, which self-assemble to a heterotetrameric ring onto the Sm binding site, a U rich motif, of their cognates RNAs [133–135]. While *in vivo* the snRNP assembly is a multistep process requiring the Survival of Motor Neurons (SMN) complex which brings together the three Sm sub-complexes and the newly cytoplasmic exported U snRNAs [136]. Here the Sm complex protects the RNAs from nucleases and promotes further RNA processing steps. One of the key processing steps is the hypermethylation of the m7G cap to TMG cap by the trimethylguanosine synthase 1 (Tgs1). [137,138]. The TMG cap together with the Sm ring serves as a nuclear import signal for the U snRNAs in humans [139]

I.3.2: Characteristics and function of Lsm sub-family

The members in the Sm-like (Lsm) protein family are involved in various steps of metabolism of coding and non-coding RNAs. The more than 15 members form several functionally distinct complexes in the eukaryotes [140], the most characterized ones being the Lsm1-7 and Lsm2-8 complexes. The Lsm1-7 interacts with the de-capping machinery regulating mRNA degradation in the cytoplasm while the Lsm2-8 protects non-coding RNAs with terminal uridyl from exonucleolytic degradation in the nucleus. Additionally, the Lsm2-8 complex is a chaperone for their cognate RNAs through their life cycle. One important distinction between Sm and Lsm complexes is that Lsm complexes can exist as a preformed ring in the absence of an RNA [141].

Lsm2-8 is a nuclear heteroheptameric complex initially identified to form part of the spliceosome [142]. The Lsm2-8 complex is composed of seven subunits (Lsm2, Lsm3, Lsm4, Lsm5, Lsm6, Lsm7, and Lsm8) forming a ring-shaped complex with high affinity to single-stranded terminal uridines which is typical of RNA pol III transcripts [124,143]. For instance, U6 is one of the major targets of the Lsm2-8 complex which binds to the UUUUU at the 3’ end of the U6 mature form (Figure 1.7). U6 is transcribed by RNA pol III and terminates at variable U tail length [144]. After transcription termination, the La protein binds to the 3’ end

temporarily stabilizing the RNA [145] till a TUTase adds uridines and an exonuclease, Usb1, trims them to a U_5 tail. Usb1 leaves the 3' end with a 2'3' cyclic phosphate which enhances the affinity of the Lsm2-8 complex. [146–148]. In budding yeast, different from vertebrates and fission yeast, the Lsm8 subunit has an extension in the C-terminus that electrostatic interacts with a 3' phosphate of U6 instead of a cyclic phosphate [146]. The binding of the Lsm2-8 complex protects the U6 from further trimming and also helps in the proper folding of the RNA, contributing to the formation of the U4/U6 di-snRNP and U4/U6.U5 tri-snRNP complexes and the regeneration of the spliceosome [149]. The Lsm2-8 complex also associates with $U6_{atac}$, and in fission yeast with the telomerase RNA subunit and stabilizes and promotes the complex assembly [119].

The Lsm1-7 is a heteroheptameric complex linked to RNA decay in the cytoplasm. Lsm1-7 is composed of subunits Lsm1, Lsm2, Lsm3, Lsm4, Lsm5, Lsm6, and Lsm7 (Figure 1.7). Here, only the Lsm1 was identified not to bind to U6 or have any effect on splicing, showing a distinct function in cell [150,151]. Indeed, the complex localizes in the cytoplasm where it affects the stability of mRNAs [152,153]. Lsm1-7 preferentially associates with oligoadenylated rather than polyadenylated mRNAs [154] and promotes degradation by associating with decapping activator Pat1 through Lsm2 and Lsm3 subunits, and the mRNA decapping enzyme Dcp1/2 [153,155]. Xrn1 can then freely degrade the target RNA from 5' to 3' after the cap removal. Lsm1-7 binding can also prevent 3' to 5' degradation of RNA from the exosome *in vitro*, perhaps regulating the 3' to 5' and 5' to 3' degradation rate [156]. In fission yeast and humans, Lsm1-7 alone binding can promote decapping and degradation of the urydilated (oligoU) RNA while Lsm1-7-Pat1 complex promotes degradation on the oligoA RNAs [157,158]. The current model proposes that the regulation of decapping reaction happens in a particular membrane-less organelle called P-bodies [142,159].

I.4: Introduction to the LARP protein family

The La and La-related proteins (LARPs) form an ancient family conserved in the majority of eukaryotes clades [160]. The family is classified by having the La Module, the combination of the La and canonical RNA Recognition Motif 1 (RRM1) motifs. The founder protein of the LARP family is the Genuine La (also known as La protein) which was first identified as a target of the autoantibodies from systemic Lupus erythematosus and Sjogren's syndrome patients [161]. Nowadays, the LARP family is composed of 5 subfamilies, La-protein, LARPs 1, 4, 6, and 7. Each subfamily contains a unique extra motif in addition to the La Module which provides a specialized function [160]. The Genuine La and LARP7, the phylogenetically closest members, associate with non-coding RNAs and are predominantly nuclear. On the contrary, the LARP1, 4, and 6 show a more cytoplasmic localization and regulate mRNAs. For instance, members of the LARP 1 and 4 commonly associate with a poly(A) binding protein PABP to regulate mRNA stability [162,163]. In the next section, I will expand on the functions of La and LARP7 proteins for their proximity to telomere biology and Lsm proteins.

I.4.1: Characterizations and function of La protein

The Genuine La proteins are ubiquitously present in eukaryotes and transiently bind and protect the 3' end of RNA pol III transcripts from exonucleolytic degradation. La protein was first recognized as a target of the autoantibodies from patients with the autoimmune disease Lupus erythematosus [164]. Biochemical studies using those autoantibodies rapidly recognized that La-protein co-immunoprecipitated a variety of nascent transcripts such as pre-tRNAs, snRNA U6, pre-5S rRNA [145,165–167]. La protein was then shown to bind the 3' end of the precursors that contain poly(U) tails [168,169] through the La module (LaM) in yeast and humans. Structurally, the LaM is composed of two domains, the La motif, and the RRM1 domain, which fold together forming a binding pocket that recognizes the 3'-UUU-OH

sequence (Maraia et al., 2017, Jacks et al., 2003). Additionally, the La protein contains an atypical RRM domain at the C-terminus, except in Ascomycota, which may function cooperatively with the LaM to recognize stem loops near the 3' end [170,171].

I.4.2: Characterizations and function of LARP7 proteins

The LARP7 family is the closest to the Genuine La proteins in structure (LaM and xRRM) and functions, but they are more exclusive on their target RNAs. Two members in the family are the p65 which is involved in telomerase RNA biogenesis and assembly in ciliates and the human LARP7 which forms a stable complex with the 7SK non-coding RNA. Since p65 was introduced in detail the previously (section I.2.2), I will focus on the human LAPR7 here. hLARP7 was identified as part of the 7SK RNP. A glycerol gradient was used to isolate the positive transcription elongation factor, P-TEFb (CDK9, Cyclin T1/2), resulted in two major complexes: a smaller P-TEFb with high kinase activity, and the larger inactive complex containing P-TEFb, a non-coding RNA named 7SK, and other proteins [172,173]. The 7SK RNA is 330 nt long and highly structured RNA [174] which forms a stable RNP complex with LARP7 and methyl phosphate capping enzyme (MePCE) (Figure 1.8) [175,176]. Since 7SK is transcribed by RNA polymerase III, it contains a uridine tail at the 3' end which is bound by LARP7 [175]. The 5' end contains triphosphate guanosine where MePCE adds a methyl group to the γ phosphate [177]. After methylation, MePCE remains bound to the 7SK due to the higher affinity to the product than the uncapped RNA substrate through extensive interaction with the stem-loop (SL) 1 in 7SK [176,178,179]. LARP7 binds to the poly(U) 3' end and stem-loop (SL) 4 with its three RNA binding domains. The LARP7 LaM binds to the UUU-3' OH and the major groove in SL4 [178,180] while the xRRM domain binds stably to SL4 [180]. It appears that the high specificity of LARP7 for its targets resides on the atypical xRRM domain containing an additional β4', a non-canonical RNP1, RNP2 sequences, and an extended helix α3 which partially occluding the RNA binding surface of the RRM [80,81,161,170]. Additionally,

the last C-terminal 21 amino acids required for stable 7SK RNP assembly interact with MePCE also inactivate the methyltransferase activity of MePCE [181,182]. Together, MePCE and LARP7 stabilize the 7SK RNA and form a stable trimeric complex. The ciliate member of the LARP7 subfamily is p65/p43 which is a chaperon of the telomerase RNA that starts the step-wise assembly of the telomerase holoenzyme.

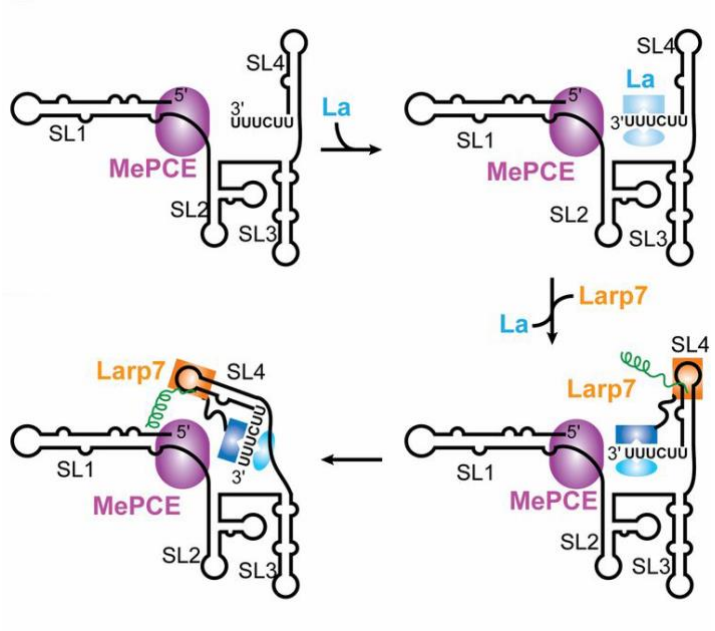

Figure 1.8: Model for the LARP7-7SK-MePCE tertiary complex formation.MePCE binds to the 5' end of nascent 7SK transcript transferring a methyl group. The LARP7 then binds to the mature 3' end assembling with MePCE to form the core 7SK RNP. Adapted from [183]

The main function of the 7SK complex is to regulate the transcription elongation step by sequestering P-TEFb. RNA pol II is stalled shortly after transcription initiation in a large subset of genes. Phosphorylation of the C-terminal domain of pol II and the negative elongation factor (NELF) is required to continue transcription of the stalled genes. The phosphorylation depends on the function of P-TEFb, a heterodimer of CDK9 and Cyclin T1/2. When P-TEFb binds the 7SK complex by HEXIM1/2 recruitment [184], it is inactivated and transcription is stalled [173,185] (Figure 1.9). P-TEFb binding is reversible but it requires

multiple proteins to dissociate such as CTIP2, hnRNP A/Q/R, DDX21 among others [186–188]. 7SK also regulates the transcription of small nuclear RNAs and enhancer RNAs [189–191].

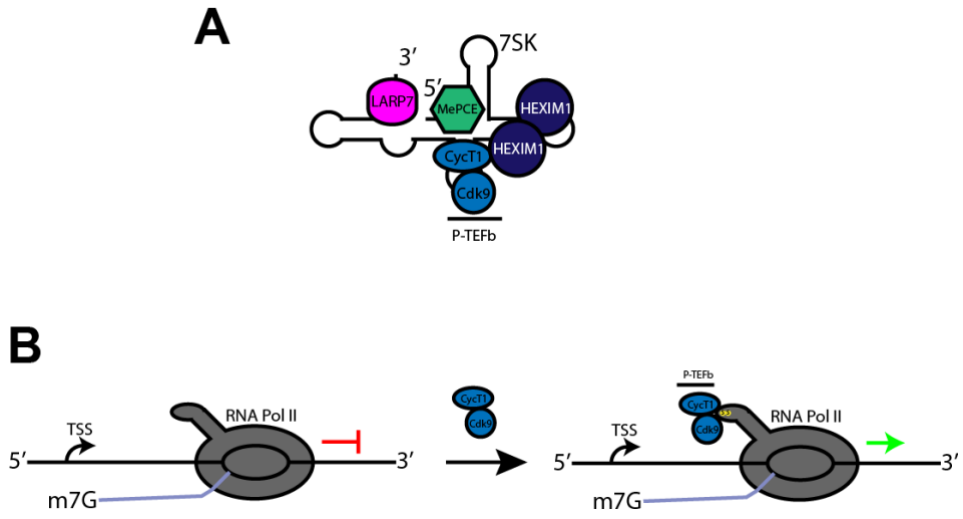

Figure 1.9: 7SK complex regulates transcription elongation.A) 7SK RNP complex together with P-TEFb and HEXIM proteins. B) Paused RNA pol II is released by P-TEFb which phosphorylates pol II CTD promoting transcription elongation.

I.5: Scope of Dissertation

The main goal of my research work as a graduate student was to understand the regulatory mechanism ensuring the stepwise assembly pathway of telomerase RNA biogenesis in fission yeast. Previous work from our lab and others reported the identification of the telomerase RNA subunit in fission yeast, showing the presence of a polyadenylated longer form [118,120]. During TER1 biogenesis, the Sm complex binds to the precursor, which is replaced by the paralog complex Lsm2-8 complex on the TER1 mature form. My project in the lab was to investigate the regulation of the sequential binding between the Sm and Lsm2-8 complexes. I identified and thoroughly characterized an RNA binding protein named Pof8 which previously was thought to be a member of the F-box proteins. Pof8 regulates the assembly of TER1 to the catalytic subunit by controlling the loading of the Lsm2-8 complex. I also showed that Pof8 is a constitutive member of the telomerase holoenzyme in fission

yeast. The results are reported in Chapter III which has been published in Nature Communications [192].

I further investigated the mechanism of how Pof8 loads Lsm2-8 complex by immunoprecipitating Pof8 followed by mass spectrometry. I identified two other members tightly interacting with Pof8, Bmc1, and Thc1. Deletion of each member reduces TER1 levels and causes misregulation of the telomere length. I further identified that Pof8 binding onto TER1 is reduced in the absence of either Bmc1 or Thc1. Finally, I propose that Bmc1 and Thc1 may bind to the TER1 cap structure due to their structural similarity to cap-binding proteins. These results revealed a connection between the 5' and 3, biogenesis of TER1, and an unusual connection between transcription regulatory machinery and telomerase biology in fission yeast. The findings are presented in Chapter IV. In addition, future directions and experiments are discussed in Chapter V.

Chapter II: Pof8 is a La-related protein and a constitutive component of telomerase in fission yeast

II.1: Abstract

Telomerase reverse transcriptase (TERT) and the non-coding telomerase RNA subunit (TR) constitute the core of telomerase. Here, we now report that the putative F-box protein Pof8 is also a constitutive component of active telomerase in fission yeast. Pof8 functions in a hierarchical assembly pathway by promoting the binding of the Lsm2-8 complex to telomerase RNA, which in turn promotes binding of the catalytic subunit. Loss of Pof8 reduces TER1 stability, causes a severe assembly defect, and results in critically short telomeres. Structure profile searches identified similarities between Pof8 and telomerase subunits from ciliated protozoa, making Pof8 next to TERT the most widely conserved telomerase subunits identified to date.

II.2: Introduction

In most eukaryotes, the DNA component of telomeres is composed of short tandem repeat sequences maintained by the reverse transcriptase telomerase. Telomerase is a ribonucleoprotein (RNP) complex in which the RNA subunit (TR, TER1 in *Schizosaccharomyces pombe*) functions as a scaffold for the assembly of telomerase reverse transcriptase (TERT, Trt1 in *S. pombe*) and other protein subunits [76]. The RNA subunit also contains the template region for telomere repeats synthesis. In all species examined, telomerase RNA subunits are transcribed as precursors that then undergo a series of processing events to produce the mature form that is assembled into the active telomerase complex. In *S. pombe*, the mature form of TER1 is ~1213 nucleotides in length[118,120] and ends just upstream of a 5′ splice site [121]. The precursor is about 200 nucleotides longer, containing an intron and a second exon. Interestingly, TER1 maturation involves only the

first step of a splicing reaction. After spliceosomal cleavage at the 5′ splice site, the first exon is released to become the mature form of telomerase. This reaction is favored by RNA elements within TER1 that promote a slow transition between the two steps of splicing resulting in the "discard" of splicing intermediates [122]. A role for spliceosomal cleavage in 3′ end processing of telomerase is conserved among many fungi, but the underlying mechanisms are surprisingly diverse [193,194].

Fungal telomerase RNAs contain an Sm-binding site near the mature 3′ end and require binding of the hetero-heptameric Sm complex for processing and stability [119,121,132,195]. The Sm complex is a member of the Hfq-family of RNA-binding proteins that is conserved in all domains of life [142]. Sm proteins also assemble on spliceosomal small nuclear RNAs (snRNAs), where they are critical for 5′ cap hypermethylation, reimport of the RNPs into the nucleus, and spliceosome function [196]. During *S. pombe* TER1 maturation, Sm proteins promote spliceosomal cleavage and recruit the methyltransferase Tgs1 that generates the 2,2,7-trimethylguanosine (TMG) cap [119]. The Sm complex then dissociates from TER1 and is replaced by the Sm-like complex Lsm2-8, which protects the 3′ end of TER1 from degradation and promotes the association of TER1 with Trt1 to generate the functional enzyme [119].

Biochemical and structural studies of telomerase from ciliated protozoa have provided fundamental insights into telomerase biogenesis and function. However, the extent to which these findings can inform studies in other organisms has remained less clear due to fundamental differences in enzyme composition and biogenesis pathways. For example, the telomerase RNA subunit is transcribed by RNA polymerase III in ciliates, but by RNA pol II in yeasts and metazoans. It has thus been thought that proteins involved in the processing and stabilization of pol III transcripts may only function in telomerase RNA biogenesis in ciliates. This includes members of the Lupus La antigen-related protein (LARP) family [161], which are

components of the telomerase holoenzyme in ciliates and are critical for the assembly, nuclear retention, and activity of telomerase [77–79,90,197].

We now demonstrate that a critical role for La family members in telomerase biogenesis and function is conserved in fission yeast, where telomerase RNA is a pol II transcript. Our results reveal that the Pof8 protein binds to telomerase RNA, functions in hierarchical assembly by promoting Lsm2-8 binding, and forms a constitutive component of the active enzyme. Profile database searches identify Pof8 as a previously unrecognized member of the LARP family displaying striking structural similarities with human and ciliate proteins. Our findings reveal an ancient role for La-related proteins (LARP) in telomerase biogenesis and indicate that evolutionary conservation in holoenzyme composition extends much further than previously thought.

II.3: Methods

II.3.1: Strains and constructs

S. pombe strains used in this study are listed in Table 2.1. The *pof8* deletion was generated by replacing the complete open reading frame with the kanamycin resistance cassette using standard laboratory techniques[198,199].

Table 2.1: *S. pombe* strains used in this study

Strain name	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:: trt1-Cmyc$_9$	Ref[200]
PP407	h +/- ade6-M210/ade6-M216 leu1-32/ leu1-32 ura4-D18/ ura4-D18 his3-D1/ his3-D1 ter1+ /ter1::kanMX6	Ref[121]
PP577	h - ade6-M216 leu1-32 ura4-D18 his3-D1 lsm4::lsm4-myc13-natMX6	Ref[119]
PP578	h - ade6-M216 leu1-32 ura4-D18 his3-D1 lsm5::lsm5-myc13-natMX6	Ref[119]
PP580	h - ade6-M216 leu1-32 ura4-D18 his3-D1 smb1:: smb1-myc13-natMX6	Ref[119]
PP582	h - ade6-M216 leu1-32 ura4-D18 his3-D1 sme1::sme1-myc13-natMX6	Ref[119]
PP585	h - ade6-M216 leu1-32 ura4-D18 his3-D1 lsm8::lsm8-myc13-natMX6	Ref[119]
PP769	h ? ade6-M? leu1-32 ura4? his3? trt1::trt1-Cmyc9 rrp6::kanMX6	This study
PP1721	h ? ade6-M? leu1-32 ura4-D18 his3-D1 lsm4::lsm4-myc13-natMX6 aur1::[pCST159 -ter1]	This study
PP1723	h - ade6-M216 leu1-32 ura4-D18 his3-D1 pof8::kanMX6	This study
PP1724	h - ade6-M216 leu1-32 ura4-D18 his3-D1 lsm4::lsm4-myc13-natMX6 pof8::kanMX6	This study
PP1725	h - ade6-M216 leu1-32 ura4-D18 his3-D1 lsm5::lsm5-myc13-natMX6 pof8::kanMX6	This study
PP1726	h - ade6-M216 leu1-32 ura4-D18 his3-D1 smb1::smb1-myc13-natMX6 pof8::kanMX6	This study
PP1727	h - ade6-M216 leu1-32 ura4-D18 his3-D1 sme1::sme1-myc13-nat, pof8::kanMX6	This study
PP1728	h - ade6-M210 leu1-32 ura4-D18 his3-D1 trt1::trt1-Cmyc9 pof8::kanMX6	This study
PP1729	h ? ade6-M210 leu1-32 ura4-D18 his3-D1 trt1::trt1-Cmyc9 pof8::kanMX6 rrp6::nat	This study
PP1797	h ? ade6-M? leu1-32 ura4-D18 his3-D1 lsm4::lsm4-myc13-nat aur1::[pCST159-ter1] pof8::kanMX6	This study
FP1546	h- ade6-M216 leu1-32 ura4-D18 his3-D1 lsm4::lsm4-myc13-natMX6 pof8::kanMX6 [pDBlet-Pof8]	This study
FP1547	h- ade6-M216 leu1-32 ura4-D18 his3-D1 lsm4:: lsm4-myc13-natMX6 pof8::kanMX6 [pDBlet-3xFLAGPof8]	This study

Knockout fragments contained ~750 base pair (bp) upstream and downstream homology and were generated by fusion PCR using primers listed in Table 2.2. Cells were grown to the late log phase and transformed by lithium acetate method. Transformants were selected on YEA plates plus geneticin disulfate (100 μg ml^{-1}). Epitope tags were introduced following the same strategy. Other strains were generated by crossing and selection of correct genotypes. The 3xFLAG tag was introduced at the N-terminus of Pof8 by fusion PCR in the context of a genomic fragment encompassing sequence from position −386 to +1769 and cloned into pDBlet plasmid[201]. All strains were verified by PCR or western blotting. Plasmids were introduced into *S. pombe* cells by electroporation.

Table 2.2: Oligonucleotides used to generate deletion, fusion and integration constructs

Product description	**Primer #**	**Sequence**
kanMX6 cassette	BLoli6138/ BLoli6139	5′-GCGAAGTAAACGGATCCCCGGGTTAATTAAG-3’/ 5′-GGAAAACATAGAATTCGAGCTCGTTTAAACTG-3’
Pof8 5' UTR	BLoli6136/ BLoli6137	5′-TCTAAAGTTCGTCTTTTGCATAAC-3’/ 5′-CGGGGATCCGTTTACTTCGCTCCTTAAAGTAC-3’
Pof8 3' UTR	BLoli6140/ BLoli6141	5′-GCTCGAATTCTATGTTTTCCTTTCTCTGGTAATAC-3’/ 5′-GCTTTCTTATTTGTAGAGACAATTG-3’
Pof8::kanMX6	BLoli6136/ BLoli6141	5′-TCTAAAGTTCGTCTTTTGCATAAC-3’/ 5′-GCTTTCTTATTTGTAGAGACAATTG-3’
cloning of Pof8 into pDBlet	BLoli6676/ BLoli6677	5′-AAAA GAATTC AACATGGCAACTGCGACCAA-3’/ 5′-CGATAAGCTT CTTCCAATAGCTCGGTTTGT-3’
3xFLAG-Pof8 5' UTR in pDBlet	BLoli6676/ BLoli6683	5′-AAAA GAATTC AACATGGCAACTGCGACCAA-3’/ 5′-ATCGTGATCTTTGTAGTCCATTTTACTTCGCTCCTTAAAG-3’
3xFLAG-Pof8 in pDBlet	BLol6684/ BLoli6677	5′-AAGTAAAATGGACTACAAAGATCACGATGGAGATTATAAAGACCATGATATAGATTATAAGGATGACGATGACAAGTTTGTGCCAAGGCAACTG-3’/ 5′-CGATAAGCTT CTTCCAATAGCTCGGTTTGT-3’
3xFLAG-Pof8 into pDBlet	BLoli6676/ BLoli6677	5′-AAAA GAATTC AACATGGCAACTGCGACCAA-3’/ 5′-CGATAAGCTT CTTCCAATAGCTCGGTTTGT-3’
Sp6-Ter1 5’ (+1 to + 97)	BLoli7098/ BLoli7099	5’-TACGATTTAGGTGACACTATAGATACTCAACGCAACGCC-3’/ 5’-CAAGGAAATTATTTCCTTCAAACTTTCAAATCAATCAC-3’
3’ arm Ter1 (+955 to +1212)-HDV	BLoli7100/ BLoli6540	5’-TGAAGGAAATAATTTCCTTGAACTTGGATTCTTTG-3’/ 5’-TTGGTCCCATTCGCCATGC-3’
Sp6-Ter1-HDV probe	BLoli7098/ BLoli6540	5’-TACGATTTAGGTGACACTATAGATACTCAACGCAACGCC-3’/ 5’-TTGGTCCCATTCGCCATGC-3’

II.3.2: Telomere length analysis and fusion assay

DNA preparation and telomere length analysis were performed based on ref [42]. Cells from 20 ml cultures (~1 × 10^9 cells per ml) were incubated with 2 ml of Z buffer (50 mM sodium citrate, 50 mM sodium phosphate dibasic, and 40 mM EDTA pH 7.8) plus 0.5 mg ml^{-1} Zymolase T100 (US Biological) and 2 mM dithiothreitol (DTT) for 1 h at 37 °C. Sodium dodecyl sulfate (SDS) was then added to a final concentration of 2% (w/v) and incubated for 10 min at 65 °C. Then 5× TE (50 mM Tris-HCl pH 8.0, 5 mM EDTA) was added to a final volume of 10 ml and proteinase K (Sigma-Aldrich, P2308) to a final concentration of 50 μg ml^{-1}. After incubation for 1 h at 50 °C, the samples were precipitated with 3 ml of 5 M potassium acetate for 30 min on ice. The precipitates were removed with two rounds of centrifugation at 3200 × *g* for 10 min at 4 °C. The supernatant was collected and mixed with 1 volume of 100% isopropanol for 1 h on ice followed by centrifugation at 10,500 × *g* for 10 min at 4 °C. Genomic DNA was resuspended in 5× TE with 50 μg ml^{-1} RNAse A. Resuspended DNA was then incubated for 1 h at 37 °C followed by two rounds of extraction with phenol:chloroform:isoamyl alcohol (25:24:1, equilibrated with 5× TE) and one round of chloroform:isoamyl alcohol (24:1, equilibrated with 5× TE). DNA was ethanol precipitated and resuspended in 1× TE. DNA concentrations were determined on a Qubit 3.0 instrument using the dsDNA BR Assay Kit (Life Technologies, Q32853) and 750 ng of each sample was digested with *Eco*RI for 12 h and then loaded onto a 1% agarose gel. The digested DNA was electrophoresed in 0.5× TBE (44.5 mM Tris-borate, 1 mM EDTA at pH 8.3) at 120 V for 6 h. Gels were stained with 1 μg ml^{-1} ethidium bromide and visualized with Typhoon 8600 scanner to confirm digestion of loaded DNA. Gels were then incubated in 0.25 M hydrochloric acid for 10 min followed by 0.5 M sodium hydroxide and 1.5 M sodium chloride buffer for 30 min and 0.5 M Tris-HCl (pH 7.5) and 1.5 M sodium chloride for 30 min at room temperature. DNA was transferred to Amersham Hybond-N+ membrane

(GE Healthcare Life Sciences) via capillary blotting. Transferred DNA was crosslinked to the membrane in a Stratalinker using a 254-nm UV light at 120 mJ cm^{-2}. A probe specific for telomeric sequences was generated by PCR from pTELO using T3 (5′-ATTAACCCTCACTAAAGGGA-3′) and T7 (5′-TAATACGACTCACTATAGGG-3′) oligos. A probe specific for the *rad16* gene was generated by PCR from wild-type genomic DNA using primers XWP9 (5′-ATGGTATTTTTTCGCCATTTACTCG-3′) and XWP10 (5′-TAGGCGGATCGTGAAGTTAA-3′). Both probes were labeled by random hexamer labeling with High Prime (Roche, 11585592001) and [α-^{32}P]-dCTP. Hybridizations were carried out with 10 million counts per minute of probe in Church–Gilbert buffer[202] at 65 °C. Blots were exposed to PhosphorImager screens and visualized with a Typhoon 8600 scanner.

To amplify chromosome end fusions, PCR reactions (25 µl) contained 1× ThermoPol buffer (NEB), 200 µM dNTPs, 0.5 µM of Bloli1256 (GGGTTGCAAAGTATGATTGTGGTAA), and Bloli1353 (TGTTGAATGTCAGAACCAACTGTTGCAT) to amplify fusion junctions, 0.1 µM of Bloli3400 (GCAAAGAAGTTTCCTGGAATAGC) and Bloli3405 (GATGTAATAAAGGGTCGGCAC) to amplify part of the *trt1* gene as loading control, 1.25U Taq polymerase (NEB, M0273) and 1 ng of genomic DNA. Reactions were incubated at 95 °C for 30 s, followed by 32 cycles of 95 °C for 15 s, 55 °C for 30 s, and 68 °C for 3 min with a final extension at 68 °C for 10 min.

II.3.3: Native protein extract and immunoprecipitation

Cultures (2 l) were grown to a density of 0.5–1 × 10^7 cells per ml and harvested by centrifugation for the preparation of cell-free extract[42,118]. Cells were washed three times with ice-cold TMG(300) buffer (10 mM Tris-HCl pH 8.0, 1 mM magnesium chloride, 10% (v/v) glycerol, 300 mM sodium acetate), and resuspended in two packed cell volumes of TMG(300) plus complete EDTA-free protease inhibitor cocktail (Roche), 0.5 mM PMSF,

1 mM EDTA, and 0.1 mM DTT and quick-frozen by dripping the cells suspension in small droplets into liquid nitrogen. Cells were lysed in a 6850 Freezer mill (SPEX SamplePrep) using eight cycles (2 min) at a rate of 10 per second with 2 min cooling time between cycles. Lysates were thawed on ice and one additional packed cell volume of TMG(300) plus supplements was added. Lysates were then cleared by centrifugation twice for 10 min at 6000 × *g* in a Beckman JA-17 rotor and then once for 45 min in a Beckman 70Ti rotor at 36,000 × *g*. All steps were carried out at 4 °C. Protein concentration was determined by Bradford assay and ranged between 6 and 11 mg ml^{-1}.

For RNA IPs, extracts (5.5 mg) were diluted to 5 mg ml^{-1} with TMG(300) buffer plus supplements. An aliquot (100 µl) was frozen as input control. Heparin was added to 1 mg ml^{-1} and Tween-20 to 0.1% (v/v). Magnetic dynabeads protein G (30 mg ml^{-1}; Invitrogen) was coated with anti-c-Myc 9E10 or anti-FLAG M2 (10 µg per 50 µl of bead suspension; Sigma-Aldrich, M4439 and F3165) by incubation for 30 min at room temperature in 200 µl of 1× PBS + 0.1% (v/v) Tween-20. Beads were washed three times with 1 ml of TMG(300). Immunoprecipitation was performed with 60 µl (Pof8) and 120 µl (Sm/Lsm) of bead suspension for 4 h at 4 °C with gentle rotation. Beads were collected using a magnet and an aliquot (100 µl) of supernatant was removed and frozen for further analysis. The beads were then washed five times with 1 ml TMG(300) plus supplements and 0.1% (v/v) Tween-20, once with TMG(200) (as TMG(300) except sodium acetate was at 200 mM) plus supplements and 0.1% (v/v) Tween-20 and once with TMG(50) plus supplements. Finally, beads were resuspended in 120 µl TMG(50) plus supplements and 0.4 U $µl^{-1}$ RNAsin (Promega) and frozen in liquid nitrogen.

II.3.4: RNA preparation

For total RNA extraction, cells (500 ml) were grown to a density of 5×10^6 cells per ml and collected by centrifugation, washed twice with ddH_2O (500 ml), resuspended in 3 ml

ddH_2O and quick-frozen by dripping the cells suspension in small droplets into liquid nitrogen. Cells were lysed in 6850 Freezer mill (SPEX SamplePrep) using seven cycles (2 min) at a rate of 10 per second with 2 min cooling time between cycles. The lysed cells were transferred into 50 ml tubes containing 10 ml phenol:chloroform:isoamyl alcohol (25:24:1, equilibrated with 50 mM sodium acetate, pH 5.2) and 10 ml 50 mM sodium acetate and 1% (w/v) SDS preheated to 65 °C. RNA was extracted four times with 10 ml phenol:chloroform:isoamyl alcohol (25:24:1, equilibrated with 50 mM sodium acetate, pH 5.2) and once with chloroform:isoamyl alcohol (24:1, equilibrated with 50 mM sodium acetate, pH 5.2). Total RNA was ethanol precipitated and resuspended in ddH_2O.

In the context of immunoprecipitation experiments, RNA was isolated from input, supernatant, and beads by incubation with proteinase K (2 μg μl^{-1} in 0.5% (w/v) SDS, 10 mM EDTA pH 8.3, 20 mM Tris-HCl pH 7.5) at 50 °C for 15 min, followed by extraction with phenol:chloroform:isoamyl alcohol and chloroform:isoamyl alcohol. RNA was ethanol precipitated for 4 h at −20 °C and resuspended in ddH_2O. RNA used for RT-PCR was further DNase treated using the RNeasy Mini Kit (Qiagen) following the manufacturer's instructions.

II.3.5: Northern blot analysis

Where indicated, RNaseH cleavage was carried out on 15 μg of DNAse-treated total RNA isolated from *S. pombe*. RNA was combined with 600 pmol of BLoli1043 (5′-AGGCAGAAGACTCACGTACACTGAC-3′) and BLoli1275 (5′-CGGAAACGGAATTCAGCATGT-3′) targeting exon 1 and exon 2, respectively. The mixture was heated to 65 °C in a thermocycler for 5 min and then allowed to slowly cool down at room temperature for 10 min. About 1× RNaseH buffer (NEB) and 5U RNaseH enzyme (NEB, M0297) were added to the mixture and incubated for 30 min at 37 °C. RNaseH-treated samples were ethanol precipitated for 4 h at −20 °C and centrifuged at

2000 × *g* for 20 min at 4 °C. RNA was then resuspended in 1× formamide loading buffer and separated on a 4% (v/v) polyacrylamide (29:1) gel containing 8 M urea and transferred to Biodyne nylon membrane (Pall Corporation) at 400 mA for 1 h in 0.5× TBE buffer. RNA was crosslinked to the membrane using 254-nm UV light at 120 mJ cm^{-2} in Stratalinker (Stratagene). Hybridization with radiolabeled probes (10 million counts per minute) were performed in Church–Gilbert buffer[202] at 60 °C for TER1 probe (nucleotides 536–998, labeled with High Prime (Roche) and [α-^{32}P]-dCTP), and at 42 °C for small nucleolar RNA snRN101 and U6 snRNA (oligonucleotide BLoli1136 (5′-CGCTATTGTATGGGGCCTTTAGATTCTTA-3′) and BLoli4628 (5′-TCTGTATCGTTTCAATTTGACCAAAGTGAT-3′), respectively, labeled with T4 polynucleotide kinase (NEB, M0201) in the presence of [γ-^{32}P]-ATP). Blots were exposed to PhosphorImager screens and analyzed with a Typhoon 8600 scanner.

II.3.6: Western blot analysis

Western blot analysis was performed with native protein extracts, prepared as described above, diluted to 6 µg $µl^{-1}$ and mixed with equal volume of 2× protein sample buffer (2× NuPAGE LDS buffer (Life Technologies), 100 mM DTT, 4% (w/v) SDS). Samples were then incubated for 10 min at 75 °C and 10 µl (30 µg) of samples was loaded onto a 4–12% NuPAGE Bis-Tris gel (Life Technologies, NP0321BOX). Electrophoresis was done in 1× MOPS buffer (Life technologies, NP0001) at 160 V for 60 min. Proteins were transferred to Protran nitrocellulose membranes (Whatman) in western transfer buffer (3.03 g l^{-1} Tris base, 14.4 g l^{-1} glycine, 20% (v/v) methanol) at 100 V for 1 h. Blots were blocked and washed with iBind Flex Western Device (Life Technologies, SLF20002). Lsm and Sm blots were probed with mouse monoclonal anti-c-Myc 9E10 (Sigma-Aldrich, M4439) at 1:5000 dilution and horse-radish peroxidase-conjugated goat anti-mouse IgG (H+L) at 1:5000 (Thermo Scientific, 31430). Trt1 blot was probed with rabbit polyclonal

anti-cMyc A14 (Santa Cruz Biotechnologies, sc-789) at 1:400 dilution and horse-radish peroxidase-conjugated goat anti-rabbit IgG (H+L) at 1:4000 (Thermo Scientific, 31460). Blots were reprobed with mouse anti-α-tubulin (Sigma-Aldrich, T5168).

II.3.7: RT-PCR

DNAse-treated RNA samples were used for RT-PCR reaction as describe[119]. Primers for RT reaction were BLoli1275 (5′-CGGAAACGGAATTCAGCATGT-3′) for precursor and spliced form; PBoli918 (5′-ACAACGGACGAGCTACACTC-3′) for first exon; and BLoli2051 (5′-GACCTTAGCCAGTCCACAGTTA-3′) for U1 as loading control. RNA samples (2.5 μg) were combined with oligos (10 pmol) and dNTP mix (10 nmol) in 13 μl, and samples were heated to 65 °C for 5 min. After cooling, the volume was increased to 20 μl by the addition of 40 U of RNasin (Promega), 5 mM DTT, 1× first-strand buffer, and 200 U of Superscript III reverse transcriptase (Invitrogen). Samples were incubated at 55 °C for 60 min. RNaseH (5 U, NEB, M0297S) was added followed by incubation at 37 °C for 20 min. Aliquots (2 μl) of the RT reactions were used for PCR amplification with Taq polymerase (NEB, M0273) under the following conditions: 5 min at 94 °C followed by 28 cycles of 30 s at 94 °C, 30 s at 57 °C, and 60 s at 72 °C, followed by 10 min at 72 °C. Primers used were BLoli1275 and Bloli1020 (5′-CAAACAATAATGAACGTCCTG-3′) for precursor and spliced form, PBoli918 and BLoli1006 (5′-CATTTAAGTGCTTGTCAGATCACAACG-3′) for first exon, and BLoli2051 and BLoli2101 (5′-ACCTGGCATGAGTTTCTGC-3′) for U1.

II.3.8: Telomerase activity assay

Telomerase was immunoprecipitated on magnetic dynabeads protein G (Invitrogen, 10003D) coated with anti-c-Myc 9E10 (Sigma-Aldrich, M4439) for Trt1 and LSm4, LSm5 or anti-FLAG M2 (Sigma-Aldrich, F3165) for Pof8 as described above. Three amounts of bead suspension (5, 10, and 20 μl) were used for the telomerase activity assay. Negative control samples were incubated in 20 μl of TMG(50) plus 20 ng of RNAse A (Invitrogen) for 10 min

at 30 °C. Beads were incubated in 10 µl of 50 mM Tris-acetate at 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, dTTP, 2 µM [α-^{32}P]-dGTP (500 Ci mmol^{-1}), and 5 µM of oligo PBoli871 (5′-GTTACGGTTACAGGTTACG-3′). Reactions were incubated for 90 min at 30 °C and stopped by the addition of proteinase K (2 µg µl^{-1} in 0.5% (w/v) SDS, 10 mM EDTA pH 8.3, 20 mM Tris-HCl pH 7.5) plus 1000 cpm 100-mer labeled with [γ-^{32}P]-ATP as the loading control at 42 °C for 15 min. Primer extended products were extracted with phenol:chloroform:isoamyl alcohol (25:24:1, equilibrated with 5× TE) and ethanol precipitated for 4 h at −20 °C. Extracted DNA was electrophoresed in 10% (v/v) polyacrylamide (19:1) sequencing gel containing 8 M urea for 1.5 h at 80 W. Gels were dried and exposed to PhosphorImager screens, and analyzed with a Typhoon 8600 scanner.

II.3.9: UV crosslinking and denaturing immunoprecipitation

A Ter1 probe corresponding to the two short arms was generated by fusing an Sp6 promoter sequence with nucleotides +1 to +97 and +955 to +1212 of TER1 and a hepatitis δ virus (HDV) ribozyme sequence to allow for production of a precisely defined 3′ end. Primers used to generate the DNA template are listed in Table 2. The TER1 +1 to +97 fragment with Sp6 promoter sequence was amplified from pJW10[118] with BLoli7098 (containing Sp6 promoter sequence) and BLoli7099 (containing 20 nt overlapping sequence with +955 to +1212 fragment). The TER1 +955 to +1212 fragment with HDV sequence (5′-GGGCGGCATGGTCCCAGCCTCCTCGCTGGCGCCGCCTGGGCAACATGCTTCGGCATGGCGAATGGGACCAA-3′) was PCR amplified from pTER1-i33 with primers Bloli7100 and Bloli6540. The entire probe sequence was then amplified by fusion PCR from the two fragments as template. PCR reactions (50 µl) contained 1× Phusion HF buffer (Life Technologies), 200 µM dNTPs, 0.5 µM of each primer, and 0.02 U µl^{-1} of Phusion Hot Start II DNA Polymerase (Life Technologies, F549). Reaction conditions were: 98 °C for 30 s,

followed by 32 cycles of 98 °C for 10 s, 65 °C for 30 s, and 72 °C for 30 s with a final extension at 72 °C for 10 min. The PCR product was cloned using Zero Blunt PCR Cloning Kit (Life Technologies, K270020) following manufacturer's instructions to give rise to plasmid pDP2 which was sequence verified. The template for in vitro transcription was generated by PCR amplification from plasmid pDP2 with primers BLoli7098 and BLoli6540 using the PCR conditions listed above.

The in vitro transcription reactions contained 1× transcription buffer (Promega), 0.5 mM each of ATP, CTP, GTP, 0.1 mM UTP, 0.66 µM [α-^{32}P]-UTP (3000 Ci mmol^{-1}), 1 µg DNA template, 40 U RNasin (Promega, N2111), and 1.9 U SP6 RNA polymerase (Promega, P1085) in a 10 µl volume. Reactions were incubated at 37 °C for 3 h (time optimized to maximize the amount HDV ribozyme cleaved product), and 2 U of DNAse I (NEB, M0303) was added, followed by incubation for 30 min at 37 °C. Reactions were stopped by the addition of formamide loading dye. The full-length HDV-cleaved TER1 probe product was gel purified on 5% (v/v) polyacrylamide gels containing 8 M urea for 2 h at 15 W. The transcribed and processed RNA is predicted to fold into the same structure as in the context of the full-length RNA using mFOLD default parameters (http://unafold.rna.albany.edu/?q=mfold).

To capture proteins that directly interact with TER1, the RNA probe (2.0 nM) was incubated on ice for 30 min with native protein extracts from strains containing cMyc-tagged Lsm4, FLAG-tagged Pof8, and no tag, respectively. Reactions (100 µl) contained 1.5% (w/v) PEG8000 (NEB), 60 mM potassium phosphate pH 7.0, 1 mM spermidine, 2 mM $MgCl_2$, 0.4 U of RNAsin (Promega, N2111), and 40 µl of 8 µg µl^{-1} native protein extract. After incubation, 50 µl of the reaction was aliquoted into two 25 µl droplets onto parafilm stretched over an aluminum block that was precooled at 4 °C and irradiated in a Stratalinker (Stratagene) using 254 nm UV light at 0.8 J cm^{-2}. After crosslinking, 1% (w/v) of SDS, 1%

(v/v) of Triton X-100, and 100 mM DTT were added to the crosslink and no crosslink reactions, and then heated in boiling water for 2 min. Denatured samples were diluted 10-fold with TMG(300) buffer and immunoprecipitated with magnetic dynabeads protein G (Invitrogen) coated with anti-c-Myc 9E10 or anti-FLAG M2 (Sigma-Aldrich, M4439 and F3165) antibody as described in native protein extract and immunoprecipitation section. After the immunoprecipitation, beads were washed four times with 1 ml TMG(300) and once with 1 ml TMG(50), and treated with proteinase K (2 μg μl^{-1} in 0.5% (w/v) SDS, 10 mM EDTA pH 8.3, 20 mM Tris-HCl pH 7.5) at 42 °C for 15 min. The supernatant was extracted with phenol:chloroform:isoamyl alcohol (25:24:1, equilibrated with 50 mM sodium acetate, pH 5.2) and ethanol precipitated for 4 h at −20 °C. RNA was resolved on a 5% (v/v) polyacrylamide containing 8 M urea for 2 h at 15 W. Gels were dried onto Whatman 3MM Chr blotting paper (GE Healthcare Life Sciences) and exposed to a PhosphorImager screen for analysis with a Typhoon 8600 scanner.

II.3.10: RNA-Seq and RIP-Seq

Ribo-depleted stranded RNA-Seq libraries were constructed for three Pof8 samples and isogenic controls using the TruSeq SBS Kit v4-HS kit (Illumina). The libraries were pooled and sequenced on an Illumina HiSeq 2500 for 100-bp single-end reads on two lanes. Read counts per library ranged from 56 million to 84 million. Reads were aligned to the *S. pombe* ASM294v2 genome from ensembl using the STAR aligner[203] (v.2.5.2b) with the following parameters: -outFilterType BySJout --alignSJDBoverhangMin 5 --alignSJoverhangMin 10 --alignIntronMin 20 --alignIntronMax 2500 --twopassMode Basic. Between 49 and 72 million reads per library passed filtering. A counts table for unambiguous uniquely mapped reads was generated using a custom R script. Coordinates for *snu4* (II: 467,488–467,615)[204] and *snu5* (II: 3,236,356–3,237,312) (GenBank: X16573.1) were manually curated based on alignment of the published sequences to the genome. Differential

gene expression analysis was performed with EdgeR[205] (v.3.14.0) using the likelihood ratio test with the Benjamini–Hochberg false discovery rate correction. Genes with fewer than one count per million in three or more libraries were filtered prior to differential expression analysis. Genes with absolute $\log_2$ fold change >1 with an adjusted *p*-value < 0.05 were considered differentially expressed.

RNA libraries from Lsm8 and control immunoprecipitations were prepared following instructions for Solexa messenger RNA sequencing and each sample was sequenced on a single lane of the Illumina GAIIx sequencer producing 40 bp single-end unstranded reads. The fastq files were aligned to the *S. pombe* reference genome ASM294v2 using the STAR aligner (v 2.5.2b) with the following parameters: --outFilterType BySJout --outFilterMultimapNmax 20 --alignSJoverhangMin 8 --alignSJDBoverhangMin 1 --alignIntronMin 20 --alignIntronMax 1000 --outSAMtype BAM SortedByCoordinate --twopassMode Basic. The corresponding reference annotation gtf file was manually curated for *snu4* using coordinates II:467,488–467,615. A counts table for unambiguous uniquely mapped reads was generated using a custom R script. There were 942,516 read counts for the control IP and 6,604,836 reads for the Lsm8 IP sample. The read counts for each gene were normalized to the median read counts per library. Genes were called enriched if the normalized counts in the IP sample exceeded the corresponding counts in the control by 10 or more and the normalized read count ratio was >2.

II.4: Results

II.4.1: Pof8 is a La-related protein family member

To gain a better understanding of the transition from Sm protein-bound TER1 precursor to the Lsm-bound mature form, we performed immunoprecipitations for each of the Sm and Lsm proteins from strains with myc epitope tags on individual subunits [119]. Precipitates were used to identify associated proteins and RNAs by mass spectrometry and

Illumina sequencing, respectively. Our attention focused on the Pof8 protein as it was reliably precipitated by Lsm2 and Lsm8 and not found in control IPs. Originally reported as a putative F-box protein [206], Pof8 had previously been implicated in telomere maintenance by screening the *S. pombe* gene deletion collection for strains with abnormal telomere length [207]. Using sequence- and profile-based searches, we were unable to independently confirm the F-box domain previously described [208]. However, a homology search readily identified an RNA recognition motif (RRM) near the C-terminus of Pof8 (Figure 2.1a, b). This RRM most closely resembled RRMs in the human LARP family. A subsequent profile sequence search of the full-length Pof8 sequence using HHpred [209] revealed a La motif and an additional RRM (Figure 2.1a, c). Both domains independently identified Pof8 as a LARP, the same family that includes the telomerase subunits p65 from *Tetrahymena thermophila* [79] and p43 from *Euplotes aediculatus* [78].

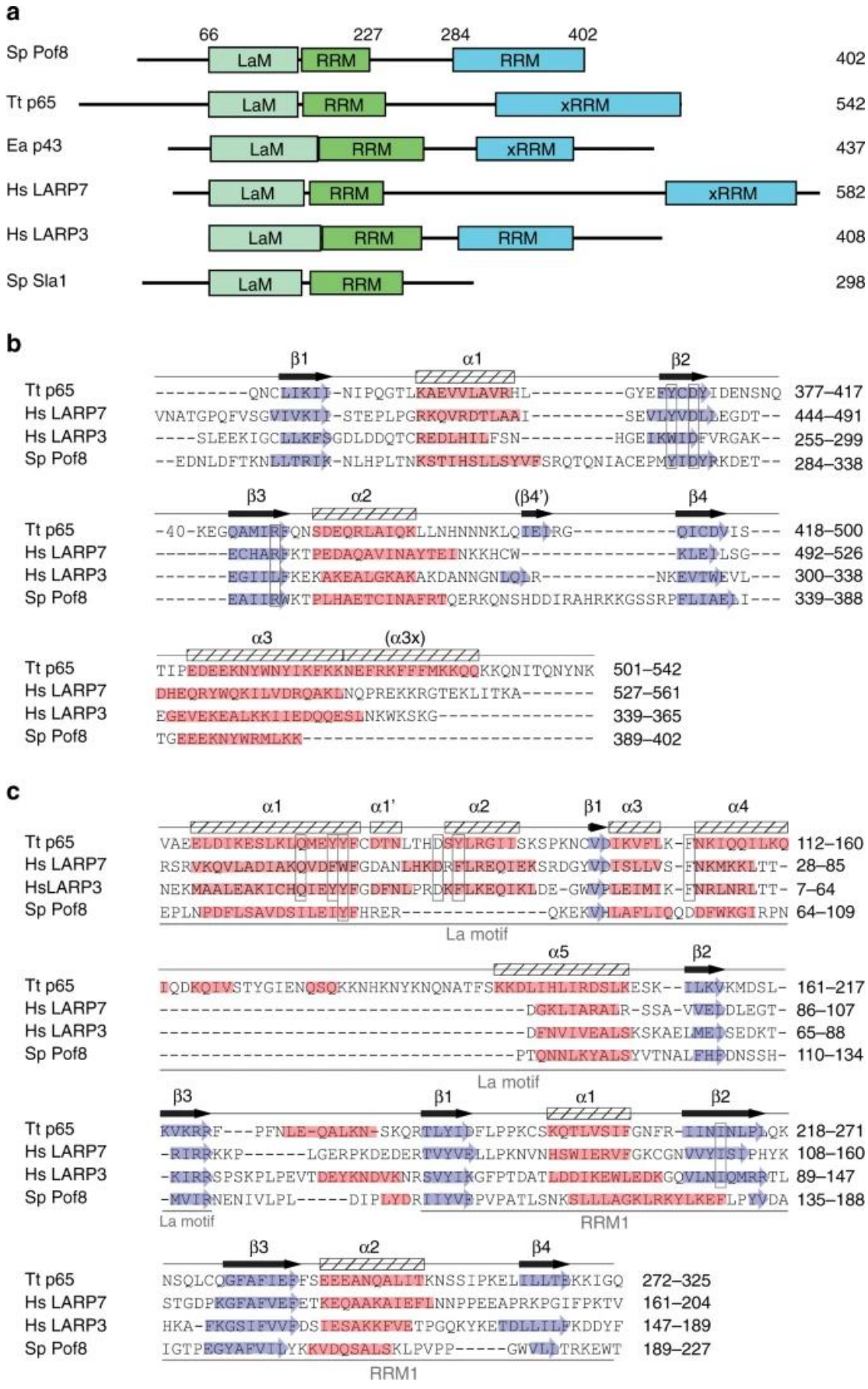
a
66
227
284
402
Sp Pof8
LaM
RRM
RRM
402
Tt p65
LaM
RRM
xRRM
542
Ea p43
LaM
RRM
xRRM
437
Hs LARP7
LaM
RRM
xRRM
582
Hs LARP3
LaM
RRM
RRM
408
Sp Sla1
LaM
RRM
298
b
β1
α1
β2
Tt p65 ---------QNCLIKII-NIPQGTLKAEVVLAVRHL---------GYEFYCDYIDENSNQ 377–417
Hs LARP7 VNATGPQFVSGVIVKII-STEPLPGRKQVRDTLAAI---------SEVLYVDLLEGDT-- 444–491
Hs LARP3 ----SLEEKIGCLLKFSGDLDDQTCREDLHILFSN---------HGEIKWIDFVRGAK-- 255–299
Sp Pof8 --EDNLDFTKNLLTRIK-NLHPLTNKSTIHSLLSYVFSRQTQNIACEPMYIDYRKDET-- 284–338
β3
α2
(β4')
β4
Tt p65 -40-KEGQAMIRFQNSDEQRLAIQKLLNHNNNKLQIEIRG----------QICDVIS--- 418–500
Hs LARP7 -------ECHARFKTPEDAQAVINAYTEINKKHCW---------KLEILSG--- 492–526
Hs LARP3 -------EGIILFKEKAKEALGKAKAKDANNGNLQLR----------NKEVTWEVL--- 300–338
Sp Pof8 -------EAIIRWKTPLHAETCINAFRTQERKQNSHDDIRAHRKKGSSRPFLIAELI--- 339–388
α3
(α3x)
Tt p65 TIPEDEEKNYWNYIKFKKNEFRKFFFMKKQQKKQNITQNYNK 501–542
Hs LARP7 DHEQRYWQKILVDRQAKLNQPREKKRGTEKLITKA------- 527–561
Hs LARP3 EGEVEKEALKKIIEDQQESLNKWKSKG---------- 339–365
Sp Pof8 TGEEEKNYWRMLKK---------- 389–402
c
α1
α1'
α2
β1
α3
α4
Tt p65 VAEELDIKESLKLQMEYYFCDTNLTHDSYLRGIISKSPKNCVDIKVFLK-FNKIQQILKQ 112–160
Hs LARP7 RSRVKQVLADIAKQVDFWFGDANLHKDRFLREQIEKSRDGYVDISLLVS-FNKMKKLTT- 28–85
HsLARP3 NEKMAALEAKICHQIEYYFGDFNLPRDKFLKEQIKLDE-GWVPLEIMIK-FNRLNRLTT- 7–64
Sp Pof8 EPLNPDFLSAVDSILEIYFHRER----------QKEKVHLAFLIQQDDFWKGIRPN 64–109
La motif
α5
β2
Tt p65 IQDKQIVSTYGIENQSQKKNHKNYKNQNATFSKKDLIHLIRDSLKESK--ILKVKMDSL- 161–217
Hs LARP7 ----------DGKLIARALR-SSA-VVELDLEGT- 86–107
Hs LARP3 ----------DFNVIVEALSKSKAELMEISEDKT- 65–88
Sp Pof8 ----------PTQNNLKYALSYVTNALFHPDNSSH- 110–134
La motif
β3
β1
α1
β2
Tt p65 KVKRRF---PFNLE-QALKN-SKQRTLYIDFLPPKCSKQTLVSIFGNFR-IINLNLPLQK 218–271
Hs LARP7 -RIRRKKP------LGERPKDEDERTVYVELLPKNVNHSWIERVFGKCGNVVYISIPHYK 108–160
Hs LARP3 -KIRRSPSKPLPEVTDEYKNDVKNRSVYIKGFPTDATLDDIKEWLEDKGQVLNIQMRRTL 89–147
Sp Pof8 -MVIRNENIVLPL-----DIPLYDRIIYVPPVPATLSNKSLLLAGKLRKYLKEFLPYVDA 135–188
La motif
RRM1
β3
α2
β4
Tt p65 NSQLCQGFAFIEFFSEEEANQALITKNSSIPKELILLTFKKIGQ 272–325
Hs LARP7 STGDPKGFAFVEFETKEQAAKAIEFLNNPPEEAPRKPGIFPKTV 161–204
Hs LARP3 HKA-FKGSIFVVPDSIESAKKFVETPGQKYKETDLLILFKDDYF 147–189
Sp Pof8 IGTPEGYAFVILYKKVDQSALSKLPVPP-----GWVLLTRKEWT 189–227
RRM1

Figure 2.1: Pof8 structural analysis reveals a La motif and two RRM domains. **a)** The domain architecture of *S. pombe* (Sp) Pof8 compared to *T. thermophila (Tt)* telomerase subunit p65, *E. aediculatus* (Ea) telomerase subunit p43, *Homo sapiens* (Hs) LARP7 and LARP3, and the *S. pombe* La protein, Sla1. Numbers above the Pof8 sequence indicate predicted domain boundaries. **b)** Alignment and secondary structure elements of the C-terminal RRM domains based on HHPred predictions; beta strands are highlighted with blue arrows, and alpha helices are highlighted in red. The secondary structure elements for p65 are based on the crystal structure [81] and are labeled above the alignment. Forty amino acids that form an extended loop between β2 and β3 in p65 have been abbreviated -40- to keep the alignment compact. **c)** N-terminal La motif and RRM1 domains. Residues shown to interact with oligo-uridine RNA are boxed and are mostly not conserved in Pof8. The p65 secondary structure elements are labeled above the alignment

II.4.2: Reduced TER1 level and short telomeres in *pof8Δ* cells

The sequence similarity with bona fide telomerase components in ciliates, combined with the interaction of Pof8 with Lsm proteins, lead us to hypothesize that Pof8 may be directly involved in telomerase biogenesis. Examination of TER1 by northern blotting revealed a four to five-fold reduction in RNA levels in *pof8Δ* cells (Figure 2.2a). The reduction in steady-state level predominantly affected the mature form generated by spliceosomal cleavage, whereas the levels of precursor and spliced form were only slightly reduced (Figure 2.2a, b). The levels of Smb1, Sme1, Lsm4, Lsm5 also remained unchanged in the absence of Pof8, and the level of Trt1 protein was only slightly reduced indicating that the reduction in the mature form of TER1 is a direct consequence of loss of Pof8 protein (Figure 2.3).

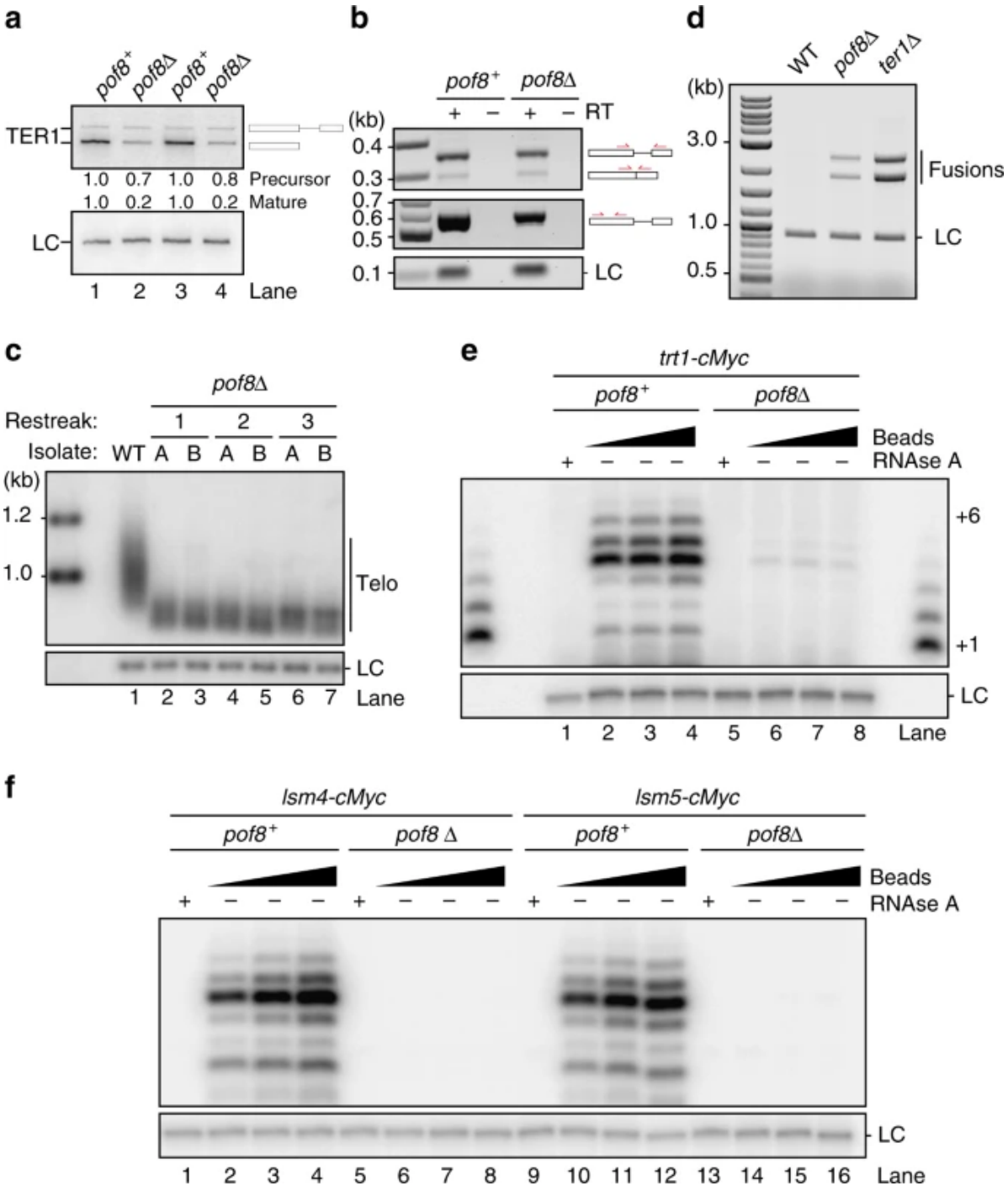

Figure 2.2: Telomerase and telomere defects associated with the deletion of *pof8*. **a)** Northern blot analysis of TER1 from wildtype and *pof8Δ* strains following RNaseH cleavage. Quantification of TER1 signal relative to wildtype and normalized to the loading control (LC, snR101) is shown below the figure. Lanes 1 and 3 represent independent isolates for *pof8*$^{+}$, and lanes 2 and 4 for *pof8Δ*. The schematics on the right represent TER1 precursor and mature forms. **b)** RT-PCR amplification of spliced and unspliced TER1 (top panel), total TER1 (middle panel), and snRNA U1 used as loading control (LC, lower panel). The schematics on the right indicate the position of primers used relative to the structure of precursor, spliced and cleaved TER1 (RT reverse transcriptase). **c)** Southern blot analysis to compare telomere length from wildtype (WT) and *pof8Δ* cells. Two independent isolates of *pof8Δ* (A and B) were restreaked three times and telomere length was analyzed (one

restreak = 20–25 generations). The *rad16*$^{+}$ locus was probed as a loading control (LC). Lane numbers are indicated below the blot. **d)** Detection of telomere–telomere fusions by PCR. **e)** Telomerase activity assay from Trt1-cMyc immunoprecipitations. To establish RNA dependence, RNAse A was added to control samples prior to incubation with DNA primer and nucleotides. A ^{32}P-labeled 100-mer oligonucleotide was used as LC and a ladder of extension products generated by terminal transferase was run on both edges of the gel as size markers; the bands corresponding to 1 and 6 nucleotide addition products are labeled on the right. **f)** As **e)** but from strains harboring Lsm4-cMyc and Lsm5-cMyc in wildtype and *pof8Δ* backgrounds

Consistent with the observations of the genome-wide telomere length screen [207], we found telomere length to be very short in the absence of Pof8 (Figure 2.2c). It is important to note that the chromosome terminal fragments following digest of genomic DNA with *Eco*RI are composed of ~800 bp subtelomeric DNA and a variable number of telomeric repeats. The difference in mobility between wildtype and *pof8Δ* is therefore indicative of critically short telomeres with most of the remaining fragment composed of subtelomeric DNA. Although telomeres were maintained at this short length over successive restreaks at the population level, a fraction of telomeres nevertheless became uncapped and chromosome end fusions were readily detected in *pof8Δ* cells, but not in wild-type cells under the same condition (Figure 2.2d). A *ter1Δ* strain in crisis served as a positive control for chromosome end fusions.

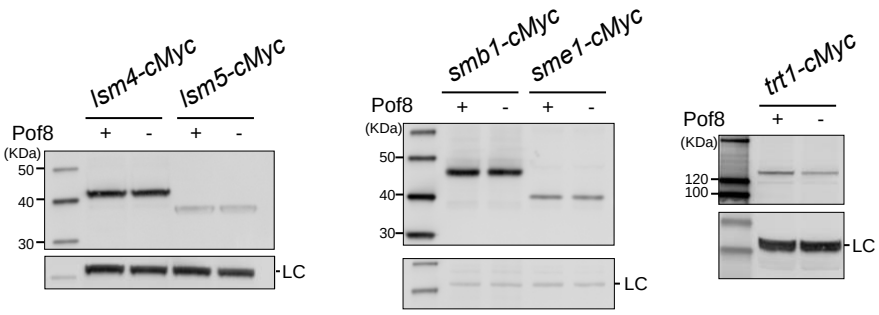

Figure 2.3: Effect of pof8 deletion on protein levels of Sm, Lsm and Trt1.Cell-free extracts from strains with c-Myc epitope tags on the indicated proteins were subjected to Western

analysis. For Lsm4, Lsm5 and Trt1, an antibody against α-tubulin was used as loading control; for Sm proteins, a non-specific band recognized by α-cMyc was used as a loading control.

II.4.3: Pof8 deletion impairs telomerase activity

The pronounced telomere defect observed here was difficult to reconcile with the modest reduction in telomerase RNA level observed by northern blotting. To investigate whether telomerase activity was affected more severely than expected from the four-fold reduction in TER1 RNA, we performed direct telomerase activity assays from Trt1 immunoprecipitants prepared from cell extracts of *pof8*$^{+}$ and *pof8Δ* cells. We observed a 20- to 30-fold reduction in telomerase activity in the absence of Pof8 (Figure 2.2e). Even less telomerase activity was detected when telomerase was immunoprecipitated with Lsm proteins from *pof8Δ* cells (Figure 2.2f).

To separate the effects of *pof8* deletion on TER1 stability from a role for Pof8 in telomerase biogenesis or regulation of activity, we expressed TER1 from the inducible *nmt1* promoter (Figure 2.4a). In the induced state, TER1 levels were 10-fold higher in the *pof8* deletion strains (lanes 5 and 6) compared to wildtype (lane 1). Regardless of the amount of TER1, telomerase activity was dramatically reduced in the absence of Pof8 (Figure 2.4b). Only after contrast adjusting beyond the point of saturation for the signal from *pof8*$^{+}$ samples, weak telomerase activity was detected in the *pof8Δ* extracts (Figure 2.4b, lower panel). Despite higher levels of TER1 RNA in the induced *pof8Δ* cells compared to the uninduced *pof8*$^{+}$ cells, the activity was over 300-fold reduced (compare lanes 2–4 with 14–16). Furthermore, despite a 16-fold higher level of telomerase RNA in induced *pof8Δ* cells vs. uninduced cells, the activity increased by less than two-fold. Consistent with the low activity despite overexpression of TER1, the short telomere phenotype of *pof8Δ* cells was not rescued (Figure 2.4c).

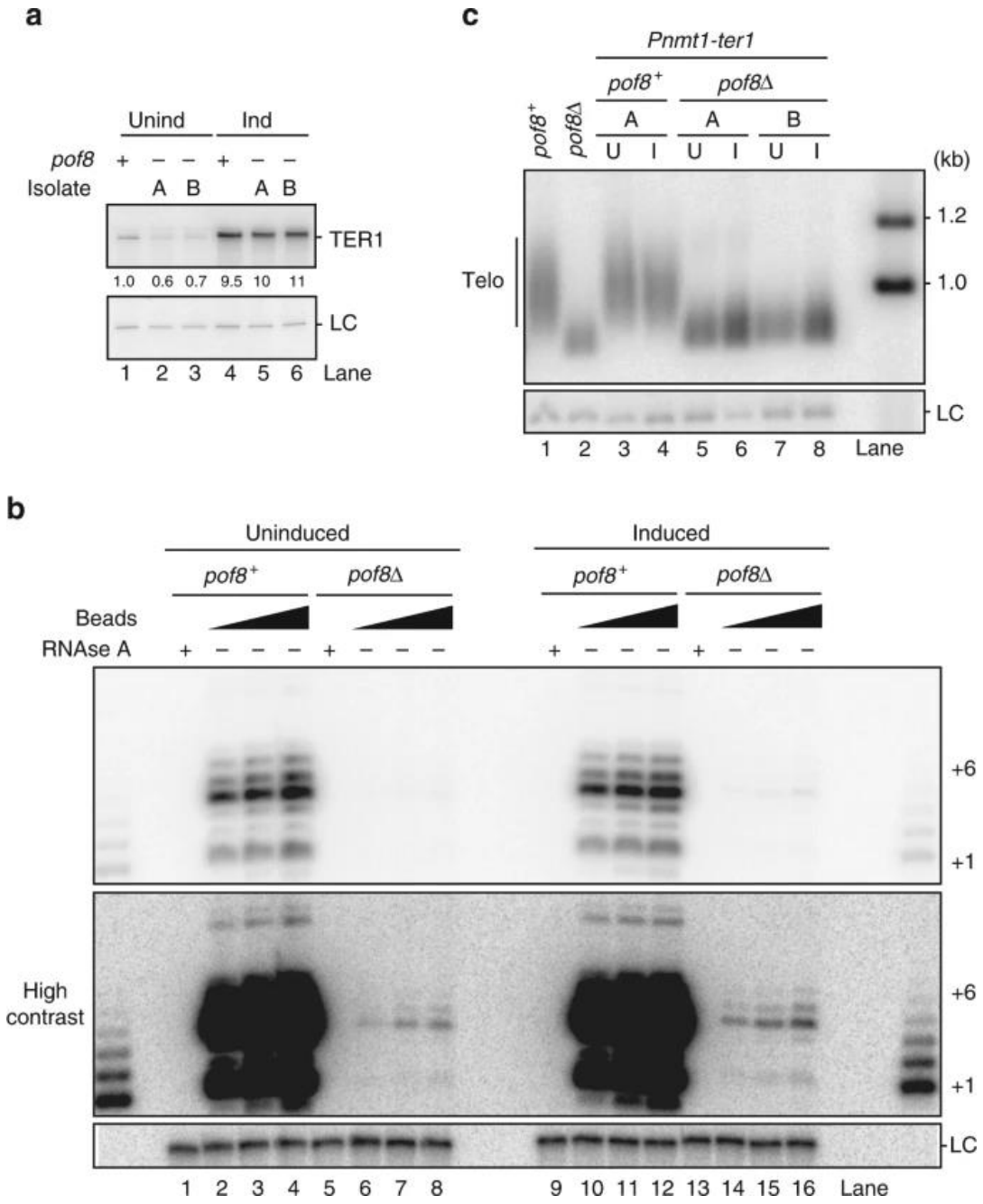

Figure 2.4: TER1 overexpression fails to rescue telomerase defects in *pof8Δ* cells. **a)** Northern blot analysis of total RNA isolated from cells expressing TER1 from a plasmid under the control of uninduced (Unind) and induced (ind) nmt1 promoter in an Lsm4-cMyc background. Two independent isolates (A and B) of *pof8Δ* transformants were used. Uninduced cells were grown in YES (with thiamine) and in EMM (no thiamine) for the induced condition. The lower panel shows snR101 probed as a loading control (LC). **b)** Telomerase activity assay from Lsm4-cMyc immunoprecipitates from *pof8*$^{+}$ and *pof8Δ* cells harboring nmt1-TER1 grown under uninduced and induced conditions. A high contrast version of the top panel is included to visualize the low levels of

telomerase activity from *pof8Δ* cells. Ladder of extension products generated by terminal transferase is flanking the assay as size markers; the bands corresponding to +1 and +6 nucleotide addition products are indicated. A ^{32}P-labeled 100-mer oligonucleotide was used as LC. **c** Southern blot analysis of telomere length of strains described in **a** and controls. U uninduced, I induced. A probe for the *rad16*$^{+}$ locus was used as the LC

TER1 levels also increased following deletion of the RNA exonuclease *rrp6* (Figure 2.5a) [123]. In line with the results obtained by overexpressing TER1, deletion of *rrp6* also failed to rescue telomerase activity (Figure 2.5b). In summary, deletion of *pof8* caused a four-fold reduction in the steady-state level of TER1, but had a far more dramatic effect on telomerase activity. Overexpressing TER1 or interfering with TER1 degradation rescued the RNA level, but failed to rescue telomerase activity and telomere shortening.

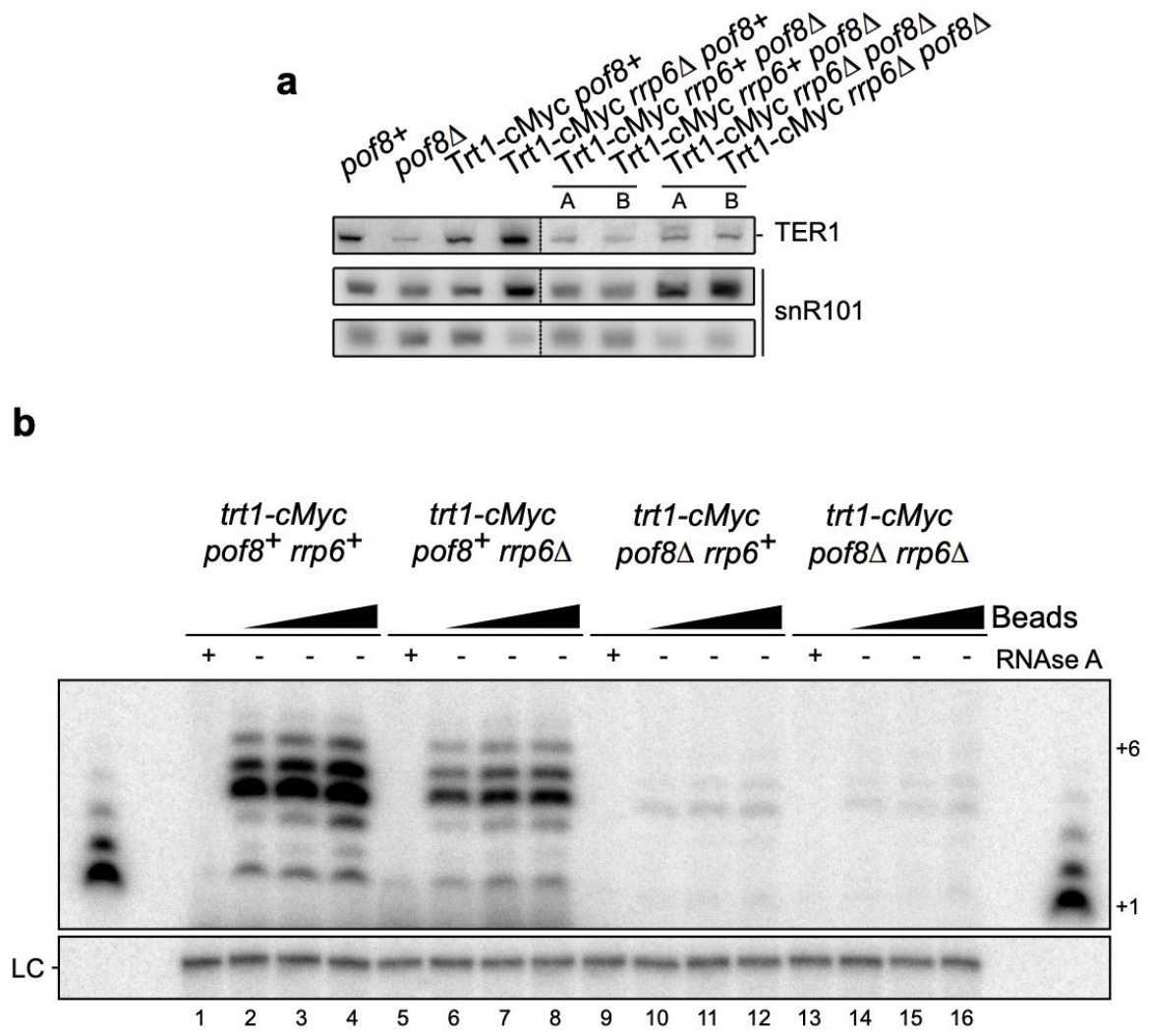

Figure 2.5: Compromising TER1 degradation does not rescue *pof8*Δ effect on telomerase activity.**a**) Northern for TER1 on total RNA samples from strains of the indicated genotypes. Both bands corresponding to snRNA101 are shown revealing a role for Rrp6 in the processing of the longer form. **b**) Telomerase activity assay from pof8Δ and rrp6Δ strains.

II.4.4: Effect of *pof8* deletion on telomerase assembly

The strong effect of *pof8* deletion on telomerase activity indicated that Pof8 may function as an assembly factor for telomerase. As shown previously [119], Lsm2-8 proteins are associated with the majority of mature telomerase RNA, whereas Sm proteins are bound to the precursor and a minor fraction of mature TER1. Deletion of *pof8* results in a 20-fold reduction in the amount of TER1 immunoprecipitated with Lsm4 and Lsm5, and a slight increase in Smb1 and Sme1 association (Figure 2.6a). These results indicate that loading of

Lsm proteins onto TER1 is compromised in the absence of Pof8. To exclude the possibility that the reduction in TER1 level caused by the deletion of *pof8* was responsible for the diminished recovery by immunoprecipitation, we repeated the experiment in the context of overexpressed TER1 (Figure 2.6b). Whereas immunoprecipitation of Lsm4 from a *pof8*$^{+}$ strain depleted over 50% of TER1 from the supernatant (lane 3), no measurable depletion was observed in the *pof8Δ* extract (lane 4, compare input with S/N) and TER1 was barely detectable in the immunoprecipitate (IP, lane 4, lower panel). We conclude that Lsm association with TER1 is compromised in the absence of Pof8. As we have previously shown that Trt1 association with TER1 requires prior binding of Lsm2-8 [119], these experiments place Pof8 upstream of Lsm and Trt1 in the hierarchical assembly of telomerase and explain the dramatic reduction in telomerase activity in the absence of Pof8.

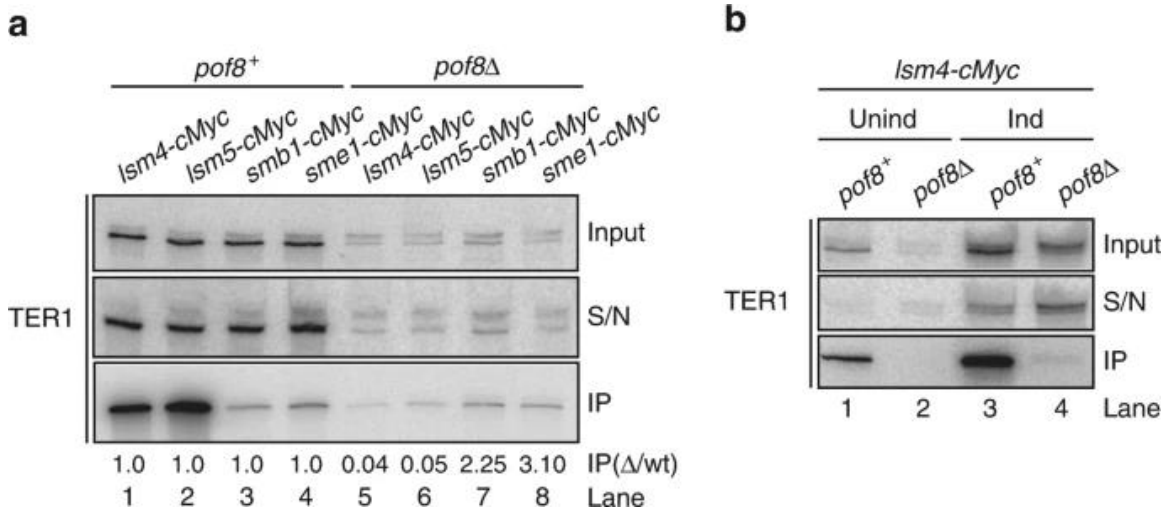

Figure 2.6: Lsm2-8 complex fails to associate with TER1 in Pof8-deficient cells. **a)** Northern blot analysis of TER1 following immunoprecipitation (IP) of cMyc-tagged Lsm and Sm proteins. Input and supernatant (S/N) represent 10% of the IP samples. Upper band corresponds to precursor, lower band to mature form. IP signal was normalized to precursor plus mature form from the input samples and fold enrichment is shown relative to wildtype

below the blots. **b)** Northern blot analysis of TER1 from anti-cMyc IP samples. Strains contain plasmid borne, inducible TER1 in an Lsm4-cMyc background. Input and S/N represent 10% of the IP samples (Unind uninduced, Ind induced).

We next wanted to know whether Pof8 directly and stably associates with TER1. Following introduction of an N-terminal 3xFLAG epitope tag, Pof8 was detected as a single band by western blotting (Figure 2.7a). Telomeres continued to be maintained at near wildtype length indicating that the 3xFLAG tag has little effect on Pof8 function (Figure 2.7b). To test whether Pof8 binds directly to TER1, we incubated a radiolabeled probe corresponding to the short arms of TER1 with *S. pombe* extract containing FLAG-tagged Pof8 and UV irradiated to crosslink RNA–protein interactions. Pof8 was then immunoprecipitated under denaturing conditions to disrupt indirect interactions. Parallel control experiments were carried out with extracts containing no tags and myc epitope-tagged Lsm4, respectively. RNA was isolated from each IP, separated by gel electrophoresis and visualized (Figure 2.7c). TER1 was found to be 2.5-fold enriched in the Lsm4 and Pof8 IPs relative to the untagged control, strongly supporting that Pof8 directly interacts with TER1, as does Lsm4.

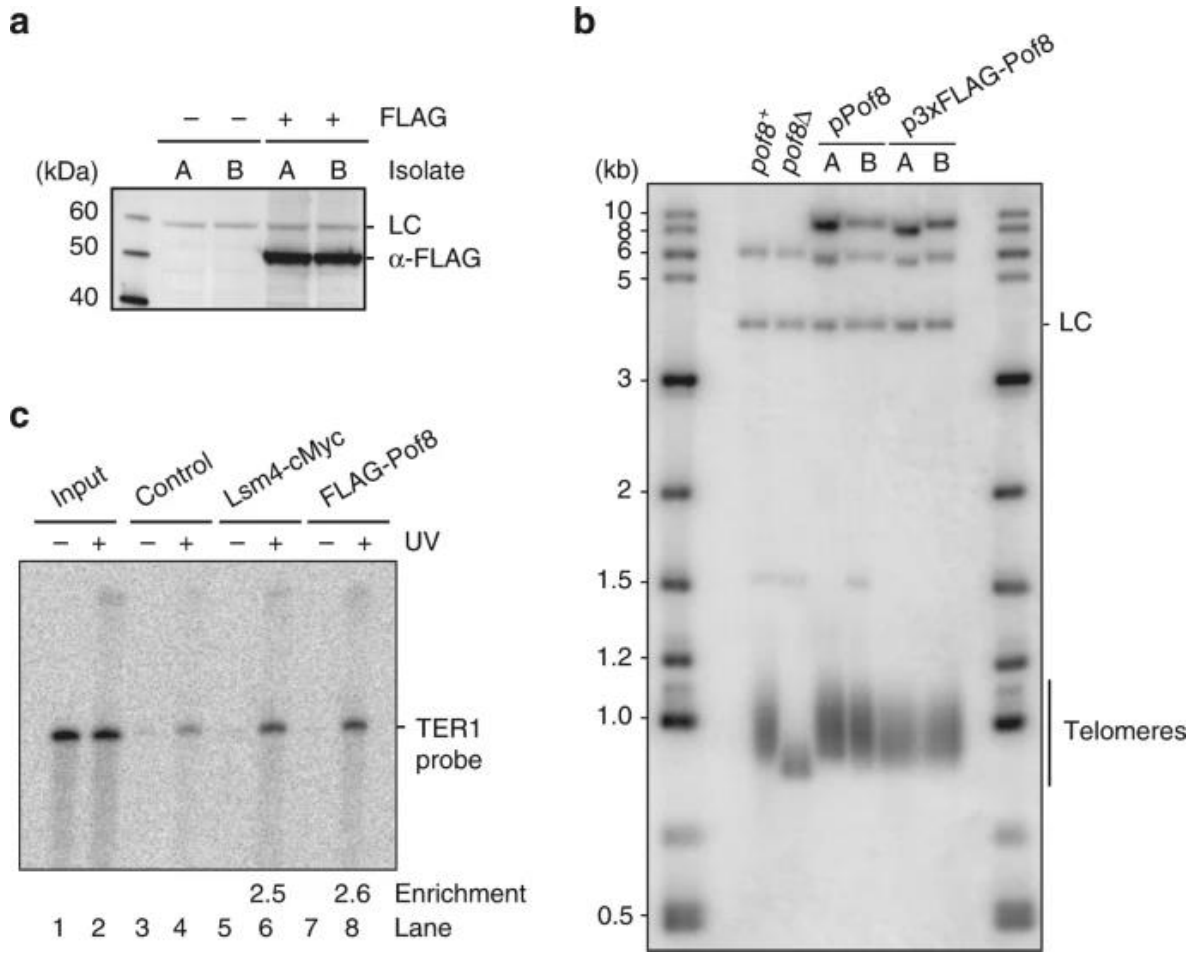

Figure 2.7: Characterization of 3xFLAG-Pof8. **a)** Western blot using α-FLAG antibody in strains containing untagged Pof8 or 3xFLAG-Pof8 with two independent isolates per strain in an Lsm4-cMyc background. A non-specific band recognized by α-FLAG in *S. pombe* extracts even in the absence of an epitope-tagged protein was used as a loading control (LC). **b)** Telomeric Southern blot from strains containing untagged Pof8 or 3xFLAG-Pof8 on a plasmid under the control of its endogenous promoter with two independent isolates per strain in an Lsm4-cMyc background. A *rad16*$^{+}$ probe was used as the LC. **c** UV crosslinking and immunoprecipitation of in vitro transcribed TER1 probe incubated in extracts containing epitope-tagged Lsm4 or Pof8

II.4.5: Pof8 is a subunit of active telomerase

To assess whether Pof8 stably associates with TER1 in the context of active enzyme, we analyzed FLAG-Pof8 immunoprecipitations carried out under native conditions by northern blotting. TER1 was readily detected in immunoprecipitates from tagged, but not from control extracts in which Pof8 was untagged (Figure 2.8a). The snRNA snR101 was not precipitated by Pof8 and thus served as specificity control. Whereas no activity was observed

in control samples, FLAG-Pof8 immunoprecipitates displayed robust telomerase activity (Figure 2.8b). To assess more quantitatively what fraction of telomerase is Pof8-associated, we generated strains containing Lsm4-cMyc in combination with FLAG-tagged or untagged Pof8. Cell-free extracts from these strains were subjected to a first round of immunoprecipitation using anti-FLAG antibody. The supernatant was then incubated with anti-cMyc to precipitate Lsm4-associated TER1. The four immunoprecipitates were then assayed for telomerase activity. As expected, no telomerase activity was precipitated in the FLAG IP from extract containing untagged Pof8 (Figure 2.8c, lanes 2–4). Subsequent IP of Lsm4-cMyc from the supernatant of the first-round IP recovered robust activity (lanes 6–8). In contrast, when Pof8-associated telomerase was first precipitated (lanes 10–12), only 1% of activity was recovered in a subsequent Lsm4 IP (lanes 14–16). Based on these results, nearly 100% of active telomerase are associated with Pof8, making this protein a bona fide component of the active holoenzyme.

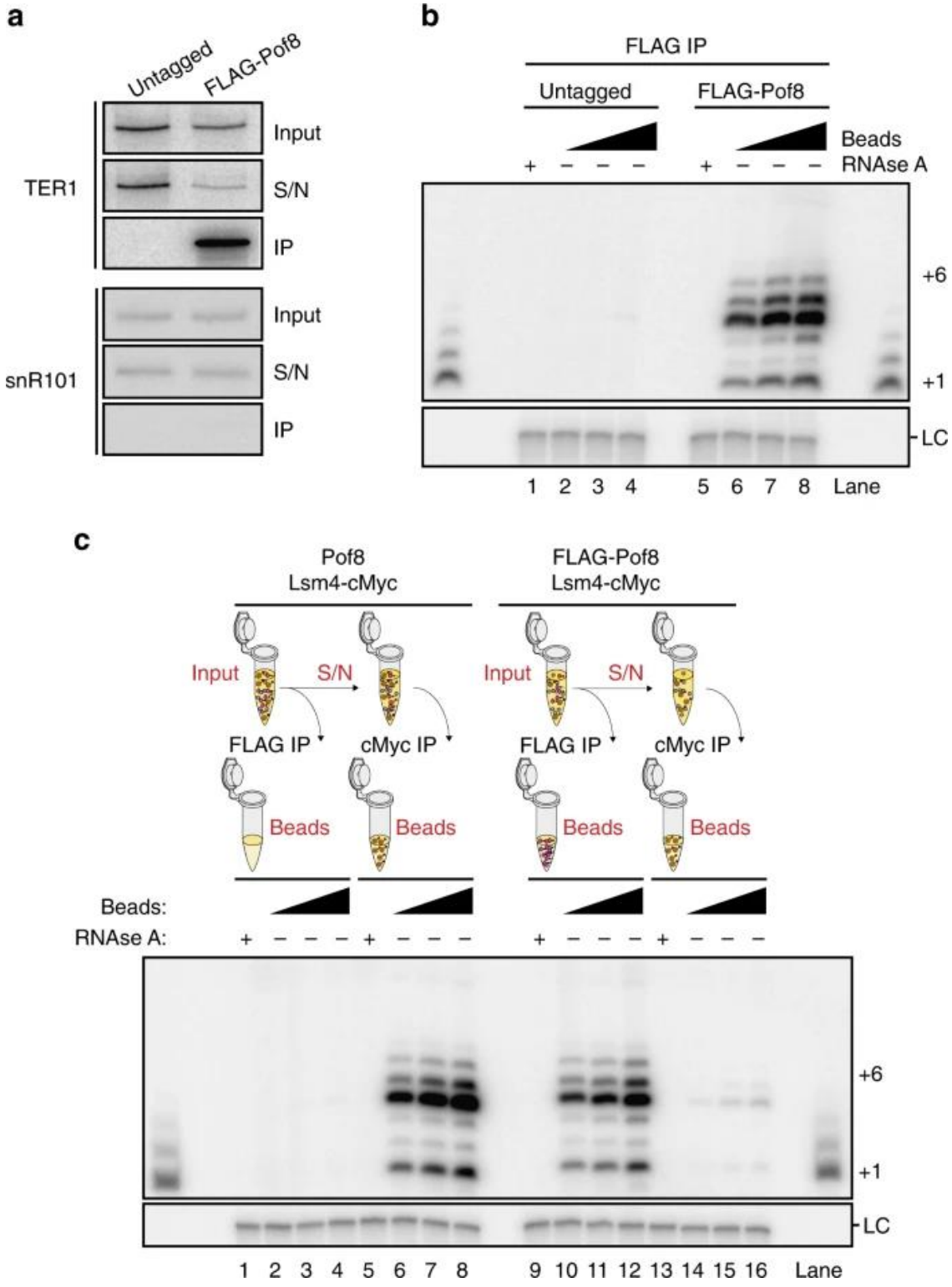
a
Untagged
FLAG-Pof8
TER1
Input
S/N
IP
snR101
Input
S/N
IP
b
FLAG IP
Untagged
FLAG-Pof8
Beads
RNAse A
+6
+1
LC
1 2 3 4 5 6 7 8 Lane
c
Pof8
Lsm4-cMyc
FLAG-Pof8
Lsm4-cMyc
Input
S/N
FLAG IP
cMyc IP
Beads
Beads:
RNAse A:
+6
+1
LC
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 Lane

Figure 2.8: Pof8 is a constitutive member of telomerase holoenzyme. **a)** Northern blot analysis of TER1 from α-FLAG immunoprecipitation (IP) from Pof8 and FLAG-Pof8 strains. Input and supernatant (S/N) represent 10% of the IP samples. snR101 was used as a negative control. **b)** Telomerase activity assay from FLAG IP samples of FLAG-Pof8 and untagged Pof8 strains in an Lsm4-cMyc background. RNAse A was added to beads before performing telomerase assays in control samples. A ladder of extension products generated by terminal transferase was run on both edges of the gel; the bands corresponding to +1 and +6 nucleotide addition products are indicated on the right. A ^{32}P-labeled 100-mer oligonucleotide was used as LC. **c)** Telomerase assay of samples generated by sequential IPs with FLAG and cMyc antibodies. A schematic of the experiment is shown above the blot: first, extracts were exposed to α-FLAG antibody-bound beads, the supernatant (S/N) from the α-FLAG IP was then exposed to α-cMyc antibody-bound beads followed by a telomerase assay for all groups of IP samples.

II.4.6: Pof8 is not a general loading factor for Lsm2-8

Considering the homology with La-related proteins and the stable association with telomerase, we wondered whether other RNAs were also affected by deletion of *pof8*. Using ribo-depleted RNA from otherwise isogenic *pof8*$^{+}$ and *pof8Δ* strains, we performed expression analysis in triplicate. Five protein-encoding transcripts and 13 non-coding RNAs including TER1 were found to be expressed at more than two-fold lower levels in *pof8Δ* cells (Table 2.3a). An even smaller number of transcripts was found to be upregulated by greater than two-fold (Table 2.3b). Among the seven upregulated protein-encoding genes was tlh2, a locus located in subtelomeric DNA and previously found to be upregulated in cells with critically short telomeres [210].

Table 2.3: Genes affected in expression level by the deletion of *pof8*.**a)** Genes with RNA levels decreased by more than 2-fold in the absence of Pof8 based on an EdgeR analysis of triplicate *pof8*$^{+}$ and *pof8Δ* RNA samples. **b)** Genes with increased expression by more than two-fold in the same dataset as in (a).

a)

gene ID	external gene ID	log FC	adj. p value	gene biotype
SPAC17G6.17	pof8	-3.1722	2.27E-44	protein_coding
SPNCRNA.214	ter1	-2.0020	6.76E-54	ncRNA
SPNCRNA.1554	SPNCRNA.1554	-1.9279	1.62E-13	ncRNA
SPAC29A4.12c	mug108	-1.8857	3.72E-17	protein_coding
SPNCRNA.690	prh1-antisense-1	-1.7065	2.58E-15	ncRNA

SPNCRNA.1340	SPNCRNA.1340	-1.6384	2.00E-20	ncRNA
SPNCRNA.716	SPAC56F8.12-antisense-1	-1.3851	5.08E-41	ncRNA
SPNCRNA.742	SPAC9.08c-antisense-1	-1.2879	6.18E-16	ncRNA
SPNCRNA.322	SPNCRNA.322	-1.2691	3.72E-17	ncRNA
SPCC569.02c	SPCC569.02c	-1.2528	1.34E-08	protein_coding
SPNCRNA.784	SPNCRNA.784	-1.2273	1.10E-06	ncRNA
SPNCRNA.1424	SPNCRNA.1424	-1.2010	1.37E-09	ncRNA
SPNCRNA.794	SPNCRNA.794	-1.1924	2.28E-08	ncRNA
SPBPB21E7.01c	eno102	-1.1252	2.39E-05	protein_coding
SPAPB24D3.07c	SPAPB24D3.07c	-1.0590	1.78E-28	protein_coding
SPBPB2B2.01	SPBPB2B2.01	-1.0412	3.06E-05	protein_coding
SPNCRNA.888	end4-antisense-1	-1.0053	1.41E-07	ncRNA
SPNCRNA.648	SPNCRNA.648	-1.0025	1.40E-07	ncRNA
SPNCRNA.779	SPNCRNA.779	-1.0016	5.71E-05	ncRNA

b)

gene ID	**external gene ID**	**log FC**	**adj. p value**	**gene biotype**
SPCC1020.09	gnr1	1.5878	2.20E-12	protein_coding
SPBCPT2R1.08c	tlh2	1.3332	5.09E-11	protein_coding
SPAC27D7.11c	SPAC27D7.11c	1.2919	1.39E-43	protein_coding
SPBC839.06	cta3	1.2484	8.90E-09	protein_coding
SPRRNA.02	15S_rRNA	1.1726	3.69E-02	rRNA
SPBC685.07c	rpl2701	1.1687	2.69E-20	protein_coding
SPAC13G7.08c	crb3	1.1250	1.49E-16	protein_coding
SPCC162.05	coq3	1.0501	4.39E-05	protein_coding
SPRRNA.01	21S_rRNA	1.0212	3.19E-02	rRNA

Although U6 snRNA associates with the Lsm2-8 complex like TER1, the U6 expression level was unaffected by the presence or absence of Pof8. Furthermore, immunoprecipitates from FLAG-Pof8 extracts contained barely more U6 than control IPs from extracts with untagged Pof8 (Figure 2.9a). This argues against Pof8 being a general loading factor for Lsm2-8 and a component of the U6 snRNP. To further assess the specificity of Pof8 for loading the Lsm complex onto RNAs, we asked how many other RNAs that are associated with Lsm8 are affected in abundance by deletion of *pof8*. We chose Lsm8 for this experiment as it is the only Lsm family member unique to the Lsm2-8 complex, all others being shared by the cytoplasmic Lsm1-7 complex. Immunoprecipitation

of Lsm8-cMyc enriched 35 RNAs by greater than two-fold, including U6 (30-fold enrichment) and TER1 (159-fold enrichment) (Table 2.4).

Table 2.4: Genes enriched in Lsm8-cMyc immunoprecipitation

gene ID	external gene ID	control normalized read count	Lsm8 IP normalized read count	FC	gene_biotype	comment
SPNCRNA.1532	SPNCRNA.1532	0.8	742.6	915.84	ncRNA	annotation overlaps with 1531
SPNCRNA.1531	SPNCRNA.1531	0.9	743.0	785.41	ncRNA	annotation overlaps with 1532
SPSNRNA.04	snu4	78.1	36400.4	466.19	snRNA	
SPNCRNA.214	ter1	3.4	534.1	159.38	ncRNA	
SPNCRNA.240	SPNCRNA.240	0.7	96.1	142.30	ncRNA	
SPSNORNA.39	snR36	0.4	45.8	112.88	snoRNA	
SPBC16E9.18	psd1	1.3	45.9	34.68	protein_coding	
SPSNRNA.06	snu6	387.3	11785.5	30.43	snRNA	
SPBTRNAGLY.09	SPBTRNAGLY.09	0.7	18.7	27.70	tRNA	
SPNCRNA.895	cox15-antisense-1	1.1	27.4	26.02	ncRNA	
SPAC1486.08	cox16	1.2	28.1	23.13	protein_coding	
SPSNRNA.02	snu2	12.2	227.9	18.74	snRNA	
SPSNORNA.21	snoU14	3.8	70.5	18.38	snoRNA	
SPAC24C9.16c	cox8	1.2	20.3	17.05	protein_coding	
SPNCRNA.445	snoR61	2.2	26.9	12.46	snoRNA	
SPSNRNA.05	snu5	8.7	108.3	12.40	snRNA	
SPAC10F6.17c	SPAC10F6.17c	1.6	13.6	8.69	protein_coding	
SPAC1B3.12c	rpb10	4.1	31.2	7.64	protein_coding	
SPBC354.06	mrps16	1.8	12.1	6.61	protein_coding	
SPNCRNA.905	SPAC8E11.01c-antisense-1	2.5	15.7	6.37	ncRNA	
SPMIT.07	atp6	5.5	31.3	5.73	protein_coding	
SPAC3H5.05c	rps1401	10.8	53.0	4.93	protein_coding	
SPBC1539.06	acb1	3.2	15.0	4.75	protein_coding	
SPAC10F6.16	mug134	3.2	15.4	4.75	protein_coding	
SPAC3H5.04	aar2	11.8	53.3	4.52	protein_coding	
SPAC4F10.20	grx1	4.3	17.5	4.04	protein_coding	
SPAPB15E9.01c	pfl2	10.7	42.3	3.95	protein_coding	
SPSNORNA.40	snR42	12.6	41.0	3.25	snoRNA	
SPSNORNA.32	sno12	34.1	103.4	3.04	snoRNA	
SPNCRNA.1464	SPNCRNA.1464	8.4	23.0	2.73	ncRNA	
SPAC24C9.03	mvd1	9.6	24.7	2.58	protein_coding	
SPAC343.12	rds1	7.9	20.1	2.53	protein_coding	
SPAC1635.01	SPAC1635.01	8.8	21.4	2.43	protein_coding	
SPAC343.20	SPAC343.20	8.0	19.2	2.40	protein_coding	
SPBC8D2.04	hht2	9.3	21.0	2.25	protein_coding	

Overlaying the differential expression data for *pof8* with the Lsm8 IP revealed TER1 as the only transcript that is bound by Lsm8 and reduced in *pof8Δ* cells (Figure 2.9b).

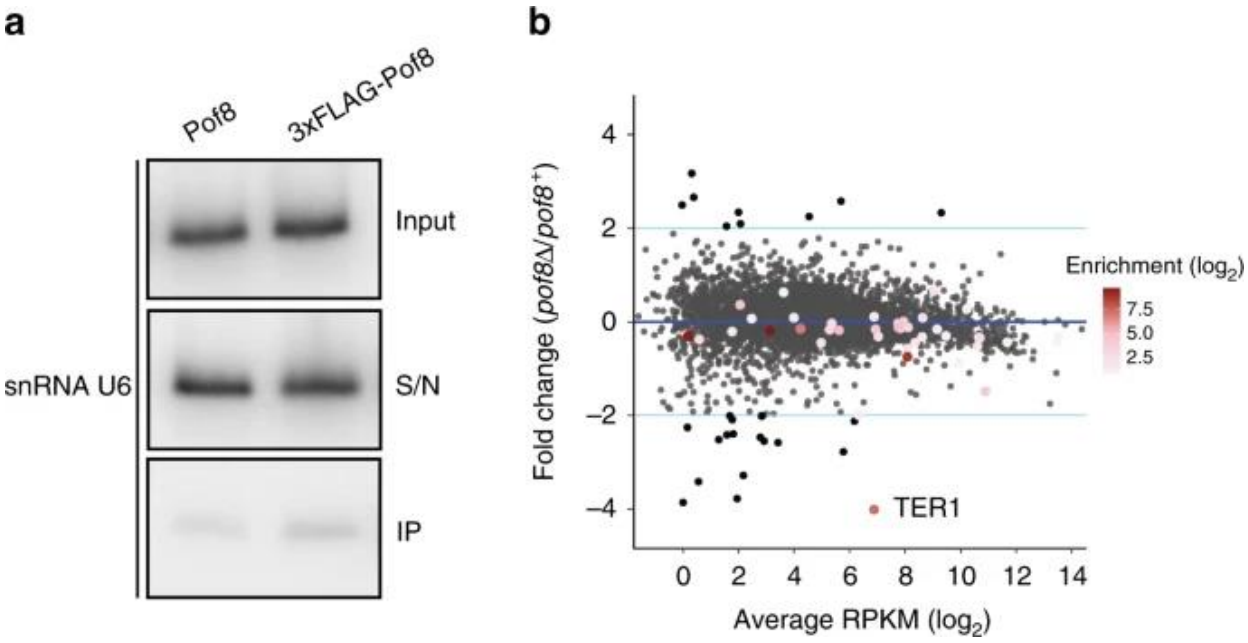

Figure 2.9: Pof8 is not a general loading factor for Lsm2-8. **a)** Northern blot for U6 snRNA from input (10%), supernatant (S/N, 10%), and immunoprecipitates (IP, 100%) with anti-FLAG from extracts of cells expressing untagged or FLAG epitope-tagged Pof8. **b)** Scatter plot relating changes in gene expression upon deletion of *pof8* to the enrichment of RNAs in Lsm8 immunoprecipitate. Gene expression in $\log_2$ average RPKM is plotted on the *x*-axis and the $\log_2$ fold change between *pof8Δ* and wildtype on the *y*-axis. Differentially expressed genes with an absolute $\log_2$ fold change of ≥1 and an adjusted *p*-value < 0.05 are colored black. Genes that are enriched in the Lsm8 IP are colored by $\log_2$ (Lsm8 IP/Control IP) as indicated by the color gradient in the legend.

II.5: Discussion

In summary, our results demonstrate that Pof8 is a previously unrecognized member of the LARP family that shares structural and functional similarity with telomerase subunits from ciliated protozoa, namely p43 and p65. Pof8 binds to telomerase RNA and promotes the loading of the Lsm2-8 complex, which in turn promotes the loading of the catalytic subunit Trt1. Pof8 is associated with nearly all telomerase activity, establishing the protein as a bona fide protein component of functional telomerase in fission yeast.

The identification of La-related proteins in association with highly purified telomerase from *Euplotes* and *Tetrahymena* was initially seen as unsurprising due to the nature of these RNAs as pol III transcripts. La protein associates rapidly and in most cases transiently with nascent pol III transcripts, and guides the RNAs through various processing

steps [211]. Interestingly though, subsequent studies revealed that p43 and p65 are telomerase-specific proteins, suggesting functions distinct from those provided by the canonical La proteins [77,79]. A series of elegant biochemical [79,82], structural [81,212], and single-molecule [85] experiments have since demonstrated that p65 binding to telomerase RNA induces structural changes in the RNA that are instrumental for the hierarchical assembly of functional telomerase in *Tetrahymena*.

Much less is known about the assembly pathways for telomerase from other organisms, but recent findings have revealed more differences than commonalities. Budding yeast TLC1 is stably associated with Sm proteins [132], whereas fission yeast TER1 undergoes the Sm to Lsm switch [119]. While divergent yeast telomerase RNAs share features with snRNAs, vertebrate TRs belong to the family of H/ACA box snoRNAs [92] and scaRNAs [102], and associate with a different set of proteins to mediate 3′ end processing and stability [95,96,98,213]. The characterization of several telomerase components that lack clear functional orthologs in other eukaryotes supports the view that telomerase evolves far more rapidly than other well-characterized RNPs, and that the presence of reverse transcriptase and a highly divergent non-coding RNA may be the only common denominators. Along these lines, it was recently shown that Pop1, 6, and 7, previously characterized as binding the P3 domains of RNase P and MRP, are also constitutive components of telomerase in budding yeast and are critical for holoenzyme integrity in this organism [115,214]. Whether this remarkable feature of a conserved functional module being shared between three highly divergent ancient RNPs is conserved in other species remains to be seen. It appears that telomerase RNA easily acquires functional sequence modules, such as the P3 or H/ACA domains, and that active telomerase can be produced via diverse pathways.

II.6: Permission from Nature Publishing

Adapted article publication permissions from Springer Nature Group

Chapter II is adapted from a previously published article (Páez-Moscoso, D. J., Pan, L., Sigauke, R. F., Schroeder, M. R., Tang, W., & Baumann, P. (2018). Pof8 is a La-related protein and a constitutive component of telomerase in fission yeast. *Nature Communications*, *9*(1), 587. https://doi.org/10.1038/s41467-017-02284-8). Nature Communications is a publication of Springer Nature and is an open-access journal wherein authors retain the permission to publish the article in part or whole as part of a published thesis (as below).

Get permission to reuse Springer Nature content online

Permission requests from authors

The author of articles published by Springer Nature do not usually need to seek permission for re-use of their material as long as the journal is credited with initial publication.

Ownership of copyright in in original research articles remains with the Author, and provided that, when reproducing the contribution or extracts from it or from the Supplementary Information, the Author acknowledges first and reference publication in the Journal, the Author retains the following non-exclusive rights:

To reproduce the contribution in whole or in part in any printed volume (book or thesis) of which they are the author(s).

Chapter III: A putative cap binding protein and the methyl phosphate capping enzyme Bin3/MePCE function with the Larp7 family member Pof8 in the biogenesis of telomerase

III.1: Abstract

Telomerase core components are the reverse transcriptase (TERT) and the noncoding RNA subunit (TR). These two subunits are enough to synthetized DNA *in vitro* but insufficient *in vivo*. Additional subunits are required for the telomerase ribonucleoprotein assembly, quality control and recruitment to the telomeres. Pof8, a member of the LARP7 protein family, was shown to help during the assembly process which it also remains associated to the holoenzyme. Here using Pof8, we have identified two previously uncharacterized proteins involved in telomerase biogenesis. Both proteins are required for telomerase activity and telomere length maintenance. We named these proteins Thc1 (Telomerase Holoenzyme Component 1) and Bmc1 (Bin3/MePCE 1) based on structural and sequence similarities to the nuclear cap binding complex and the methyl phosphate capping enzyme (Bin3/MePCE) in metazoans, respectively. Thc1 and Bmc1 function together with Pof8 in recognizing telomerase RNA and promoting the recruitment of the Lsm2-8 complex and the catalytic subunit to assemble functional telomerase.

III.2: Introduction

Telomeres are protective caps at the ends of eukaryotic linear chromosomes that prevent natural DNA ends from being recognized as double-strand breaks. Deficiency in maintaining proper telomere structure results in cellular senescence or chromosome instability that could lead to cancer [215,216]. An essential event to maintain telomeres is the extension of shortened telomeres by the telomerase, a specialized reverse transcriptase [217,218]. The telomerase consists of a catalytic protein telomerase reverse transcriptase (TERT, Trt1 in *Schizosaccharomyces pombe*) and the non-coding telomerase RNA subunit (TR, TER1 in *S. pombe*), functioning as the template and a scaffold for the holoenzyme assembly [76,219]. Besides the core components, there are unknown number of other subunits involved in the biogenesis, regulation, and recruitment of the enzyme to telomeres. Mutations in these subunits have been linked to human premature aging such as dyskeratosis congenital, aplastic anemia, idiopathic pulmonary fibrosis[220]. In addition, re-activation of telomerase is observed in ~85% of cancers [221]. Given the critical role of telomerase in cell proliferation, it is not surprising that telomerase RNP biogenesis is a tightly regulated hierarchical assembly process involving numerous factors[4,5].

TER1 is transcribed by RNA polymerase II as a precursor containing an intron flanked by two exons. The precursor contains important cis-regulatory elements that dictate the biogenesis process. Among them are the 5' splice site (5'SS), a branch site, a 3' splice site (3'SS), and a uridine-rich Sm binding site, GU_6GA[118,120,121]. The Sm ring binds to the Sm site where it recruits the spliceosome [121]. Here, the spliceosome processes TER1 precursor by cleaving exon 1 and discarding the intron–exon 2, a process called spliceosomal cleavage [121,122]. The Sm complex also recruits the methyltransferase Tgs1 to hyper-methylate the m7G cap to 2,2,7-trimethylguanosine (TMG) cap [119]. Perhaps after cap methylation, TER1 pre-

mature form is further trimmed at the 3' end leaving a molecule of ~1213 nucleotides in length with a poly(U) tail [119,121]. The poly(U) tail increases the affinity for the nuclear pre-formed Lsm2-8 ring, which replaces its paralog Sm complex. The Lsm2-8 complex protects the RNA against exonucleases like the exosome and promotes assembly to the catalytic subunit, Trt1[119,123,222].

The efficient sequential binding of Sm and Lsm2-8 complexes involves the La-related protein Pof8[192,223,224]. Pof8 was originally described as an F-box protein [206], but based on sequence conservation and domain organization, Pof8 is closer to the LARP7 family. Pof8 is structurally similar to the telomerase subunits p43 and p65 from ciliated protozoa and the LARP7 protein from the 7SK complex in mammalian [225]. The human Larp7 proteins have also been implicated in telomere maintenance [226]; Larp7 loss of function mutations in patients with Alazami syndrome have very short telomeres and Larp7 knockdown causing a reduction in telomerase activity in cancer cells.

Pof8 has a conserved N-terminal La module (La motif plus RNA recognition motif 1 (RRM1)) and a C-terminal RRM2. The conservation of the RNP3 motif (Y/W-X-D/Q) and the presence of an α3 helix, as revealed by the crystal structure and NMR studies of the C-terminal domain, identifies RRM2 as an xRRM despite the absence of an extended α3 helix[29]. Loss of Pof8 reduces TER1 stability, causes a prominent assembly defect, and results in critically short telomeres [192,223,224]. Recent work suggested Pof8 binds primarily to the pseudoknot region of TER1 via its RRM1 and xRRM domains, and it recruits the Lsm2-8 complex to the U-rich 3' end via direct protein-protein interactions through its most N-terminal region [227]. This extensive contact among Pof8-TER1-Lsm2-8 ensures the quality of the TER1 and final recruitment to telomerase [227]. The observation that Pof8 is associated with nearly 100% of telomerase enzyme [192] supports its role in stabilizing an active conformation of the holoenzyme, rather than functioning solely as an assembly chaperon.

In this study, we wonder if Pof8 has binding partners in addition to the known telomeric components resembling the behavior of other members of the LARP7 family. We identified that Pof8 forms a complex with at least two additional proteins. These proteins are Bmc1, the putative fission yeast ortholog of metazoan methyl phosphate capping enzyme, and Thc1, a previously uncharacterized gene (SPCC18b5.09c). Deletion of *bmc1* or *thc1* results in decreased TER1 steady-state levels, telomere shortening, and assembly defect by reducing Pof8 loading onto TER1. However, Lsm2-8 complex loading appears not to be affected in *bmc1* deletion cells. Together, the results provide new insights into telomerase biogenesis as a hierarchical assembly process. Also, the structural similarity of Thc1 with human PARN, and Bmc1 with MePCE provides a new evolutionary link in the biogenesis of telomerase in distant organisms and interactions between distant non-coding RNA machinery such as 7SK RNA.

III.3: Methods

III.3.1: Strains and constructs

S. pombe strains used in this study are listed in Table 3.1. The *bmc1* deletion was generated by replacing the entire open reading frame with the NatMX6 resistance cassette in a diploid strain generated fresh by crossing PP137 and PP139. The knockout fragment was generated by fusion PCR using primers listed in Table 3.2 and contained the NatMX6 cassette and ~750 base pair (bp) of homology upstream and downstream of the *bmc1* open reading frame. Cells were transformed as described[198] and allowed to recover in a shaker at 25 °C overnight. Diploid transformants were then selected on yeast extract low adenine (YEA) plates supplemented with nourseothricin (NAT, 100 μg ml^{-1}). The correct insertion of the NatMX6 cassette was confirmed by PCR and haploid cells were obtained by sporulation

and selection on YEA plus NAT plates. The absence of the *bmc1* ORF in the haploid cells was confirmed by PCR and sequencing.

Table 3.1: *S. pombe* strains used in this study

Strain	Genotype	Source	Figure
FP1546	*h⁻ ade6-M216 leu1-32 ura4-D18 his3-D1 lsm4::lsm4-myc13-natMX6 pof8::kanMX6 [pDBlet-Pof8]*	Páez-Moscoso et al 2018	1a
FP1547	*h⁻ ade6-M216 leu1-32 ura4-D18 his3-D1 lsm4:: lsm4-myc13-natMX6 pof8::kanMX6 [pDBlet-3xFLAGPof8]*	Páez-Moscoso et al 2018	1a
FP1913	*h? ade6-M21? leu1-32 ura4-D18 his3-D1 bmc1(SPBC2A9.10)::natMX6 [ura4, pDBlet]*	This study	S4a,b
FP1914	*h? ade6-M21? leu1-32 ura4-D18 his3-D1 bmc1(SPBC2A9.10)::natMX6 [pDBlet–bmc1 WT]*	This study	S4a,b
FP1915	*h? ade6-M21? leu1-32 ura4-D18 his3-D1 bmc1(SPBC2A9.10)::natMX6 [pDBlet–bmc1-Twinstrep]*	This study	S4a,b
FP1916	*h? ade6-M21? leu1-32 ura4-D18 his3-D1 bmc1(SPBC2A9.10)::natMX6 [pDBlet–bmc1-Twinstrep F246A]*	This study	S4a,b
FP1917	*h? ade6-M21? leu1-32 ura4-D18 his3-D1 bmc1(SPBC2A9.10)::natMX6 [pDBlet–bmc1-Twinstrep VLD39-41AAA]*	This study	S4a,b
FP1918	*h? ade6-M21? leu1-32 ura4-D18 his3-D1 bmc1(SPBC2A9.10)::natMX6 [pDBlet–bmc1-Twinstrep GCN43-45ACA]*	This study	S4a,b
FP1919	*h? ade6-M21? leu1-32 ura4-D18 his3-D1 bmc1(SPBC2A9.10)::natMX6 [pDBlet–bmc1-Twinstrep EPQ188-190APA]*	This study	S4a,b
PP137	*h⁺ ade6-M216 leu1-32 ura4-D18 his3-D1*	Lab stock	
PP138	*h⁻ ade6-M216 leu1-32 ura4-D18 his3-D1*	Lab stock	2a, 2b, 2c, 2d, S3, S5
PP139	*h⁻ ade6-M210 leu1-32 ura4-D18 his3-D1*	Lab stock	
PP1723	*h⁻ ade6-M216 leu1-32 ura4-D18 his3-D1 pof8::kanMX6*	Páez-Moscoso et al 2018	2a, 2b, 2d, 5a
PP1797	*h? ade6-M21? leu1-32 ura4-D18 his3-D1 lsm4::lsm4-myc13-nat aur1::[pCST159-ter1] pof8::kanMX6*	This study	3g
PP1839	*h? ade6-M210/M213 ura4-D18 leu1-32 his3-D1 pof8::3xFLAG-pof8-kanMX6*	This study	2c
PP1843	*h⁻ ade6-M216 leu1-32 ura4-D18 his3-D1 thc1 (SPCC18B5.09c):: thc1-TEV-2xV5-natMX6*	This study	1d, 2c
PP1844	*h⁻ ade6-M216 leu1-32 ura4-D18 his3-D1 bmc1(SPBC2A9.10)::bmc1-TEV-TwinStrep-natMX6*	This study	1c, 2c
PP1845	*h⁻ ade6-M216 ura4-D18 leu1-32 3xFLAG-pof8-kan thc1 (SPCC18B5.09c):: thc1-TEV-2xV5-natMX6*	This study	1d, 3c
PP1846	*h⁻ ade6-M216 ura4-D18 leu1-32 3xFLAG-pof8-kan bmc1(SPBC2A9.10)::bmc1-TEV-TwinStrep-natMX6*	This study	1c, 3d
PP1847	*h⁻ ade6-M216 leu1-32 ura4-D18 his3-D1 thc1 (SPCC18B5.09c)::his3*	This study	2a, 2b, 2d, S3, S5
PP1857	*h? ade6-M210/M216 leu1-32 ura4-D18 his3-D1*	This study	3a, 3b
PP1858	*h? ade6-M210 leu1-32 ura4-D18 his3-D1 pof8::KanMX6*	This study	3a, 3b
PP1859	*h⁺ ade6-M210 leu1-32 ura4-D18 his3-D1 bmc1(SPBC2A9.10)::natMX6*	This study	3a, 3b
PP1860	*h? leu1-32 ura4-D18 his3-D1 ade6-M210/ade6-M216*	This study	2a, 2b, S3
PP1861	*h? leu1-32 ura4-D18 his3-D1 ade6-M210/ade6-M216 bmc1(SPBC2A9.10)::natMX6*	This study	2a, 2b, 2d, S3

PP1862	*h? ade6-M210 leu1-32 ura4-D18 his3-D1 thc1 (SPCC18B5.09c)::his3*	This study	3a, 3b
PP1863	*h? ade6-M210/216 leu1-32 ura4-D18 his3-D1 pof8::kanMX6 bmc1(SPBC2A9.10)::natMX6*	This study	3a, 3b
PP1864	*h? ade6-M216 leu1-32 ura4-D18 his3-D1 pof8::kanMX6 thc1 (SPCC18B5.09c)::his3*	This study	3a, 3b
PP1865	*h? ade6-M210/216 leu1-32 ura4-D18 his3-D1 bmc1(SPBC2A9.10)::natMX6 thc1 (SPCC18B5.09c)::his3*	This study	3a, 3b
PP1866	*h? ade6-M210 leu1-32 ura4-D18 his3-D1 bmc1(SPBC2A9.10)::natMX6 thc1 (SPCC18B5.09c)::his3*	This study	3a, 3b
PP1882	*h? ade6-M210 leu1-32 ura4-D18 his3-D1 smb1:: smb1-cMyc-natMX6, pof8::kanMX6*	This study	4c, 4d
PP1883	*h? ade6-M210 leu1-32 ura4-D18 his3-D1 smb1:: smb1-cMyc-natMX6, thc1 (SPCC18B5.09c)::his3*	This study	4c, 4d
PP1884	*h? ade6-M216 leu1-32 ura4-D18 his3-D1 smb1:: smb1-cMyc-natMX6, bmc1(SPBC2A9.10)::natMX6*	This study	4c, 4d
PP1885	*h? ade6-M210/M216 leu1-32 ura4-D18 his3-D1 smb1::smb1-cMyc-natMX6*	This study	4c, 4d
PP1886	*h? ade6-M216 leu1-32 ura4-D18 his3-D1 lsm4::lsm4-cMyc-natMX6, pof8::kanMX6*	This study	4a, 4b
PP1887	*h? ade6-M210 leu1-32 ura4-D18 his3-D1 lsm4::lsm4-cMyc-natMX6, thc1 (SPCC18B5.09c)::his3*	This study	4a, 4b
PP1888	*h? ade6-M216 leu1-32 ura4-D18 his3-D1 lsm4::lsm4-cMyc-natMX6, bmc1(SPBC2A9.10)::natMX6*	This study	4a, 4b
PP1889	*h? ade6-M210/M216 leu1-32 ura4-D18 his3-D1 lsm4::lsm4-cMyc-natMX6*	This study	4a, 4b
PP1892	*h? ade6-M216 leu1-32 ura4-D18 his3-D1 thc1 (SPCC18B5.09c)::spcc18b5.09c-TEV-2xV5-natMX6 bmc1(SPBC2A9.10)::bmc1-TEV-Twinstrep-natMX6*	This study	1e, 3e
PP1894	*h? ade6-M216 leu1-32 ura4-D18 his3-D1 bmc1(SPBC2A9.10)::bmc1-TEV-Twinstrep-natMX6*	This study	1e
PP1895	*h? ade6-M216 leu1-32 ura4-D18 his3-D1 pof8::kanMX6 thc1 (SPCC18B5.09c)::spcc18b5.09c-TEV-2xV5-natMX6 bmc1(SPBC2A9.10)::bmc1-TEV-Twinstrep-natMX6*	This study	3e
PP2014	*h? ade6-M21? leu1-32 his3-D1 ura4-D18 pof8::3xFLAG-pof8-kanMX6*	This study	5a, 5b, 5c
PP2015	*h? ade6-M21? leu1-32 his3-D1 ura4-D18 pof8::3xFLAG-pof8-kanMX6 thc1 (SPCC18B5.09c)::his3*	This study	5a, 5b, 5c
PP2016	*h? ade6-M21? leu1-32 his3-D1 ura4-D18 pof8::3xFLAG-pof8-kanMX6 bmc1(SPBC2A9.10)::natMX6*	This study	5a, 5b, 5c
PP2017	*h? ade6-M21? leu1-32 his3-D1 ura4-D18 pof8::3xFLAG-pof8-kanMX6 thc1 (SPCC18B5.09c)::his3 bmc1(SPBC2A9.10)::natMX6*	This study	5a, 5b, 5c

The *thc1* deletion was generated by replacing the open reading frame with the *his3*$^{+}$ gene using primers listed in Table 3.2 and the same strategy as above except that haploid strain PP138 was used for transformation. Epitope tags were introduced as follows: 3xFLAG tag was introduced at the N-terminus of Pof8, 2xV5 at the C-terminus of Thc1 and Twinstrep tag at the C-terminus of Bmc1. Additional strains were generated by crossing and selection of correct genotypes followed by verification by PCR or western blotting. A DNA fragment encompassing Twinstrep-tagged Bmc1 from position -408 to +682 was generated by gene

synthesis and cloned into pDBlet plasmid[201]. Point mutants were generated by site directed mutagenesis in this context and were verified by Sanger sequencing. Plasmids were introduced into *S. pombe* cells by electroporation.

Table 3.2: Oligonucleotides and gene synthesis products used to generate deletion, fusion and integration constructs.

Product description	**Primer #**	**Sequence**
Nat MX6 for *bmc1* deletion	BLoli6066/BLoli2491	5'-TTTAGCTTGCCTCGTCCCCG-3'/5'-TGGATGGCGGCGTTAGTATC-3'
bmc1 5' homology	BLoli7736/BLoli7737	5'-ATCCTGAAGCGATGATGCCA-3'/ 5'-CGGGGACGAGGCAAGCTAAACCCAAGTCGAGGAGGTTTTT-3'
bmc1 3' homology	BLoli7738/BLoli7739	5'-GATACTAACGCCGCCATCCATTGTCTAGTAAAACGTTAAG-3'/5'-TTGGCGAGTATAACCAATGT-3'
bmc1::natMX6	BLoli7736/BLoli7739	5'-ATCCTGAAGCGATGATGCCA-3'/5'-TTGGCGAGTATAACCAATGT-3'
his3^{+} for *thc1* deletion	BLoli7486/BLoli7487	5'-GTTTTGAAGACGGTGATACACGTTGTAATG-3'/5'-ATTTATCTGTTTGCTTATCGAACTATACGG-3'
thc1 5' homology	BLoli7484/BLoli7485	5'-CAATAATAACTTTGCTTACGATTAATAG-3'/5'-TGTATCACCGTCTTCAAAACTTTTGGTACC-3'
thc1 3' homology	BLoli7488/BLoli7489	5'CGATAAGCAAACAGATAAATTAGAACACAGC-3'/5'-GAGACTAATTGGGTTAAACAAAG-3'
thc1::his3	BLoli7484/BLoli7489	5'-CAATAATAACTTTGCTTACGATTAATAG-3'/5'-GAGACTAATTGGGTTAAACAAAG-3'
bmc1 5' homology for tagging	BLoli7718/BLoli7709	5'-AGCTCTCGAATTGGCCCTG-3'/5'-TTTAGAAGTGTTATTTCTCAAATTGAGGATGACTCCATG-3'
bmc1 3' homology for tagging	BLoli7541/BLoli7542	5'-GCCATCCAGTTAATTGTCTAGTAAAACGTTAAGAATAG/5'-TGATGTTGGCGAGTATAAC-3'
natMX6 cassette for *bmc1* tagging	BLoli7710/BLoli7540	5'-TGAGAAATAACACTTCTAAATAAGCGAATTTC-3'/5'-TAGACAATTAACTGGATGGCGGCGTTAG-3'
bmc1-Twinstrep::natMX6	BLoli7718/BLoli7542	5'-AGCTCTCGAATTGGCCCTG-3'/5'-TGATGTTGGCGAGTATAAC-3'
thc1 5' homology for tagging	BLoli7711/BLoli7712	5'-TCTTAAGATATTTGGGCTATAAATG-3'/5'-TTTAGAAGTGTTATTACGTGGAATCTAATCC-3'
thc1 3' homology for tagging	BLoli7715/BLoli7716	5'-GCCATCCAGTACAGATAAATTAGAACACAGC-3'/5'-AACAAAGTAGTAACCAAGG-3'
natMX6 cassette for *thc1* tagging	BLoli7713/BLoli7714	5'-CACGTAATAACACTTCTAAATAAGCGAATTTCTTATGATTTATG-3'/5'-ATTTATCTGTACTGGATGGCGGCGTTAG-3'
Thc1-2xV5::natMX6	BLoli7711/BLoli7716	5'-TCTTAAGATATTTGGGCTATAAATG-3'/5'-AACAAAGTAGTAACCAAGG-3'
3xFLAG-Pof8 5' homology for tagging	BLoli6676/BLoli6400	5'-AAAAGAATTCAACATGGCAACTGCGACCAA-3'/5'-TTTAGAAGTGTTACTTTTTTAACATACGCCAATAATTC-3'
Pof8-kanMX6 3' homology for tagging	BLoli6401/BLoli6141	5'-GGCGTATGTTAAAAAAGTAACACTTCTAAATAAGCG-3'/5'-GCTTTCTTATTTGTAGAGACAATTG-3'
3xFLAG-Pof8-kanMX6	BLoli6676/BLoli6141	5'-AAAAGAATTCAACATGGCAACTGCGACCAA-3'/5'-GCTTTCTTATTTGTAGAGACAATTG-3'
Cloning of bmc1-Twinstrep into pDBlet	BLoli8069/BLoli8070	5'-CCGATAAGCTTAAACTATCTTAACCTGTCTACG-3'/5'-GGTGGCGGCCGCTACAGTTTTGGTATACCAGG-3'
bmc1 F246A 5' arm	Bloli8069/BLoli8006	5'-CCGATAAGCTTAAACTATCTTAACCTGTCTACG-3'/5'-GTTCGTTTGGCAGCATTCTTGTAC-3'
bmc1 F246A 3' arm	BLoli8007/BLoli8070	5'-GTACAAGAATGCTGCCAAACGAAC-3'/5'-GGTGGCGGCCGCTACAGTTTTGGTATACCAGG-3'

bmc1 VLD39–41AAA 5' arm	Bloli8069/BLoli7792	5'-CCGATAAGCTTAAACTATCTTAACCTGTCTACG-3'/5'-CATTATTGCATCCTATCGCTGCGGCTGAAGCCTC-3'
bmc1 VLD39–41AAA 3' arm	BLoli7793/BLoli8070	5'-CAGCCGCAGCGATAGGATGCAATAATGGGAC-3'/5'-GGTGGCGGCCGCTACAGTTTTGGTATACCAGG-3'
bmc1 GCN43–45ACA 5' arm	Bloli8069/BLoli7795	5'-CCGATAAGCTTAAACTATCTTAACCTGTCTACG-3'/5'-GAGCAGACACTGTCCCATTAGCGCACGCTATTGCT-3'
bmc1 GCN43–45ACA 3' arm	BLoli7796/BLoli8070	5'-GTGCGCTAATGGGACAGTGTCTGCTCAAATTG-3'/5'-GGTGGCGGCCGCTACAGTTTTGGTATACCAGG-3'
bmc1 EPQ188–190APA 5' arm	Bloli8069/BLoli7802	5'-CCGATAAGCTTAAACTATCTTAACCTGTCTACG-3'/5'-TCAAGTACGAGTCCCATCCGGCAGGTGCTAAAATAAG-3'
bmc1 EPQ188–190APA 3' arm	BLoli7803/BLoli8070	5'-ACCTGCCGGATGGGACTCGTACTTGAAAGCTG-3'/5'-GGTGGCGGCCGCTACAGTTTTGGTATACCAGG-3'
Bmp1 full length PCR verification	BLoli8048/BLoli7740	5'-GAGAGACTAGCACCAATGTATCC-3'/5'-TAGCTCTTGAAAAGGTACGG-3'
5' arm bmp1::natMX6 PCR verification	BLoli7735/BLoli3688	5'-CACTACAATTCAAGACTACC-3'/5'-GTAAGCCGTGTCGTCAAGAG-3'
3' arm bmp1::natMX6 PCR verification	BLoli7740/BLoli3791	5'-TAGCTCTTGAAAAGGTACGG-3'/5'-GGCGCTCTACATGAGCATGC-3'
5' arm bmp1 PCR verification	BLoli8048/BLoli8049	5'-GAGAGACTAGCACCAATGTATCC-3'/5'-AGGGAATCAGGCAAACATTTAAG-3'
thc1::his3 full length PCR verification	BLoli7484/BLoli7717	5'-CAATAATAACTTTGCTTACGATTAATAG-3'/5'-TGAAGCACCAAGATACGAGT-3'

Gene synthesis products

thc1-TEV-2xV5 fragment	BLgs1466	5'- TGCTGTATCGGAGCCTAATTCTAGTGGTCTTAAGATATTTGGGCTATAAATGATTGAAAATTGGC AATCTTGTTCACGCTGATCAACAATAAAAAAACGGCTACTTAGTATTTACCCTATACTCTTAGCA TGGATACCTTCAATATTCATTGAAGTTGTAGTTTTAGCTTATTTGGTACCAAAAGTTTTGAAGAAT GGAAGAGAAAAATACTGTTTCTTTATCTAAGCATATTGAACGTCCAGTAGAAGTTGTTGAAAGTC ATTCTACGTACATTTTAAGTGCACAAGGACTTTATCTTACAGAACGCGTTTTAAGAAGTTATTTTA AACAACCTGATCTAATTATTACTTGGAAGGATAGTATGAGAGCTTACCTGACATTTTCTTCGCCG CAGGAAGCTCAAAAGGCTTACTTAGATTCACTTCGTTGGGGCAGTCAACTGAATGCTATCATTAA ACCATTCTACGGTTCGCACGATGAGGTACTTCGTTTATGTAAAAGGAAAAGAATTATTCCACTAC AAAATTTCTTGACTTCAGGTCTCGAGCCAACCACTGAGGATCTGTACTTTCAGAGCGATAACGAT GGTAAGCCTATCCCTAACCCTCTCCTCGGTCTCGATTCTACGGGCGGAGGCTCCGGCGGCGGCTC TGGAGGATCTGCAGGTAAGCCAATACCCAACCCACTTCTTGGATTAGATTCCACGTAATAACACT TCTAAATAAGCGAAT-3'

bmc1-TEV-Twin-Streptag fragment	BLgs1465	5'- AGCTCTCGAATTGGCCCTGTACGCAATCCTGGTTCCATTGTAGAAGACCAGTTTAATTATTACCCC ATTTCAAGCATTAAAAAGTTTTCCAGGATACCAGTGCAACTTCAACCACCTCTCAATAAGCAAAA TTTCCCTCACAATATAGAATTTGAGACCGCTGACTTCTTGCGCTGGGAATCGAAACGAAAATTCA AAATAATACTAGCATTATCCGTATCTAAATGGGTGCATCTAAATAACCACGATGAAGGAATCATT AAATTCTTTGGGAAGATTAGTTCTTTATTGGAAACGAATGGTGTTCTTATTTTAGAACCTCAAGGA TGGGACTCGTACTTGAAAGCTGCAAAAAAAATATCTGTAAGTCACGACTATCAATACTTCTAACT CTTACTTTTCTAGGTTTTTAATCAAACACCTGAGAACCTCAAAATCCAACCTGATGCGTTTGAACA TTTGCTTAATCAAGCAGGACTAGTGCTTGAATACAGTATCGAACCTCAAGTAAATAACTCTGAGT ACAAGAATTTTGCCAAACGAACAATGTATATCTATAAAAAAAAAGGAATTGGAATCATAAAACT ATTAACTTCTACTCTCGAGCCAACAACTGAAGATTTATATTTTCAATCTGACAATGATTGGTCTCA CCCACAGTTTGAAAAAGGCGGAGGCTCCGGCGGCGGCTCTGGAGGATCTGCATGGAGTCATCCT CAATTTGAGAAATAACACTTCTAAATAAGCGAATTTC-3'

III.3.2: Protein extracts

Cultures (2 l) were grown to a density of 0.5–1 × 10^7 cells per ml and harvested by centrifugation. Cells were washed three times with ice-cold TMG (300) buffer (10 mM Tris-

HCl pH 8.0, 1 mM magnesium chloride, 10% (v/v) glycerol, 300 mM sodium acetate), and resuspended in two packed cell volumes of TMG(300) plus supplements (complete EDTA-free protease inhibitor cocktail (Roche), 0.5 mM PMSF, 1 mM EDTA, and 0.1 mM DTT) and quick-frozen by dripping the cell suspension into liquid nitrogen. Cells were lysed in a 6850 Freezer mill (SPEX SamplePrep) using eight cycles (2 min) at a rate of 10 per second with 2 min cooling time between cycles. Lysates were thawed on ice and one additional packed cell volume of TMG(300) plus supplements was added. Lysates were then cleared by centrifugation twice for 10 min at 6000 × g in a Sorvall Linx 6000 with F13-14x50cy rotor and then once for 50 min in a Beckman 70Ti rotor at 36,000 × g. All steps were carried out at 4 °C. For the affinity purification followed by mass spectrometry experiment the extract was prepared as above with the following modification: the lysis buffer contained 50 mM Tris-HCl pH 7.5, 150 mM sodium chloride, 0.1% (v/v) NP40 and 10% (v/v) glycerol, complete EDTA-free protease inhibitor cocktail (Roche), 0.5 mM PMSF, 1 mM EDTA, and 0.1 mM DTT. Cell lysates were centrifuged for 15 min at 8000 x g in a Beckman JA-17 rotor immediately after thawing. The supernatant was collected, and one additional packed cell volume of lysis buffer was added to the pellet and centrifuged for 15 min at 12000 xg in a Beckman JA-17 rotor. The two supernatants were pooled together and then cleared by centrifugation for 40 min in a Beckman 70Ti rotor at 40,000 × g. For all extracts protein concentrations were determined by Bradford assay and ranged from 5 to 10 mg ml^{-1}.

Denatured protein extracts for Western blotting were prepared from 1 x 10^8 cells through lysis by vortexing with 0.5-mm glass beads in 10% trichloroacetic acid for 8 min at 4°C. Beads were washed once with 0.8 ml 10% trichloroacetic acid, the bead supernatant was transferred to a fresh tube and protein precipitate was collected by centrifugation at 16,000 × g for 2 min. Protein pellets were washed once with acetone and resuspended in 120 µl of 1× protein sample buffer (1× NuPAGE LDS buffer, 50 mM dithiothreitol, 2% [w/v] sodium

dodecyl sulfate). Samples were incubated for 5 min at 75°C and centrifuged at 16,000 × g for 1 min. The soluble fraction was collected and analyzed by SDS PAGE and western blotting.

III.3.3: Immunoprecipitations and mass spectrometry

For the identification of proteins that interact with Pof8, cell-free extracts of strains FP1547 harboring 3xFLAG epitope-tagged Pof8 and untagged control (FP1546) were diluted to 5 mg/ml in 8 ml of lysis buffer (50 mM Tris-HCl pH 7.5, 150 mM sodium chloride, 0.1% (v/v) NP40 and 10% (v/v) glycerol) plus complete EDTA-free protease inhibitor cocktail (Roche), 0.5 mM PMSF, 1 mM EDTA, and 0.1 mM DTT and incubated with 500 µl of EZview Red Flag-agarose 50% slurry equilibrated in lysis buffer (Sigma, F2426). The mix was incubated for 8 hours at 4 °C with gentle rotation. Agarose beads were collected by centrifugation at 300 x g for 5 min with low brake and washed 5 times for 5 minutes with 5 ml of wash buffer (50 mM Tris-HCl pH 7.5, 250 mM sodium chloride, 0.1% (v/v) NP40 and 10% (v/v) glycerol) at 4 °C with gentle rotation. Proteins were eluted from beads by incubation for 30 min with 500 µl elution buffer (50 mM Tris-HCl pH 7.5, 250 mM sodium chloride, 0.1% (v/v) NP40 and 10% (v/v) glycerol, and 100 µg ml^{-1} 3xFLAG peptide) at 4 °C. The elution step was repeated two more times and the eluates were pooled. A 4 µl aliquot of the first eluate was used for Silver stained SDS-PAGE analysis following manufacturer's instructions (SilverXpress Silver Staining Kit; Invitrogen, LC6100). The elutes were treated with 0.1U of benzonase (Sigma, E8263) for 30 min at 37ºC and further diluted with the same volume of 100 mM Tris-HCl, pH 8.5. The samples were then split in three 400µl aliquots, mixed with 100 µl of 100% trichloroacetic acid (TCA) and incubated at 4°C overnight. After centrifugation at 14,000 rpm for 30 min at 4°C, the pellets were washed twice with 500 µl cold acetone and centrifuged at 14,000 rpm for 10 min after each wash. Air dried pellets were resuspended in 30 µl 100 mM Tris-HCl, pH 8.5, 8M Urea and pooled. 4.5 µl of 0.1M Tris(2-carboxyethyl)-phosphine hydrovhloride (TCEP) was added to the pooled solution to a final

concentration of 5 mM and incubated for 30 min at room temperature (rt). After the incubation, 1.8 µl of 0.5 M CAM was added and incubated for 30 min at rt in the dark. Proteins were then digested with 0.1 µg $µl^{-1}$ Lys-C at 37 °C overnight. The samples were then diluted to 2M Urea with 100 mM Tris-HCl, pH8.5, added with $CaCl_2$ to 2 mM and digested with 0.1 µg $µl^{-1}$ Trypsin at 37°C overnight. Finally, 90% formic acid was added to a final concentration of 5%. and samples were analysed by MudPIT mass spectrometry as described[51]. Tandem mass (MS/MS) spectra were interpreted using SEQUEST against the NCBI database of *S. pombe* proteins supplemented with 177 sequences from common contaminants (human keratins, IgGs, proteolytic enzymes). To estimate relative protein levels, (dNSAFs) were calculated for each detected protein. The dNSAF takes into account the diversity of the peptides and the length of the protein for normalization [1].

For RNA immunoprecipitations, cell-free extracts (5.5 mg) were diluted to 5 mg ml^{-1} with TMG(300) buffer plus supplements. An aliquot (100 µl) was frozen as input control. Magnetic protein G dynabeads (30 mg ml^{-1}; Invitrogen) were coated with anti-c-Myc 9E10, anti-FLAG M2, or anti-V5 (20 µg per 100 µl of bead suspension; Sigma-Aldrich, M4439 and F3165, and Invitrogen, R960-25) by incubation for 30 min at room temperature in 200 µl of 1× PBS + 0.1% (v/v) Tween-20. The magnetic beads and Strep-Tactin Sepharose (28935599, GE Healthcare) were washed three times with 1 ml of TMG(300). Immunoprecipitations were performed with 60 µl of bead suspension for Pof8, Thc1, Bmc1 and 120 µl for Sm and Lsm for 4 hours at 4 °C with gentle rotation. Recombinant Pof8 / 6xHis-Thc1 was added to 50 nM during the incubation for experiments in Figure 3.7g. Beads were collected using a magnet or centrifugation at 300 x g for 2 min (Strep-Tactin Sepharose). The beads were then washed four times with 1 ml TMG(300) plus supplements and 0.1% (v/v) Tween-20, and once with TMG(50) plus supplements. Finally, beads were

resuspended in 120 µl TMG(50) plus supplements and 0.4 U/µl RNAsin (Promega) and frozen in liquid nitrogen.

For co-immunoprecipitation analysis, the input aliquot (100 µl) was mixed with equal volume of 2× protein sample buffer (2× NuPAGE LDS buffer (Life Technologies), 100 mM DTT, 4% (w/v) SDS) and incubated for 10 min at 75 °C. The bead suspension was divided into 4 equal volumes, three samples were incubated in 20 µl of TMG (50) plus 40 ng of RNAse A (Invitrogen), 500 U Benzonase (Sigma), or buffer for 15 min at 37 °C and then washed twice with TMG (300) and TMG (50). Finally, all the samples were resuspended in 20 µl of 1x protein sample buffer (1× NuPAGE LDS buffer (Life Technologies), 100 mM DTT, 4% (w/v) SDS), incubated for 10 min at 75 °C, and supernatants were analysed by western blotting.

III.3.4: Western blot analysis

Denatured protein extract or immunoprecipitates were loaded onto 4–12% NuPAGE Bis-Tris gel (Life Technologies, NP0321BOX) for electrophoresis in 1× MOPS buffer (Life technologies, NP0001) at 160 V for 60 min. Proteins were transferred to Protran nitrocellulose membranes (Whatman) in western transfer buffer (3.03 g l^{-1} Tris base, 14.4 g l^{-1} glycine, 20% (v/v) methanol) at 100 V for 1 h with a cold pad stored at -20ºC prior to use. Blots were processed using an iBind Flex Western Device (Life Technologies, SLF20002). The FLAG epitope tag was detected with mouse monoclonal anti-FLAG M2 antibody (Sigma Aldrich, F3165) at 1:4000 dilution and horse-radish peroxidase-conjugated goat anti-mouse IgG (H+L) at 1:5000 (Thermo Scientific, 31430). The V5 epitope-tag was detected with rabbit polyclonal anti-V5 antibody (Abcam, ab9116) at 1:2000 dilution and the Twinstrep tag with rabbit polyclonal anti-Strep-tag II antibody (Abcam, ab76949) at 1:2000 dilution; and horse-radish peroxidase-conjugated goat anti-rabbit IgG (H+L) at 1:5000 (Thermo Scientific,

31460). Secondary antibodies were visualized with ECL plus substrate (Pierce, 32132) on a Typhoon FLA9500 (GE Healthcare Life Sciences) for co-IP blots and a ChemiDoc MP Imager (Bio-rad) for denatured protein extract blots.

III.3.5: RNA extraction

For total RNA extraction, cell cultures (250 ml) were grown to a density of 6-8.5 × 10^6 cells ml^{-1} and collected by centrifugation, washed twice with H_2O (500 ml), resuspended in 3 ml H_2O and quick-frozen by dripping the cell suspension into liquid nitrogen. Cells were lysed in a 6850 Freezer mill (SPEX SamplePrep) using seven cycles (2 min) at a rate of 10 per second with 2 min cooling time between cycles. Lysed cells were transferred into 50 ml tubes containing 5 ml phenol:chloroform:isoamyl alcohol (25:24:1, equilibrated with 50 mM sodium acetate, pH 5.2) and 5 ml 50 mM sodium acetate and 1% (w/v) SDS preheated to 65 °C. RNA was extracted four times with 5 ml phenol: chloroform:isoamyl alcohol (25:24:1, equilibrated with 50 mM sodium acetate, pH 5.2) and once with chloroform:isoamyl alcohol (24:1, equilibrated with 50 mM sodium acetate, pH 5.2). RNA was ethanol precipitated and resuspended in H_2O. In the context of immunoprecipitation experiments, RNA was isolated from input and beads by incubation with proteinase K (2 μg μl^{-1} in 0.5% (w/v) SDS, 10 mM EDTA pH 8.3, 20 mM Tris-HCl pH 7.5) at 50 °C for 15 min, followed by extraction with phenol:chloroform:isoamyl alcohol and chloroform: isoamyl alcohol. RNA was ethanol precipitated for 4 h or overnight at -20 °C and dissolved in H_2O.

RNA used for RT-PCR was treated with 8 U of DNAse I (RNAse-free, NEB, M0303) for 15 min at 37 °C in a 50 μl reaction. The reaction volume was increased to 100 μl with H_2O and further purified using RNA Clean & Concentrator (Zymo Research, R1013) following the manufacturer's instructions. In brief, 100 μl RNA samples were mixed with 200 μl of RNA Binding Buffer and 300 μl of 100% ethanol. The mix was transferred to the

Zymo-Spin IC Column and centrifuged for 30 sec at 14,000 × g. The column then was washed with 400 µl of RNA Prep Buffer, and once with 700 µl of RNA Wash Buffer and centrifuge for 30 sec at 14,000 × g followed by a final wash used 400 µl RNA Wash Buffer and centrifugation for 1 min at 14,000 × g. The flow-through was discarded in each step. The column was transfer to a new tube, and RNA was eluted in 20 µl of H_2O by centrifugation for 30 sec at 14,000 × g.

III.3.6: Northern blot analysis

DNAse-treated RNA isolated from immunoprecipitation was resuspended in 1× formamide loading buffer, incubated for 2 min at 75°C, and separated on a 4% (v/v) polyacrylamide (29:1) gel containing 8 M urea and transferred to Biodyne nylon membrane (Pall Corporation) at 400 mA for 1 h in 0.5× TBE buffer. RNA was crosslinked to the membrane using 254-nm UV light at 120 mJ/cm^2 in a Stratalinker (Stratagene). Hybridization with radiolabeled probes (10 million counts per minute) were performed in Church–Gilbert buffer at 60 °C with TER1 probe (nucleotides 536–998, labeled with High Prime (Roche) and [α-^{32}P]-dCTP). Blots were washed briefly once with 100 ml of 0.1X SSC, 0.1 % (w/v) SDS and twice with 100 ml of 0.1 SSC, 0.1 % (w/v) SDS for 15 min at 60 °C each. Blots were exposed to PhosphorImager screens and analyzed with a Typhoon 8600 scanner.

III.3.7: RT-qPCR

Purified and DNase-treated RNA was reverse transcribed in 20 µl reactions containing 2.5 µg of total RNA or RNA isolated from the input of immunoprecipitations or ¼ of the RNA isolated from IPs using 5 mg of cell-free extract, 4 µl of 5 × Vilo reaction mix (Life Technologies, 11754), and 2 µl of 2 × Superscript III enzyme blend (Life Technologies, 11754). All the reactions were incubated at 25°C for 10 min, 42°C for 1 hr, and 85°C for 5

min in a thermocycler. RNaseH (5 U, NEB, M0297S) was added followed by incubation at 37 °C for 20 min and 65°C for 20 min. Reactions were then diluted with water 10-fold for total or input RNA and 20-fold for RNA isolated from IPs. Reactions in 384-well microplates contained 2 µl of the diluted cDNA, 5 µl of 2× PerfeCTa SYBR Green FastMix (Quantabio, 733-2086), 2 µl of forward and reverse primers from stocks at 2.5 µM(listed in Table 3.3) and 1 µl of H_2O. The qPCR reactions were performed in technical and biological triplicates in a ViiA 7 Real-Time PCR system (Thermo Fisher Scientific). The qPCR data from total RNA was normalized to three reference genes (*act1*, *his1*, and *snR101*) by the QuantStudio Software V1.2.4 (Thermo Fisher Scientific) while IP RNA was normalized to the input (TER1 all, TER1 precursor). Finally, the relative fold change was calculated by comparing either to an untagged control or to WT as stated in the figure legends.

Table 3.3: Oligonucleotides used for RT-qPCR.

Product	**Primer #**	**Sequence**
TER1 all	BLoli7827	5'-CAGTGTACGTGAGTCTTCTGCCTT-3'
	BLoli7828	5'-CAAAAATTCGTTGTGATCTGACAAGC-3'
TER1 exon 2	BLoli7825	5'-AATTGCGTATTTAGTAAGAACGCG-3'
	BLoli7826	5'-GATTCATCACTTTCTCAAAATTTTGAAACCG-3'
act1	BLoli7829	5'-GGATTCCTACGTTGGTGATGA-3'
	BLoli7830	5'-CGTTGTAGAAAGTGTGATGCC-3'
his1	BLoli7818	5'-CGAAGACGTGCTTCAGCGA-3'
	BLoli7819	5'-TGTCCACCTCGGAATCACTG-3'
snR101	BLoli7866	5'-CGCTCTAGAAATTGGAATGAG-3'
	BLoli7867	5'-TCTTAAAGGTGTGTCTCTCC-3'

Statistical analysis was performed on the log2 transformed data. Each sample was compared with the wildtype using two tailed Dunnett's test from the R package PMCMRplus. Homoscedasticity was tested with Levene's and Bartlett's tests and equal variances for the samples was assumed. In addition, the data were plotted as bar graphs in original scale along

with mean and standard deviation. Significant values (“.”at $p < 0.05$, “*” at $p < 0.01$, “**” $p < 0.001$, and “***” at $p = 0$) are marked above each bar.

III.3.8: Telomerase activity assay

Telomerase was immunoprecipitated on magnetic protein G dynabeads (Invitrogen, 10003D) coated with anti-c-Myc 9E10 (SigmaAldrich, M4439) for Lsm4 or anti-FLAG M2 (Sigma-Aldrich, F3165) for Pof8 as described above. Bead suspension aliquots (10 and 20 µl) were used in telomerase activity assays. Control samples were incubated in 20 µl of TMG(50) plus 20 ng of RNAse A (Invitrogen) for 10 min at 30 °C prior to the telomerase assay. Supernatants were removed and beads resuspended in 10 µl of 50 mM Tris-acetate at 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, dTTP, 2 µM [α32P]-dGTP (500 Ci/mmol), and 5 µM of oligo PBoli871 (5′-GTTACGGTTACAGGTTACG-3′). Reactions were incubated for 90 min at 30 °C and stopped by the addition of proteinase K (2 µg/µl in 0.5% (w/v) SDS, 10 mM EDTA pH 8.3, 20 mM Tris-HCl pH 7.5). Following addition of 1000 counts per minute of a 100-mer oligonucleotide labeled with [γ-^{32}P]-ATP as recovery and loading control reactions were incubated at 42 °C for 15 min. DNA was extracted with phenol:chloroform:isoamyl alcohol (25:24:1, equilibrated with 5× TE) and ethanol precipitated for 4 hours at −20°C. Products were separated by electrophoresis through a 10% (v/v) polyacrylamide (19:1) sequencing gel containing 8 M urea for 1.5 h at 80 W. Gels were dried and exposed to PhosphorImager screens, and analyzed with a Typhoon PhosphorImager.

III.3.9: RNA-seq

Ribosomal RNA was depleted from 2 µg total RNA using Illumina's Ribo-Zero Magnetic Kit for Yeast (Part #15065382 Rev A, November 2014). Library prep was performed with Illumina's TruSeq Stranded Total RNA LT Sample Prep Kit, following the TruSeq Stranded Total RNA Reference Guide (Oct. 2017; Document # 1000000040499v00) and starting with the Elute Fragment Prime High Conc Mix. Libraries were amplified with 8 PCR cycles, profiled with a High Sensitivity DNA Chip on a 2100 Bioanalyzer (Agilent Technologies) and quantified using the Qubit dsDNA HS Assay Kit (Life Technologies). All 12 samples were pooled in equimolar ratio and sequenced on a NextSeq 500 High Output flow cell (1 x 84 cycles, plus 7 cycles for the index read). Read counts per library ranged from 38.36 to 42.88 million. Reads were trimmed and aligned to the *S. pombe* ASM294v2 genome from ENSEMBL using the STAR aligner (v. 2.5.2b)[203] with the following parameters: --alignSJDBoverhangMin 5 --alignSJoverhangMin 10 --twopassMode Basic --alignIntronMin 20 --alignIntronMax 2500. Between 34.55 and 41.55 million reads per library passed filtered and aligned uniquely. DESeq2 (v. 1.22.2) in R was used for differential gene expression analysis[228]. Genes with absolute $\log_2$ fold change >= 1 and an adjusted *p*-value < 0.05 were considered differentially expressed. Coverage plots of RNA-seq data were generated using pyGenomeTracks[229].

III.3.10: Telomere length analysis

DNA preparation and telomere length analysis were performed as described[192]. Cells from 20 ml of stationary culture (~2×10^9 cells) were incubated with 2 ml of Z buffer (50 mM sodium citrate, 50 mM sodium phosphate dibasic, and 40 mM EDTA pH 7.8) plus 0.5 mg/ml Zymolase T100 (US Biological) and 2 mM dithiothreitol (DTT) for 1 hour at 37 °C. Sodium dodecyl sulfate (SDS) was then added to a final concentration of 2% (w/v) and

incubated for 10 min at 65°C. Then 5× TE (50 mM Tris-HCl pH 8.0, 5 mM EDTA) was added to a final volume of 10 ml and proteinase K (Sigma-Aldrich, P2308) was added to a final concentration of 50 μg ml^{-1}. After incubation for 1 h at 50°C, the samples were precipitated with 3 ml of 5 M potassium acetate for 30 min on ice. The precipitates were removed with two rounds of centrifugation at 3200 × g for 10 min at 4°C. The supernatant was collected and mixed with 1 volume of 100% isopropanol for 1 hour on ice followed by centrifugation at 10,500 × g for 10 min at 4 °C. Genomic DNA was resuspended in 5 × TE with 50 μg ml^{-1} RNAse A. Resuspended DNA was then incubated for 1 hour at 37°C followed by two rounds of extraction with phenol: chloroform:isoamyl alcohol (25:24:1, equilibrated with 5× TE) and one round of chloroform:isoamyl alcohol (24:1, equilibrated with 5× TE). DNA was ethanol precipitated and resuspended in 1× TE. DNA concentrations were determined on a Qubit 3.0 instrument using the dsDNA BR Assay Kit (Life Technologies, Q32853) and 750 ng of each sample was digested with 40 U of EcoRI (NEB, R0101L) and 2 μl of 10× EcoRI buffer in a 20 μl reaction for 12 hours. The digested DNA was separated by electrophoresis on a 1% agarose gel in 0.5× TBE (44.5 mM Tris-borate, 1 mM EDTA at pH 8.3) at 120 V for 6 h. Gels were stained with 1 μg ml^{-1} ethidium bromide and visualized with a Typhoon 8600 scanner to confirm digestion of loaded DNA. Gels were then incubated in 0.25 M hydrochloric acid for 10 min followed by 0.5 M sodium hydroxide and 1.5 M sodium chloride buffer for 30 min and 0.5 M Tris-HCl (pH 7.5) and 1.5 M sodium chloride for 30 min at room temperature. DNA was transferred to Amersham Hybond-N+ membrane (GE Healthcare Life Sciences) via capillary blotting. Transferred DNA was crosslinked to the membrane with 120 mJ per cm^2 with 254-nm UV light in a Stratalinker. A telomere specific probe was generated by PCR from pTELO using T3 (5′-ATTAACCCTCACTAAAGGGA-3′) and T7 (5′-TAATACGACTCACTATAGGG-3′) oligos. A probe specific for the *rad16* gene was generated by PCR from wild-type genomic

DNA using primers XWP9 (5′-ATGGTATTTTTTCGCCATT TACTCG-3′) and XWP10 (5′-TAGGCGGATCGTGAAGTTAA-3′). Both probes were labeled by random hexamer labeling with High Prime (Roche, 11585592001) and [α-^{32}P]-dCTP. Hybridizations were carried out with 10 million counts per minute of each probe in Church–Gilbert buffer at 65 °C. Blots were exposed to PhosphorImager screens and visualized with a Typhoon Imager.

III.3.11: Recombinant protein purification

Untagged *pof8* and 6xHis N-terminally tagged *thc1* were cloned into the multicistronic vector pET-His6 (9B) (addgene #48284) and transformed into a *E. coli* Rosetta (DE3) pLysS strain. Cultures (500 ml) were grown at 37 °C until an OD_{600} of ~0.5 when expression was induced by the addition of 1 mM isopropyl β-d-1-thiogalactopyranoside (IPTG), and further incubated for 2 hours at 37 °C. Cultures were then harvested by centrifugation once for 20 min at 4000 × g in a Beckman JA-10 rotor, and cell pellet was resuspended in 10 ml lysis buffer (50 mM NaH_2PO_4 buffer pH 8.0, 300 mM NaCl, 10 mM imidazole). Cells were lysed by sonication with 30% amplitude, 15 sec on, 15 sec off for total of 1 min 30 sec, and lysates were cleared by centrifugation twice for 20 min in a Beckman JS-7.5 at 7500 rpm and JA-17 at 13000 rpm at 4 °C. Cleared lysates were mixed with 0.5 ml Ni-NTA beads (250 µl bed volume, Qiagen) in 15 ml tubes and incubated for 30 min at 4 °C with gentle agitation. The lysate-resin mix was transfer to a pre-cooled column and beads were allowed to settle for 15 min prior to release of flow through. The column was then washed twice with 2.5 ml of wash buffer (50 mM NaH_2PO_4 buffer pH 8.0, 300 mM NaCl, 40 mM imidazole, 10% (v/v) glycerol). The bound proteins were eluted in 1.25 ml elution buffer and collected 10 fractions every ~125 µl volume (50 mM NaH_2PO_4 buffer pH 8.0, 300 mM NaCl, 250 mM imidazole, 10% (V/V) glycerol). Finally, the collected fractions were dialyzed

against 50 mM NaH_2PO_4 pH 8.0, 150 mM NaCl, 0.5 mM EDTA, 1 mM DTT, 20% (v/v) glycerol, snap-frozen, and stored at –80°C.

III.4: Results

III.4.1: Pof8 forms a complex with Thc1 and Bmc1

To understand how Pof8 controls the loading of Lsm2-8 onto TER1, we performed immunoprecipitation (IP) of native protein extracts from strains containing untagged and N-terminal 3xFLAG epitope-tagged Pof8 under the control of its endogenous promoter. Co-purified proteins were first detected on a silver staining gel which showed a specific band pattern on the tagged sample compared to the untagged control (marked with asterisk in Figure 3.1a). Co-precipitants were then identified using multidimensional protein identification technology (MudPIT). 19 proteins with an average distributed Normalized Spectral Abundance Factor (dNSAF) greater than 0.005 in the tagged samples were found to be enriched by more than two-fold (Figure 1b and Table 3.4). Among them were Pof8 itself and Lsm4, a member of the Lsm2-8 complex which had previously been shown to directly interact with Pof8 by crosslinking [227]. Other members of the Lsm2-8 complex, except for Lsm2 and Lsm7, were preferentially co-precipitated with the 3xFLAG-Pof8 but were not enriched under our criteria (Table 3.4). The two most prominent bands enriched in the epitope-tagged Pof8 samples are Pof8 (49 kDa) and the heat-shock protein Ssa2 (70 kDa) based on their molecular weight and dNSAF values. Ssa2 is an abundant chaperon and was also detected in the untagged control IP. We speculate that the enrichment of Ssa2 may be an artifact from the tagging of Pof8 itself, and we did not further investigate. However, we cannot rule out the possibility that Ssa2 is involved in TER1 biogenesis. Two of the proteins uniquely detected in the tagged Pof8 samples got our attention: sequence orphan

SPCC18b5.09c and SPBC2A9.10, a predicted transcriptional and translational regulator of the Bin3 family.

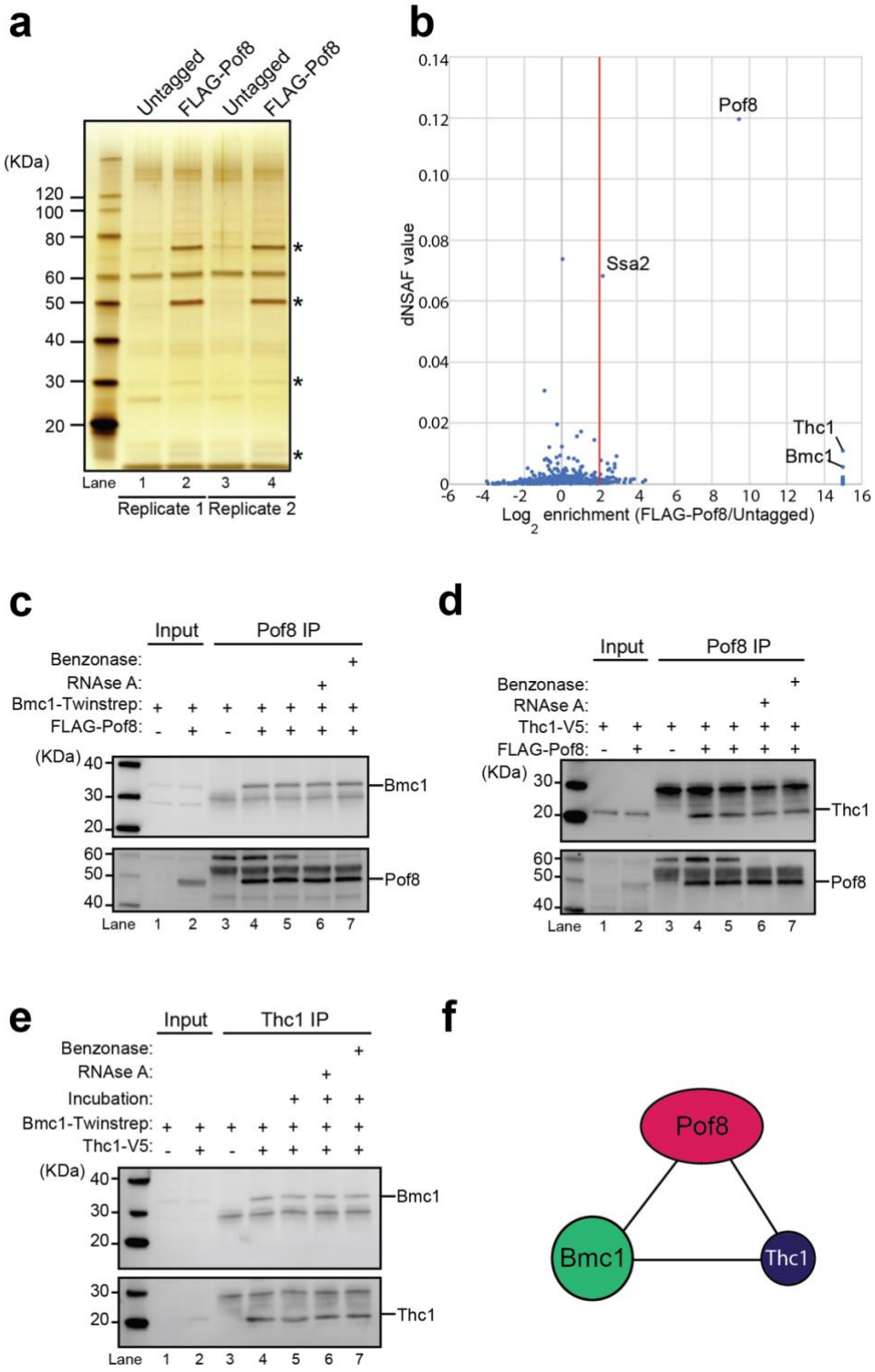

Figure 3.1: Identification of Pof8 interacting proteins. **a)** Silver stained SDS-PAGE of immunoprecipitations (IPs) for 3xFLAG-Pof8 expressed from a plasmid under the control of

its endogenous promoter. The two bands marked with asterisks at ~50 kDa and ~75 kDa represent 3xFLAG-Pof8 and Ssa2, a heat shock protein, espectively. Two biological replicates are shown. Two lower bands are more prominent in lanes 2 and 4 compared to lanes 1 and 3 and are also marked with asterisks. **b)** Scatter plot of proteins enriched in 3xFLAG-Pof8 IP by mass spectrometry. The x-axis shows the average enrichment in the 3xFLAG-Pof8 samples compared to untagged controls. On the y-axis is the distributed Normalized Spectral Abundance Factor (dNSAF) mean value from tagged samples. The vertical red line marks a 4-fold enrichment in the tagged over untagged samples. **c)** Bmc1 coimmunoprecipitates with Pof8. IP with anti-FLAG antibody in extracts from strains harboring Twinstrep-tagged Bmc1 and 3xFLAG-tagged or untagged Pof8. Plus indicates presence of tag, minus presence of the untagged protein. Western blot for input (10%, lanes 1 and 2) and IP (lanes 3-7) probed with anti-strep-tag II antibody for Bmc1-Twinstrep (top panel) and reprobed with anti-FLAG antibody for 3xFLAG-Pof8 (bottom panel). Samples in lanes 5-7 were incubated at 37ºC for 15 min to assess whether the interactions are mediated by nucleic acid. **d)** Thc1 co-immunoprecipitates with Pof8. Co-IP of 2xV5-tagged Thc1 with 3xFLAG-tagged or untagged Pof8 by anti-FLAG antibody. Western blot for input (10%, lanes 1 and 2) and IP (lanes 3-7) probed with α-V5 antibody to detect Thc1-2xV5 (top panel) and reprobed with α-FLAG antibody for 3xFLAG-Pof8 (bottom panel). Samples in lanes 5-7 were treated as in (c). **e)** Bmc1 co-immunoprecipitates with Thc1. Co-IP of Twinstrep-tagged Bmc1 with 2xV5-tagged or untagged Thc1. Western blot for input (10%, lanes 1 and 2) and immunoprecipitated (lanes 3-7) probed with α-strep-tag II antibody for Bmc1-Twinstrep (top panel) and α-V5 antibody for Thc1-2xV5 (bottom panel). Samples in lanes 5-7 were treated the same way as in (c). **f)** Schematic of interactions among the three proteins.

Sequence orphan SPCC18b5.09c encodes for a 13.5 kDa protein whose sequence homology searches were unable to identify a potential homolog. However, a subsequent structural profile search using HHpred [209] revealed a putative motif related to the cap-binding domains of NCBP3 and the poly(A)-specific ribonuclease PARN (Figure 3.2a,b). The strongest similarity between these proteins resides on the highly conserved surface pocket motif (WXDD) [230]. The human nuclear cap binding complex (CBP20 and 80) and PARN have been implicated in the biogenesis of telomerase [95,96,98,99], suggesting that SPCC18b5.09c may carry out an analogous function in fission yeast. Based on our results detailed below, we will subsequently refer to SPCC18b5.09c as Thc1 (Telomerase holoenzyme component 1).

Table 3.4: Proteins enriched in Pof8-FLAG affinity purification with average dNSAF values >0.005 for the tagged Pof8 samples. Proteins are sorted by average fold enrichment in tagged Pof8 over untagged control.

Pof8:Control	NCBI_Gene	Locus	Description	MW
∞	∞	NP_587939.1	sequence orphan	13507
∞		NP_596220.1	Bin3 family, transcriptional and translational regulator (predicted)	30485
711.60	pof8	NP_594264.1	F-box protein Pof8	46808
7.73529412	leu1	NP_595804.2	3-isopropylmalate dehydrogenase Leu1	39733
7.22566372		NP_594058.1	ubiquitin-40S ribosomal protein S31 fusion protein	17215
7.20027435	rvb2	NP_595640.1	AAA family ATPase Rvb2	51562
5.5539629	hsp90	NP_594365.1	Hsp90 chaperone	80596
4.66244783	ssa2	NP_588421.1	heat shock protein Ssa2 (predicted)	70233
4.3323991	mcp60	NP_592894.1	mitochondrial heat shock protein Hsp60/Mcp60	62168
3.41423488	rpp203	NP_594358.1	60S acidic ribosomal protein P2C	11114
2.15492958	rps3	NP_596763.1	40S ribosomal protein S3	27553
2.07858473	tdh1	NP_596154.1	glyceraldehyde-3-phosphate dehydrogenase Tdh1	35870
2.06499759	rpp201	NP_596513.1	60S acidic ribosomal protein P2A	11158
1.81178753	ssc1	NP_593459.1	mitochondrial heat shock protein Hsp70	72977
1.79822926	sks2	NP_595438.1	heat shock protein, ribosome associated molecular chaperone Sks2	67206
1.29202037		NP_592891.1	conserved fungal protein	45724
1.15761234	lsm4	NP_596279.1	U6 snRNP-associated protein Lsm4 (predicted)	13941
1.04284559		NP_596601.2	mRNA processing factor (predicted)	51705
1.02180998	pabp	NP_593377.1	mRNA export shuttling protein	71513
0.96570304	rps2802	NP_588343.1	40S ribosomal protein S28	7664
0.87724193	act1	NP_595618.1	actin Act1	41765
0.8607053		NP_594161.1	sorbose reductase (predicted)	27437
0.82987552		NP_595189.1	homocysteine synthase Met17	46409
0.6725372	ef1a-c	NP_595255.1	translation elongation factor EF-1 alpha Ef1a-c	49675
0.63651325	plr1	NP_594584.1	pyridoxal reductase Plr1	36815
0.62809643	pyk1	NP_594346.1	pyruvate kinase (predicted)	55515
0.60922916	htb1	NP_588181.1	histone H2B Htb1	13819
0.57449627		NP_594844.1	conserved fungal protein	38445
0.5428472		NP_593467.1	transcription factor (predicted)	80990
0.53773336	hrp1	NP_593660.1	ATP-dependent DNA helicase Hrp1	158548

On the other hand, a Clustal alignment[231] confirmed a high similarity between the open reading frame SPBC2A9.10 and proteins in the MePCE / Bin3 family (Figure 3.2c). As the name Bin3 is ambiguous, referring on one hand to Bicoid interacting protein 3, first identified in *Drosophila* and characterized by the presence of an AdoMet binding domain, but also to the unrelated but highly conserved Bridging-Integrator 3 protein, members of the former family are now preferentially referred to as Methyl phosphate capping enzymes

(MePCE). To maintain the link to the founding member of the family, drosophila Bin3 and in adherence to the nomenclature conventions for fission yeast genes, we will therefore refer to SPBC2A9.10 as Bmc1 (Bin3/MePCE 1). MePCE in metazoans is responsible for methylating the γ-phosphate of RNA pol III transcripts such as snRNA U6 and 7SK, a major transcriptional regulator [177]. As TER1 biogenesis has some similarities with the biogenesis of pol III-transcribed U6 snRNA[119], we wondered whether Bmc1 may be a further link between these two non-coding RNAs, despite the absence of a 5' triphosphate cap structure on TER1.

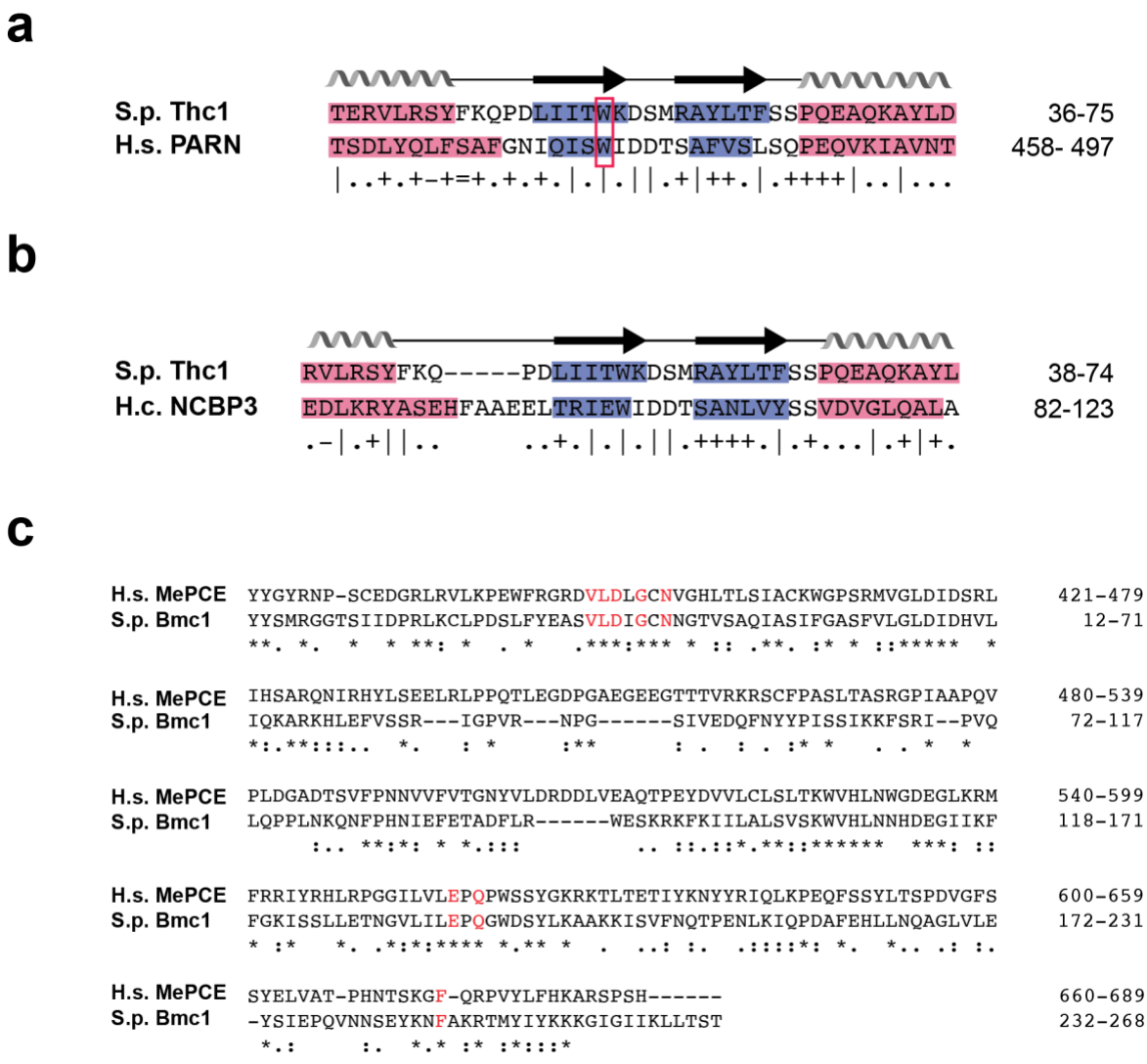

Figure 3.2: Thc1 structural analysis reveals a putative cap-binding domain.Alignment and secondary structure elements of a putative cap binding domain in Thc1 based on HHPred predictions; β-strands are highlighted in blue and α-helices in red. The secondary structure elements predicted for Thc1 are labelled above the alignment. **a)** alignment with the human

(H.s.) poly-adenosine specific ribonuclease (PARN) and **b)** alignment with *Histoplasma capsulatum* (H.c.) Nuclear Cap Binding Protein 3 (NCBP3). The conserved tryptophane that stacks against the guanosine base in the PARN structure is boxed in red. Symbols indicating the quality of the alignment match are depicted below the sequence. "|" very good, "+" good, "." Neutral, "–" bad, and "=" very bad. **c)** SPBC2A9.10 is the likely *S. pombe* ortholog of the methylphosphate capping enzyme in vertebrates. Clustal alignment of the SAM-binding domain of the methylphosphate capping enzyme from human (H.s. MePCE) with *S. pombe* SPBC2A9.10 (Bmc1). Residues in red depict the mutated amino acids used in Figure 3.6. The following symbols are used below the alignment: '*' = fully conserved residue, ':' = amino acid groups of similar properties, '.' = amino acids groups of weakly similar properties.

Next the interaction of Thc1 and Bmc1 with Pof8 was confirmed by co-IP followed by western blot (Figure 3.1c, d). Since Pof8 has been shown to directly bind TER1 [192], we tested whether the Pof8-Thc1 and Pof8-Bmc1 interactions are nucleic acid-dependent. Neither RNase A nor Benzonase treatment affected the co-IP confirming a protein-protein interaction (Figure .31c, d). Similarly, Thc1 was able to coprecipitate Bmc1 in a nucleic acid-independent manner (Figure 3.1e). Altogether the results show that Pof8 forms a strong interaction with at least two other proteins named Thc1 and Bmc1, perhaps forming a stable complex (Figure 3.1f).

III.4.2: Reduced TER1 levels and short telomeres in *thc1Δ* and *bmc1Δ* cells

Since Pof8 is involved in TER1 biogenesis, we then speculated that Pof8-Thc1-Bmc1 forms a stable complex with TER1. To test this, we deleted *thc1*$^{+}$ in haploid cells and one copy of *bmc1*$^{+}$ in diploid cells. Subsequent sporulation gave rise to healthy haploid colonies confirmed by PCR and absence of reads mapping to the respective loci in RNA-seq experiments (Figure 3.3a-f).

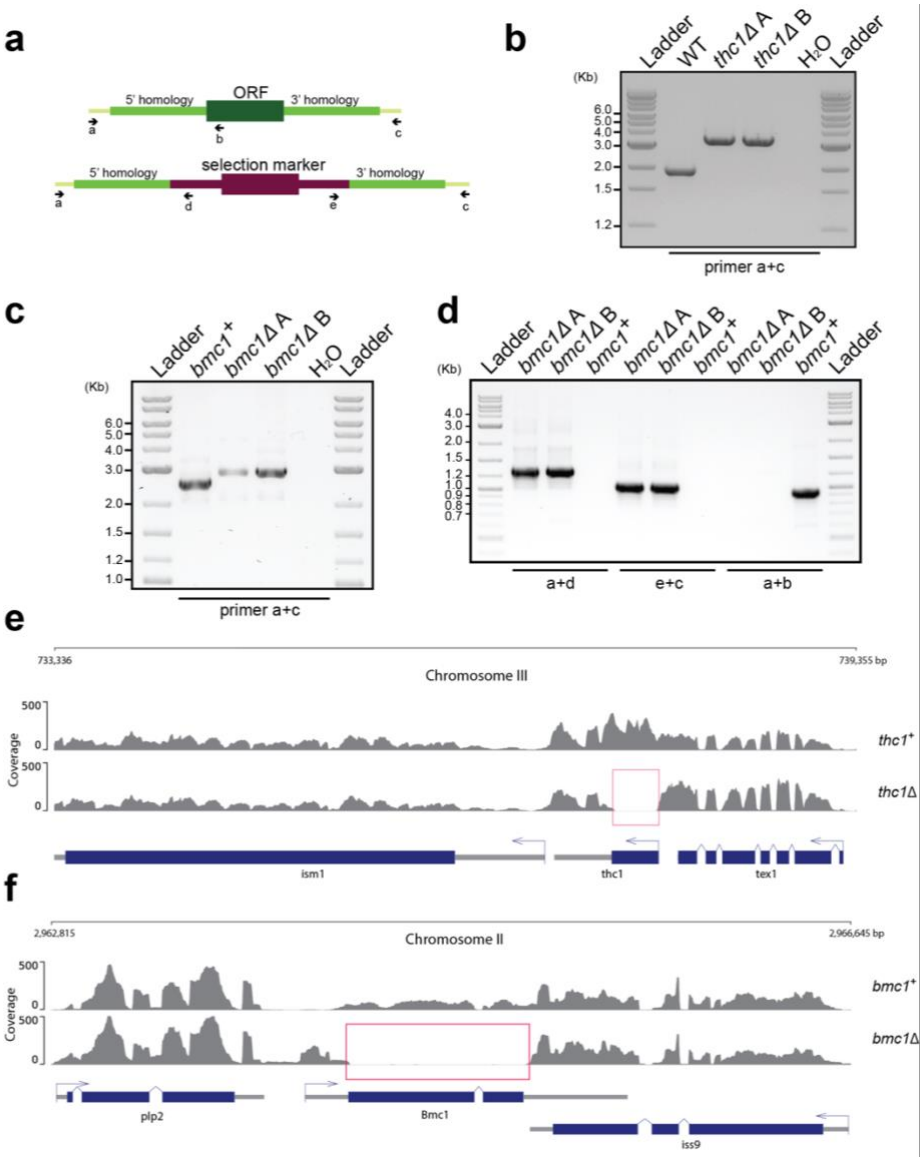

Figure 3.3: Strains verification.**a)** Schematic of primers used to verify the deletion of *thc1* and *bmc1*, respectively. **b)** Diagnostic PCR across the locus using genomic DNA from two *thc1D* isolates and an isogenic control strain. The lane labelled H_2O contained no genomic DNA. **c)** Diagnostic PCR as in b but for the bmc1 locus. **d)** Diagnostic PCR for bmc1 locus in the 5' and 3' regions **e)** Coverage tracks from RNA sequencing analysis for the *thc1* locus in a *thc1D* and *thc1*$^+$ control strain. The absence of reads mapping to the *thc1* open reading frame in the deletion strain serves as an independent confirmation of the deletion. **f)** Coverage tracks from RNA sequencing analysis for the *bmc1* locus in a *bmc1D* and *bmc1*$^+$ control strain.

The steady-state levels of TER1 were measured in these strains by RT-qPCR using primers annealing to the exon 1 and exon 2 (to measure precursor). *pof8Δ* resulted in ~70%

reduction in TER1 mature form levels, in agreement with previous reports [192] while *thc1Δ* and *bmc1Δ* cells had ~70% and ~50% lower TER1 levels, respectively (Figure 3.4a). TER1 precursor was only slightly affected in any of the mutants (Figure 3.4b). We next wondered whether Thc1 and Bmc1 stably associate with TER1 as previously reported for Pof8. TER1 co-immunoprecipitated with V5-tagged Thc1 and FLAG-tagged Pof8 was highly enriched compared to untagged controls (Figure 3.4c) while a lower enrichment was observed with Twinstrep-tagged Bmc1, indicating either a weaker interaction *in vitro* or Bmc1 not being a constitutive member of the telomerase complex.

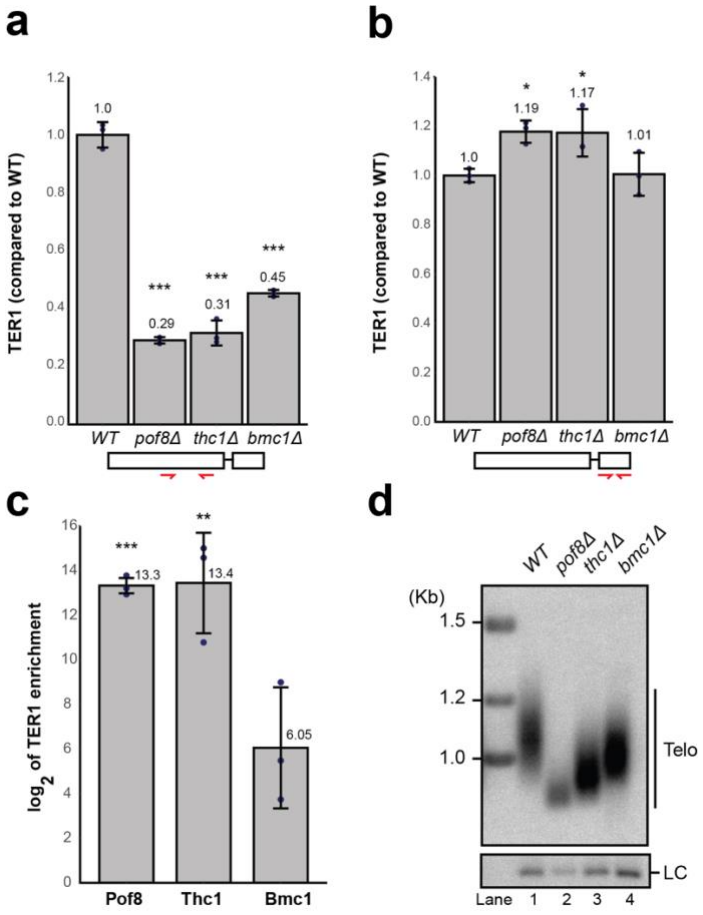

Figure 3.4: Thc1 and Bmc1 stabilize TER1 and are required for wildtype telomere length maintenance.**a)** Quantification of TER1 by RT-qPCR from DNase-treated total RNA extracted from wildtype (WT), *pof8Δ*, *thc1Δ*, and *bmc1Δ* strains. TER1 was normalized to three endogenous control genes (act1, his1, and snR101) and is shown relative to expression in wildtype. The bars represent mean values for biological triplicates (individual values shown as dots on each bar). Error bars represent standard deviation. Three asterisks denote p

value < 0.001, two askerisks, p value < 0.01, one askerisks, p value < 0.05. The schematic on the bottom indicates the position of the primers (red) used for TER1 amplification. **b)** RT-qPCR quantifications as in (a) but using a primer pair in exon 2 of TER1, which is present in the precursor but not in the mature form of telomerase RNA. **c)** RT-qPCR quantification of TER1 recovered from immunoprecipitations for 3xFLAG-Pof8, Thc1-2xV5, and Bmc1-Twinstrep normalized to untagged controls. The bars show the mean values of enrichment of TER1 RNA relative to untagged control after normalization to RNA levels in the input using the same primer sets as in (a). Biological triplicates shown as dots on each bar. Error bars represent standard deviation. **d)** Telomeric Southern blot using genomic DNA from WT, *pof8Δ, thc1Δ, and bmc1Δ* strains. Genomic DNA was digested with EcoRI, separated by gel electrophoresis, transferred to a membrane and probed for telomeric repeats and the *rad16*$^{+}$ locus as loading control (LC).

Since telomere length regulation is sensitive to any change in TER1 levels as shown in *pof8Δ*, we measured telomere length in the deletion of the other two genes. Telomeres were approximately shorter than wildtype in *thc1Δ* cells and in a lesser extent *bmc1Δ* cells (Figure 3.4d), and stable through successive re-streaks (Figure 3.5). Overall, these results confirm that Thc1 and Bmc1 are involved in telomerase biogenesis or quality control. These interactions may be the reminiscence of the human LARP7 which forms a stable complex with MePCE and the 7SK non-coding RNA in metazoans, although the 7SK RNA was unidentified in fission yeast, [192,232–234].

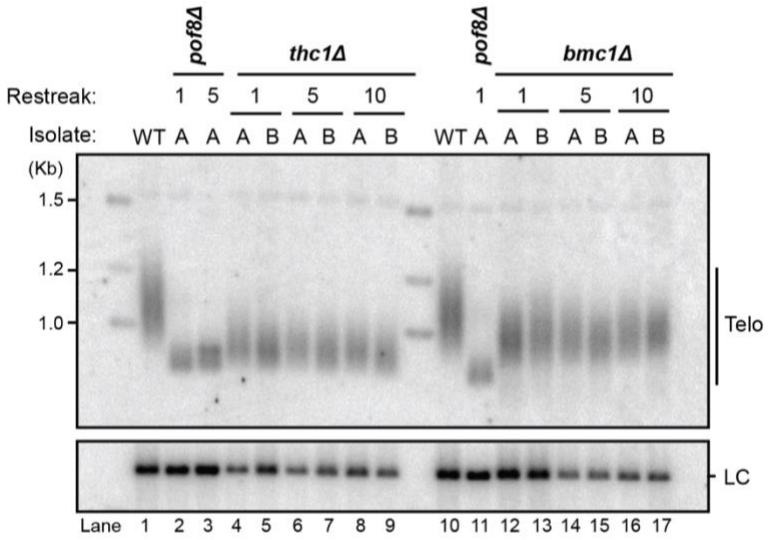

Figure 3.5: Stable but short telomeres in *thc1* and *bmc1* deletion strains.Telomeric Southern blot comparing telomere length in wildtype (WT) and *thc1Δ* or *bmc1Δ* cells. Two independent isolates of each deletion (A and B) were restreaked ten times in series and telomere length was analyzed (one restreak equals 20-25 generations). A *pof8Δ* strain was included as control. The *rad16*$^{+}$ locus was probed to assess integrity and equal loading of the genomic DNA samples.

III.4.3: Bmc1 function in telomerase biogenesis is independent of its catalytic activity.

Metazoan 7SK RNA stability depends on the binding of LARP7 at the 3' end and MePCE at the 5' end. MePCE caps the RNA by transferring a methyl group to the γ-phosphate [176,235]. In the case of Bmc1, it is unclear how TER1 with its trimethylguanosine (TMG) cap could be further methylated. Curiously, structural and biochemical analyses of MePCE showed a capping-independent function where MePCE stably binds to the 5' even with a TMG capped RNA [176,179,236]. This led us to speculate that the methyltransferase activity of Bmc1 is dispensable for its function on TER1. Since sequence alignments showed that the amino acids involved in AdoMet binding and methyltransferase activity are conserved in fission yeast, we generated several mutants designed to abolish the capping activity[176,179,232]. Out of these mutants depicted in red (Figure 3.2c), only VLD39AAA was approximately 3-fold reduced in the protein level (Figure 3.6a). Using telomere length as the final readout of telomerase function *in vivo*, all the catalytic mutants rescued the telomere length phenotype caused by *bmc1* deletion except for the mutant VLD39-41AAA (Figure 3.6b). It is likely that mutation of the VLD motif causes structural changes that compromises more than just the catalytic activity which is reflected in its lower stability. In summary, the catalytic activity of Bmc1 is not required for its function in telomerase biogenesis. However, a more thorough separation of function mutations to distinguish between methyltransferase and RNA binding activities, and biochemical functional assays will be necessary to further dissect Bmc1 function on TER1 biogenesis.

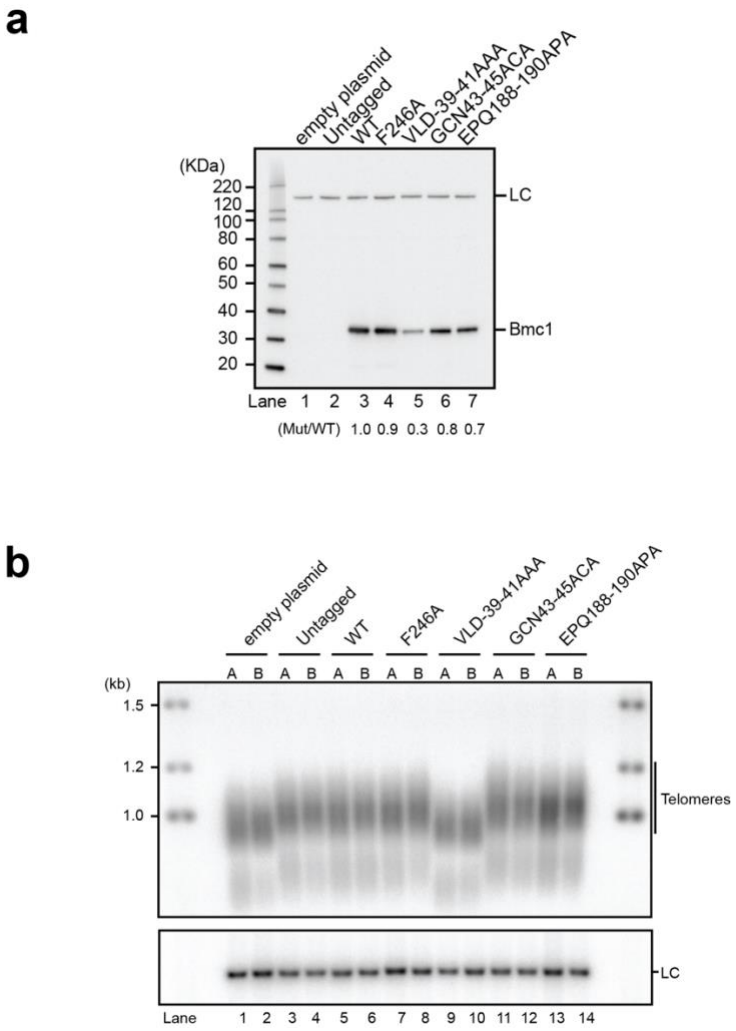

Figure 3.6: The function of Bmc1 in telomerase biogenesis is independent of its catalytic activity. **a)** Western blot analysis of Twinstrep-tagged WT and mutant Bmc1 designed to affect the methyltransferase activity using α-Strep-tag II antibody. All versions of Bmc1 were expressed from a plasmid under the control of the endogenous promoter in a *Bmc1Δ* background. A non-specific band recognized by α-Strep-tag II in the absence of the epitope-tag (lane 2) was used as intrinsic loading control (LC). Quantification of mutants compared with WT are shown below the lane numbers. **b)** Telomeric Southern blot for Bmc1 mutants. Two independent isolates of each strain are shown. A probe against the *rad16*$^{+}$ locus was used as a loading control (LC). Lane numbers are indicated below the blot.

III.4.4: Pof8, Thc1 and Bmc1 function in the same pathway.

Since the phenotype penetrance is not equal among the deletion of each of the three genes, we asked whether they have distinct roles during TER1 biogenesis. We speculated that

measuring the TER1 levels by RT-qPCR and telomere length by Southern blot in double and triple deletion would result in an additive phenotype if they function in different pathways. Neither TER1 levels (Figure 3.7a) nor telomere length (Figure 3.7b) was further reduced in any deletion combination compared to single *pof8* deletion. These results suggest that *pof8* is epistatic to *thc1* and *bmc1*.

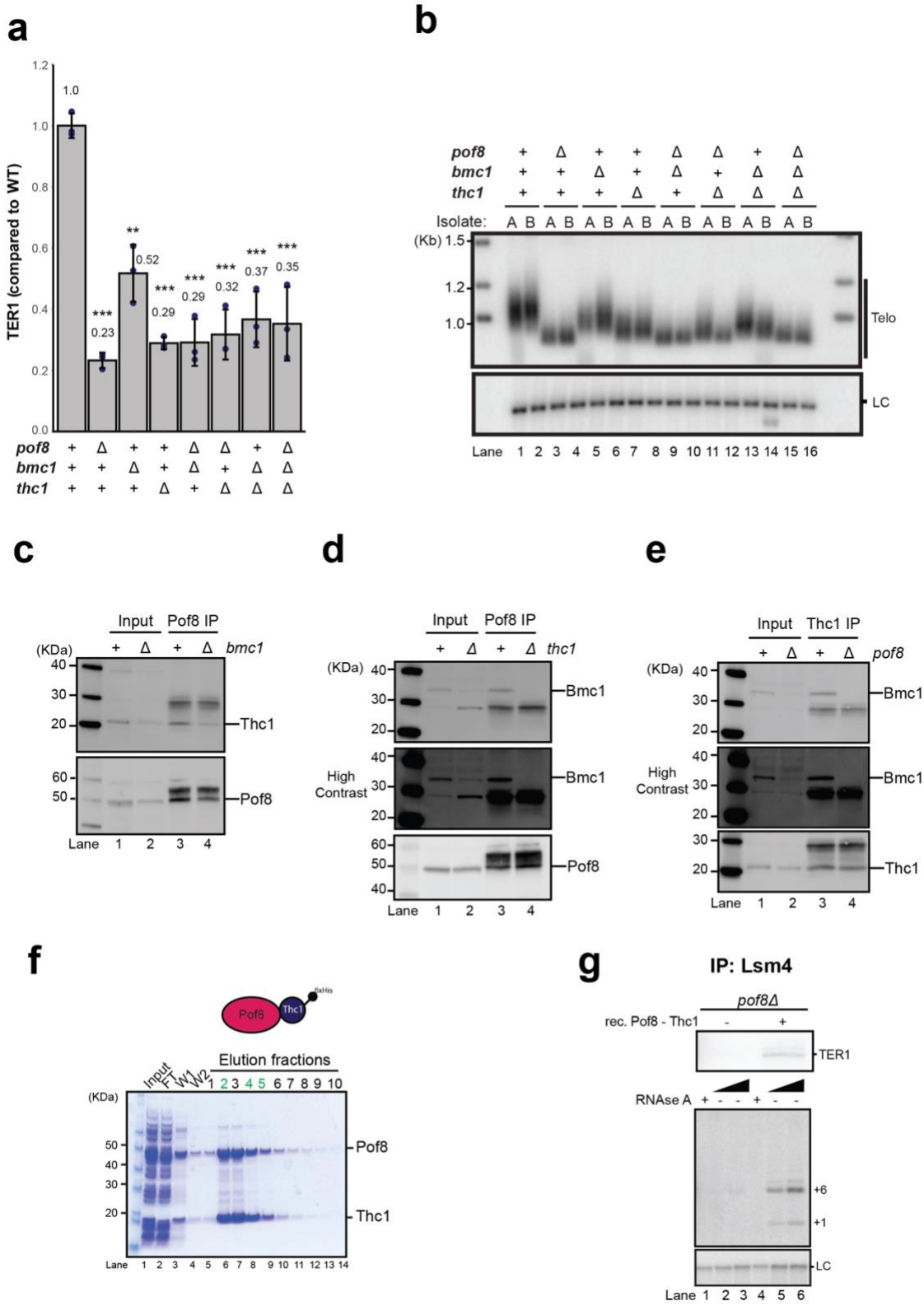

Figure 3.7: Functional interactions among Pof8, Thc1, and Bmc1.**a)** Effect of single, double, and triple deletions of *pof8*, *thc1*, and *bmc1* genes on TER1 levels. RT-qPCR data normalized to *act1*, *his1*, and *snR101* genes and presented relative to expression in WT. Mean fold change values were calculated from biological replicates (shown as dots on the bars). Error

bars show standard deviation. **b)** Telomeric Southern blot of the strains used in (a). Two independent isolates of each deletion are shown. The *rad16*$^{+}$ locus was used as a loading control (LC). **c)** Western blot analysis of co-IP using α-FLAG antibody in strains containing 3xFLAG-Pof8 and Thc1-2xV5 in *bmc1*$^{+}$ or *bmc1Δ* backgrounds. Bmc1 stabilizes Thc1 protein, but is not essential for Thc1 to interacts with Pof8. The blot was probed with α-V5 antibody for Thc1-2xV5 (*top panel*) and reprobed with α-FLAG antibody for 3xFLAG-Pof8 (*bottom panel*). Input represents 10% of IP. **d)** Western blot analysis of co-IP using α-FLAG antibody in strains containing 3xFLAG-Pof8 and Bmc1-Twinstrep in *thc1*$^{+}$ or *thc1Δ* backgrounds. Thc1 stabilizes Bmc1 protein and is required for the Bmc1-Pof8 interaction. The top panel is probed with α-Strep-tag II antibody for Bmc1-Twinstrep. A higher contrast of the same blot is shown in the middle panel to visualize the low levels of Bmc1-Twinstrep in the input and the absence of Bmc1 in the IP (lanes 2 and 4). The bottom panel shows the reprobe with α-FLAG antibody for 3xFLAG-Pof8. **e)** Western blot analysis of co-immunoprecipitation using α-Strep-tag II antibody in strains containing Bmc1-Twinstrep and Thc1-2xV5 in *pof8*$^{+}$ or *pof8Δ* backgrounds. Pof8 stabilizes Bmc1 but not Thc1 and is required for the Bmc1-Thc1 interaction. The top panel is probed with α-Strep-tag II antibody for Bmc1-Twinstrep. A higher contrast of the same blot was shown in the middle panel to visualize the low levels of Bmc1-Twinstrep in the input and the absence of Bmc1 in the IP (lanes 2 and 4). The bottom panel is the reprobe with α-V5 antibody for Thc1-2xV5. **f)** Coomassie stained SDS-PAGE gel of immobilized-metal affinity chromatography (IMAC) purified recombinant 6xHis-Thc1 and untagged Pof8 co-expressed from a multicistronic plasmid in *E. coli*. The gel shows the input (lane 1), flow-through (FT, lane 2), two washes (W, lanes 3 and 4), and 10 fractions of eluate (lanes 5-14). The fractions labeled in green were pooled, quantified and used in g. **g)** (*top panel*) Northern blot analysis of TER1 following immunoprecipitation of cMyc-tagged Lsm4. Extracts from *pof8* deficient cells were mixed with buffer or 50 nM recombinant Pof8 / 6xHis-Thc1 complex (*bottom panel*). Telomerase activity assay with IP fractions from the same IP as in the top panel. RNAse A was added to beads before performing the activity assay in lanes 1 and 4 as controls. The bands corresponding to +1 and +4 nucleotide addition products are indicated on the right. A ^{32}P-labeled 100-mer oligonucleotide was used as LC.

III.3.5: Deletion of *pof8* affects stability and interaction between Thc1 and Bmc1

To further understand the different phenotype for each single deletion, we assessed the stability of the other subunits and integrity of the complex in the absence of each protein. When native protein extract was prepared from *bmc1Δ* cells, the interaction between Pof8 and Thc1 was conserved as well as the stability (Figure 3.7c). On the other hand, deletion of *pof8* or *thc1* resulted in a decrease in Bmc1 in the input and non-detectable signal for coimmunoprecipitation of Bmc1 in both Pof8 and Thc1 IPs (Figure 3.7d, e). Hence the absence of Pof8 or Thc1 affected Bmc1 stability and even further its interaction with Thc1 and Pof8, respectively. Further supporting these results, co-expressing 6xHis-Thc1 and

untagged Pof8 in *E. coli* was enough to form a Thc1–Pof8 subcomplex (Figure 3.7f). Addition of recombinant Thc1–Pof8 subcomplex to native protein extract from *pof8Δ* cells increased the recovery of TER1 and telomerase activity from immunoprecipitation of epitope-tagged Lsm4 (Figure 3.7g). The results suggest physical and functional interactions within the complex, in which the stability of individual factors and the integrity of the complex depends on other members. In addition, the different phenotype penetrance in the single deletion may be due to the loss of Pof8 causing a defect in the complex formation which is not the case for a *bmc1* deletion where Thc1 – Pof8 interaction is unaffected.

III.4.6: Different effects on the Sm to Lsm2-8 switch

TER1 precursor is bound by Sm complex which is replaced by the Lsm2-8 complex after spliceosomal cleavage produces the mature telomerase RNA [119]. As shown previously, Pof8 helps during this transition via direct protein-protein interaction with the Lsm2-8 complex, and deletion of *pof8* resulted in a drastic reduction in the amount of TER1 immunoprecipitated with Lsm4 [192,223,224,227]. To dissect the roles of Thc1 and Bmc1, the amount of TER1 co-immunoprecipitated with Lsm4 and Smb1 subunits was measured in the single deletion background by RT-qPCR. As shown previously, *pof8Δ* had a prominent reduction of TER1 co-precipitated with Lsm4 while an increase with Smb1 (Figure 3.8a, d). In stark contrast, deletion of *thc1* affected TER1 binding to Lsm4 only by two-fold, and an increase to Smb1 (Figure 3.8a , b). *bmc1Δ*, interestingly, caused slight increase of TER1 association with both Lsm and Sm complexes (Figure 3.8a, d). The relatively similar amount of TER1 associated with Lsm2-8 complex in *bmc1Δ* may explain the milder effect on the TER1 steady-state levels and telomere length regulation in comparison to *pof8* or *thc1* deletions (Figure 3.4a, c). Even though Sm binds to TER1 precursor, the complex temporarily remains bound to TER1 mature form after spliceosomal cleavage [119,121]. To

further dissect the identity of the TER1 species more stably bound to the Sm complex in the single deletions, TER1 precursor associated with Smb1 were quantified unable to detect any differences among the different background strain (Figure 3.8c). These observations support the notion that the switch from Sm-bound to Lsm-bound form which occurs after spliceosomal cleavage is an active process and that TER1 remains associated with the Sm complex for longer in the absence of Pof8-Thc1-Bmc1 binding and Lsm2-8 complex replacement.

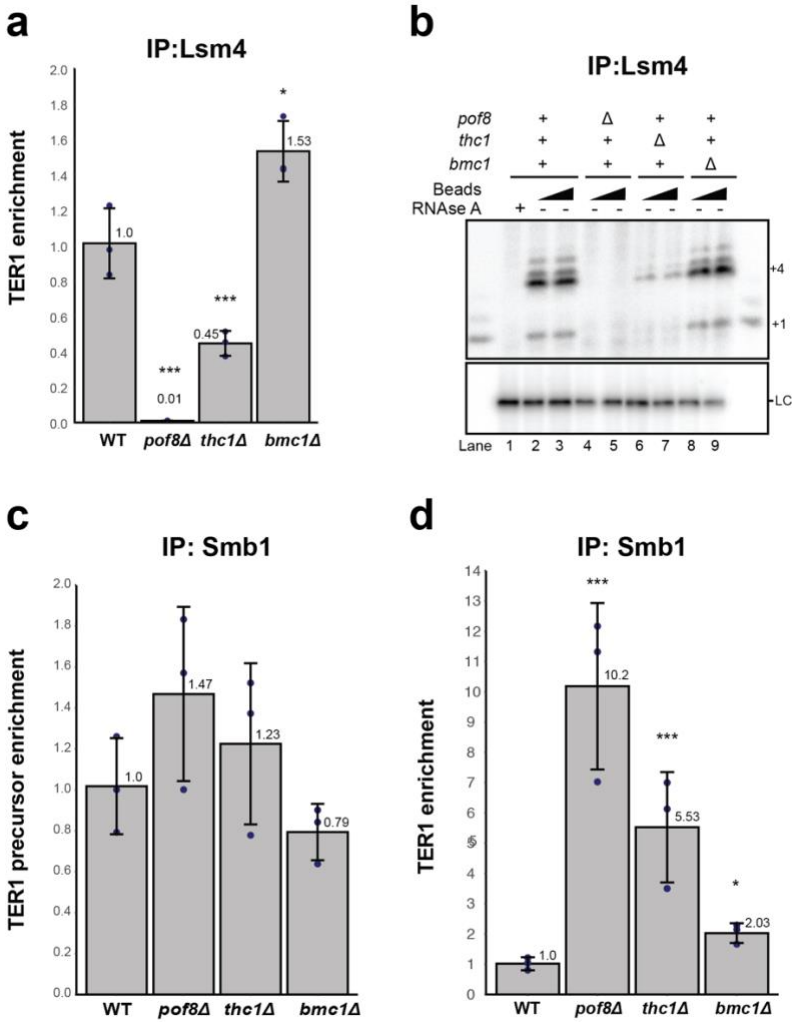

Figure 3.8: Pof8 and Thc1 - but not Bmc1 - are required for loading the Lsm2-8 complex onto TER1. **a)** RT-qPCR analysis of TER1 from immunoprecipitation of cMyc-tagged Lsm4 from extracts prepared from WT, *pof8Δ*, *thc1Δ*, and *bmc1Δ* cells, respectively. The bars represent mean fold enrichment of TER1 relative to WT after normalization to TER1 levels in the input. Error bars represent standard deviation of biological triplicates (individual values

shown as dots). **b)** Telomerase activity assay using the same Lsm4 IP samples as in (a) shows that the activity is compromise din the absence of Pof8 or Thc1. **c)** Association of Sm proteins with the TER1 precursor is not compromised in the absence of Pof8, Thc1 or Bmc1. RT-qPCR analysis of TER1 precursor following IP of cMyc-tagged Smb1 from WT, *pof8Δ*, *thc1Δ*, and *bmc1Δ* cells. **d)** The Sm complex remains associated with TER1 following spliceosomal cleavage in the absence of Pof8, Thc1 or Bmc1. RT-qPCR analysis of TER1 following IP of cMyc-tagged Smb1 from WT, *pof8Δ*, *thc1Δ*, and *bmc1Δ* cells.

Finally, to confirm the RT-qPCR results, telomerase activity levels were tested by primer extension from cMyc epitope-tagged Lsm4 immunoprecipitation in the same strains. Interestingly, despite the modest effect on TER1–Lsm association, loss of Thc1 had 15-fold reduction in telomerase activity associated with Lsm4 (Figure 3.8b). Under the same conditions, the activity was undetectable in the absence of Pof8 and was unchanged in the absence of Bmc1. Taking all together, we propose a model where Pof8 and Thc1 are mainly involved in regulating the sequential binding of Sm to Lsm complexes while Bmc1 may optimize Pof8 function, perhaps by assessing the processing state at the 5' end of TER1.

III.4.7: Thc1 and Bmc1 ensure stable association of Pof8 with telomerase

Since the deletion of *pof8* consistently shows the strongest effect on Lsm recruitment, telomerase activity, and telomere length maintenance, we assessed the extent to which Thc1 and Bmc1 affect Pof8 stability or Pof8–TER1 association. The stability of Pof8 protein in *thc1Δ*, *bmc1Δ* and double deletions measured by western blot showed insignificant changes in the different deletion backgrounds (Figure 3.9a). We next measured the affinity of Pof8 for TER1 by co-immunoprecipitation in the same strains. A reduction greater than 3-fold in Pof8 binding to TER1 was obtained in the deletion of *thc1* and *bmc1*compared to WT cells (Figure 3.9b). A five-fold reduction was observed following the deletion of both *thc1* and *bmc1* demonstrating that the stability of the Pof8-TER1 complex is enhanced by the presence of Thc1 and Bmc1 proteins.

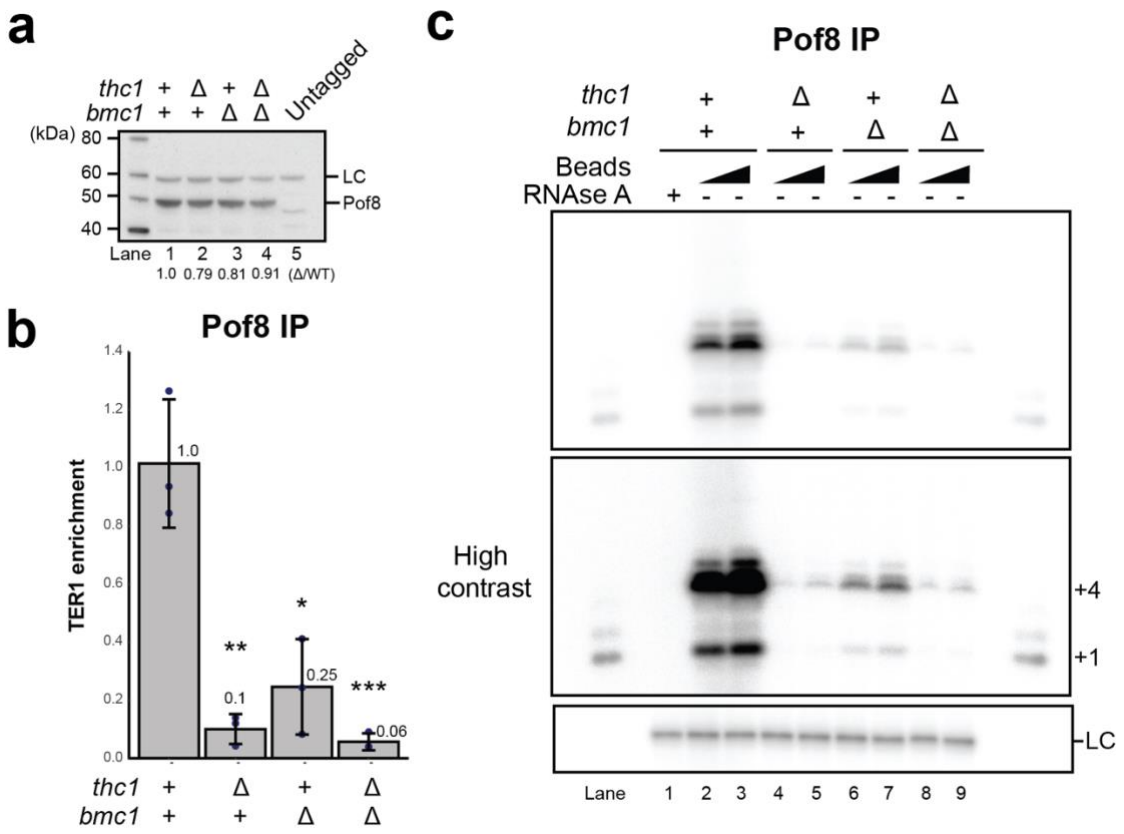

Figure 3.9: The absence of Bmc1 and/or Thc1 does not affect Pof8 protein stability, but weakens the Pof8 – TER1 interaction. **a)** Western blot analysis of FLAG-tagged Pof8 in WT, *thc1*$^{-}$, *bmc1*$^{-}$, and double deletion backgrounds using α-FLAG antibody. Extract from cells harboring untagged Pof8 was used as the negative control (lane 5). A non-specific band cross-reacting with α-FLAG in all lanes was used as intrinsic loading control (LC) for normalization. Quantification of FLAG-tagged Pof8 in single and double deletion backgrounds relative to WT shown in lane 1. **b)** RT-qPCR analysis of TER1 following IP of FLAG-tagged Pof8 from *Thc1*, *Bmc1*, and combined deletion backgrounds relative to WT background. Three biological replicates (shown as dots on the bars) were used and error bars represent standard deviation. **c)** Telomerase activity assay from Pof8 IP samples of the same strains used in (b).

Previously Pof8 was shown to be a constitutive member of telomerase holoenzyme and associated with strong telomerase activity *in vitro* [192]. Since Pof8 shows less association with TER1 in the absence of Thc1 and Bmc1, telomerase activity was measured by pulling down Pof8 in the same background. The activity was reduced by more than 100-fold in *thc1 Δ*, 28-fold in *bmc1Δ*, and over 300-fold in the double deletion background (Figure 3.9c). It is

worth notice that the levels of telomerase activity associated with Pof8 in *bmc1Δ* cells were also significantly lower compared to *bmc1+* while the Lsm4 associated activity did not change in the absence of *bmc1* (Figure 3.8b). The above results suggest a model where the affinity of Pof8 to TER1 is dependent on the complex.

Our current model proposes Pof8 function in a complex with at least two other members to ensure the stepwise biogenesis and quality of TER1. As a complex, Pof8 directly promotes the assembly of Lsm-TER1 via direct protein-protein interaction[227], while Bmc1 and Thc1 fine-tune this process. Bmc1 and especially Thc1 enhance Pof8 loading or stability onto TER1 (Figure 3.10) probably by binding to the 5' end and integrating the biogenesis and quality control at both ends of TER1.

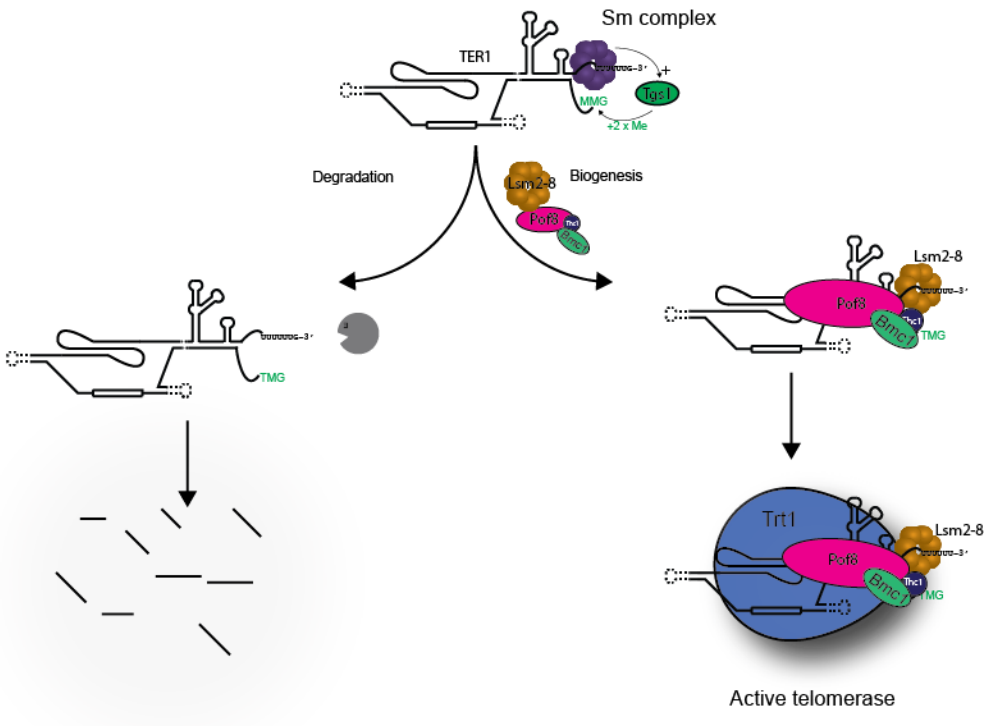

Figure 3.10: Proposed model of the role of the Pof8 complex during the stepwise assembly of telomerase.After spliceosomal cleavage, the Sm complex (purple) facilitates cap hypermethylation by Tgs1 and is subsequently replaced by the Lsm2-8 complex (brown). The Pof8 – Thc1 – Bmc1 complex functions during the Sm to Lsm transition. Pof8 (pink) promotes the loading of Lsm2-8 onto TER1, while Bmc1 (green) and Thc1 (blue) stabilize

Pof8 binding through protein-protein and likely through protein-RNA interactions. In the absence of Pof8 on TER1, the Lsm2-8 complex cannot load leading to TER1 degradation.

III.4.8: Deletion of *thc1* affects the expression of several other ncRNAs

In light of the similarity of Thc1 and Bmc1 with general RNA processing factors, we asked whether the expression of RNAs other than TER1 is also affected by the deletion of either gene. For the deletion of *thc1*, 77 genes met the criteria of more than a two-fold change in expression levels with an associated p-value <0.05 (Figure 3.11a). Next to TER1 and thc1, 26 genes, including 11 non-coding RNAs were downregulated (Table 3.5a). We also identified 49 loci that were upregulated, the majority of which were non protein-coding RNAs (Table 3.5b).

Table 3.5: Genes affected in expression level by the deletion of *thc1*.**a)** Genes with RNA levels decreased by more than 2-fold in the absence of *thc1*$^+$ based on DESeq2 analysis of RNA samples from triplicate cultures of *thc1*$^+$ and *thc1*$^+\Delta$ cells. **b)** Genes with increased expression by more than 2-fold by the same analysis as in (a).

a

gene_ID	external_gene_ID	log2_FC	adj. p-value	gene_biotype
SPNCRNA.863	SPNCRNA.863	-2.0640	9.6E-141	ncRNA
SPNCRNA.1344	SPBC1271.08c-antisense-1	-1.9046	6.7E-34	ncRNA
SPCC569.02c	SPCC569.02c	-1.8719	1.2E-12	protein_coding
SPBCPT2R1.07c	SPBCPT2R1.07c	-1.6810	2.5E-02	pseudogene
SPNCRNA.742	SPAC9.08c-antisense-1	-1.6559	3.5E-16	ncRNA
SPBP4G3.02	pho1	-1.5304	2.7E-236	protein_coding
SPNCRNA.214	ter1	-1.5083	7.8E-128	ncRNA
SPNCRNA.923	SPNCRNA.923	-1.5039	1.5E-05	ncRNA
SPNCRNA.130	omt3	-1.4787	4.7E-06	ncRNA
SPNCRNA.1255	SPNCRNA.1255	-1.4776	6.9E-04	ncRNA
SPAC821.10c	sod1	-1.4115	1.0E-138	protein_coding
SPCC70.08c	SPCC70.08c	-1.4017	1.2E-42	protein_coding
SPCC663.14c	trp663	-1.3878	4.1E-02	protein_coding
SPCC1393.14	ten1	-1.3405	5.2E-07	protein_coding
SPNCRNA.1605	SPNCRNA.1605	-1.3271	3.2E-07	ncRNA
SPCC18B5.09c	thc1	-1.3033	2.0E-76	protein_coding

SPNCRNA.888	end4-antisense-1	-1.2969	8.5E-23	ncRNA
SPAC186.06	SPAC186.06	-1.2292	5.1E-09	protein_coding
SPNCRNA.1626	SPNCRNA.1626	-1.2100	4.9E-44	ncRNA
SPBC1271.07c	SPBC1271.07c	-1.1767	4.9E-24	protein_coding
SPBPB2B2.01	SPBPB2B2.01	-1.1699	1.4E-24	protein_coding
SPBPJ4664.03	mfm3	-1.1656	4.3E-03	protein_coding
SPNCRNA.1340	SPNCRNA.1340	-1.1631	5.2E-15	ncRNA
SPAC8E11.12	SPAC8E11.12	-1.0941	1.5E-02	protein_coding
SPAC27D7.03c	mei2	-1.0916	3.0E-17	protein_coding
SPNCRNA.944	SPNCRNA.944	-1.0809	3.6E-05	ncRNA
SPBPB21E7.07	aes1	-1.0504	6.4E-25	protein_coding
SPAC1F7.08	fio1	-1.0350	3.4E-02	protein_coding

b

gene_ID	external_gene_ID	log2_FC	adj. p-value	gene_biotype
SPNCRNA.466	SPNG2151	4.2720	2.13E-02	ncRNA
SPAC977.05c	SPAC977.05c	4.1810	5.59E-03	protein_coding
SPBC1348.06c	SPBC1348.06c	4.0169	8.64E-03	protein_coding
SPMITTRNASER.01	SPMITTRNASER.01	2.5719	1.10E-05	tRNA
SPNCRNA.287	SPNCRNA.1300,SPNG1093	2.5602	2.39E-12	ncRNA
SPNCRNA.100	SPNCRNA.100	2.5543	7.41E-09	ncRNA
SPMITTRNALYS.01	SPMITTRNALYS.01	2.3763	1.32E-11	tRNA
SPNCRNA.1656	nup120-antisense-1	2.3445	1.38E-05	ncRNA
SPCC1393.10	ctr4	2.2131	3.19E-201	protein_coding
SPAC1F8.04c	SPAC1F8.04c	2.1839	3.42E-15	protein_coding
SPBP4H10.09	rsv1	1.9912	6.16E-04	protein_coding
SPNCRNA.1282	SPNCRNA.1282	1.8133	9.99E-11	ncRNA
SPNCRNA.1677	doa10-antisense-1	1.7962	8.50E-03	ncRNA
SPMITTRNAMET.01	SPMITTRNAMET.01	1.7436	1.54E-03	tRNA
SPCC1235.14	ght5	1.7393	8.95E-134	protein_coding
SPNCRNA.601	SPNCRNA.601	1.6964	1.15E-07	ncRNA
SPAC1F8.03c	str3	1.6478	2.32E-07	protein_coding
SPAC1142.05	ctr5	1.6164	3.39E-148	protein_coding
SPBPB2B2.05	SPBPB2B2.05	1.5930	1.59E-03	protein_coding
SPNCRNA.1558	qcr10-antisense-1	1.4771	4.57E-08	ncRNA
SPRRNA.01	21S_rRNA	1.4280	7.37E-12	rRNA
SPNCRNA.1307	SPBPB10D8.02c-antisense-1	1.4233	9.50E-19	ncRNA
SPNCRNA.945	SPNCRNA.945	1.4199	5.60E-07	ncRNA
SPAC1F8.06	pfl8	1.3844	2.63E-05	protein_coding
SPNCRNA.983	tsn1-antisense-1	1.3631	6.86E-03	ncRNA
SPRRNA.02	15S_rRNA	1.3599	1.52E-10	rRNA
SPBP4H10.10	rbd3	1.3464	1.10E-09	protein_coding

SPCC330.06c	pmp20	1.3422	2.45E-49	protein_coding
SPBC11B10.02c	his3	1.3419	2.23E-54	protein_coding
SPAC4G8.03c	puf5	1.2856	8.00E-49	protein_coding
SPBC1778.04	spo6	1.2527	2.95E-07	protein_coding
SPNCRNA.634	shk2-antisense-1	1.2495	5.43E-64	ncRNA
SPNCRNA.1557	lid2-antisense-1	1.2051	2.49E-03	ncRNA
SPAC56F8.15	SPAC56F8.15	1.2029	8.85E-28	protein_coding
SPNCRNA.1438	cdc2-antisense-1	1.1829	6.63E-03	ncRNA
SPBC1289.14	SPBC8E4.10c	1.1804	2.69E-03	protein_coding
SPNCRNA.1409	SPNCRNA.1409	1.1604	6.19E-07	ncRNA
SPNCRNA.33	prl33	1.1499	3.74E-02	ncRNA
SPCC1739.08c	SPCC1739.08c	1.1300	2.14E-04	protein_coding
SPBC1D7.02c	scr1	1.1176	2.29E-42	protein_coding
SPNCRNA.1330	SPNCRNA.1330	1.1154	4.05E-02	ncRNA
SPBC23G7.08c	rga7	1.1060	7.65E-27	protein_coding
SPNCRNA.892	tif211-antisense-1	1.1049	2.85E-04	ncRNA
SPNCRNA.953	SPNCRNA.953	1.0861	1.32E-35	ncRNA
SPNCRNA.1031	wsp1-antisense-1	1.0573	1.98E-04	ncRNA
SPNCRNA.875	gcv1-antisense-1	1.0473	2.35E-08	ncRNA
SPNCRNA.1378	cbp6-antisense-1	1.0417	6.87E-09	ncRNA
SPNCRNA.774	SPNCRNA.774	1.0215	1.55E-02	ncRNA
SPCC794.01c	gcd1	1.0186	1.95E-09	protein_coding

It, therefore, appears that Thc1 is involved in the biogenesis or regulation of more RNAs than just TER1. Far fewer changes in gene expression were observed in the deletion of *bmc1*, with no gene meeting the criteria described above (Figure 3.11b). In light of U6 snRNA being a target of vertebrate MePCE, we specifically examined this RNA but found no significant change in expression in either deletion. However, it must be noted that the knockdown of MePCE in human cells also did not reduce the steady-state levels of U6 snRNA[37]. While changes in expression levels are indicative of a specific RNA being a target for modification, RNA stability is not necessarily affected by the absence of the methyl group. It therefore will require further studies to assess whether Bmc1 functions specifically as a telomerase biogenesis factor or has other targets in fission yeast.

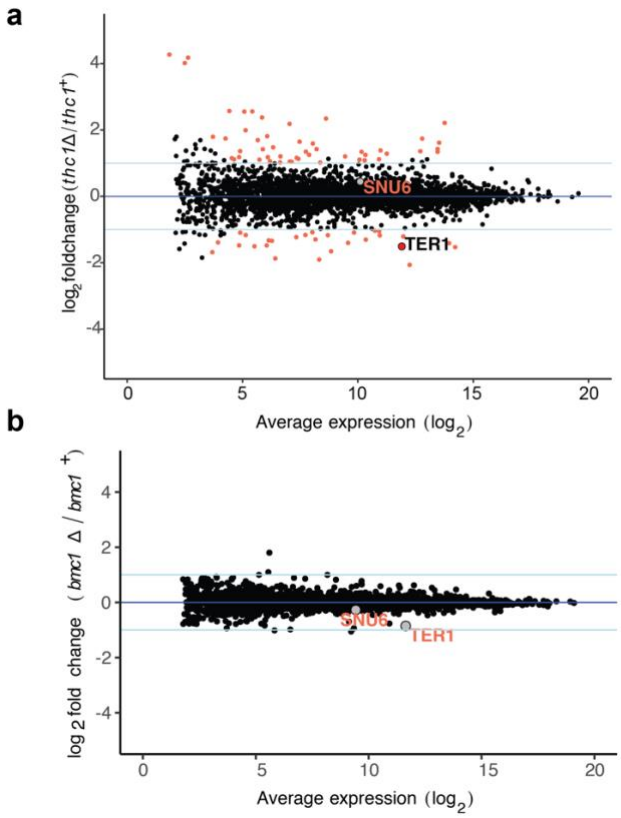

Figure 3.11: MA plot of differential expression analysis for the deletion of *thc1* **a**) and *bmc1* **b**). Average expression is plotted on the x-axis and $\log_2$ fold change between the deletion and wildtype is plotted on the y-axis. Differentially expressed genes with an absolute $\log_2$ fold change of ≥ 1 and an adjusted p-value < 0.05 are colored in red.

III.5: Discussion

In chapter II, I showed that Pof8, a member of the LARP7 family, regulates the transition from Sm to Lsm2-8 complex onto TER1 biogenesis. In chapter III I identified Thc1 and Bmc1 two previously uncharacterized proteins forming a complex with Pof8. Deletion of *thc1* or *bmc1* decreases the association between Pof8 and TER1, lower steady-state TER1 levels, and telomere length maintenance defects. This work is the first report of an association among a LARP7 family member and MePCE/Bin3 outside the 7SK complex suggesting a more ancient origin of the association of these proteins.

It was proposed that the 7SK RNP complex is an innovation of metazoans to regulate transcription elongation [233]. At its core, the 7SK is composed of three subunits: 7SK non-coding RNA, LARP7, and MePCE. The two protein subunits have broad distributions along

the tree of life. LARP7 family members (Pof8 in *S. pombe*) have been identified in all animals, some protists, and fungi, except in the most basal lineages [237]. MePCE (Bmc1 in *S. pombe*) has an even older origin being present in more basal eukaryotes [232,233]. The presence of non-coding 7SK RNA is limited only to metazoans. However, the last statement is based on a lack of evidence, and it remains to be seen whether a 7SK-like complex exists in *S. pombe* and other unicellular organisms. The identification of Bmc1 and Pof8 functioning together in telomerase biogenesis demonstrate more ancient roots of the Larp7 – MePCE/Bin3 partnership than previously recognized. Interestingly, the stability of Bmc1 is largely dependent on the presence of Thc1 and Pof8 indicating that most if not all Bmc1 in the cell are in complex with these two proteins. They are likely forming a functional unit on their cognates RNAs.

It was recently reported that the 7SK complex interacts with the SMN complex in the cytoplasm to regulate the assembly of the Sm complex on snRNAs to connect spliceosome production with transcriptional output [238]. Interestingly, telomerase RNA in yeasts share features with snRNAs, including the association of the Sm and Lsm2-8 complexes, an Sm binding site, and a TMG cap structure [119–121]. Whether this resemblance was critical in telomerase acquiring these biogenesis factors or whether telomerase is the more ancient target RNA remains to be investigated.

The LARP7 family shares an atypical RNA recognition motif, named xRRM, at the C-terminus which confers high affinity and specificity to their cognate RNAs. The xRRM domain has three unique features, a non-canonical RNP1 and RNP2 sequences, an RNP3 motif, and an extended helix at the C-terminus named α3 [80]. The α3 helix is a fundamental part of the binding to the target RNAs creating a binding pocket between the RNA and the third and fourth turns of the α3-helix in p65 and human LARP7[239]. Truncation of the p65 and LARP7 α3-helix extra turns resulted in reduced binding affinity to substrate RNA[81,239,240].

Curiously, Pof8 α3-helix is only three turns long which makes it shorter than the other family members hinting that the shorter α3-helix in Pof8 will also result in a weaker binding affinity for TER1. Nevertheless, the α3 helix in Pof8 is important for RNA binding as mutations lower the affinity for TER1, reduce TER1 levels and cause telomere shortening similar to a full deletion [223,224,227]. Hu *et al.,* 2020 proposed an alternative model where the RRM1-xRRM and Pof8-Lsm2-8 cooperatively bind to the TER1. However, they did not observe an increase in telomerase activity when recombinant Pof8 was added to the *pof8Δ* native protein extract. In contrast, we find that the addition of a preformed Thc1–Pof8 complex increased the association of TER1 with Lsm2-8 and strongly enhanced *in vitro* telomerase activity (Figure 3.7g). These results suggest that Thc1 and Pof8 bind the RNA cooperatively.

Finally, the similarity of Thc1 with the cap-binding domain of PARN and NCBP3 raises the exciting possibility that Thc1 interacts with the 5' end of TER1 while Pof8 binds the pseudoknot region and Lsm2-8 the 3' end of TER1, thereby integrating information on the processing and folding status from distant regions of the RNA and providing quality control for subsequent biogenesis events such as the loading of the catalytic subunit. It will be important to identify the binding site of Thc1 to TER1 and how it interacts with Pof8 and Bmc1 to unveil its function during TER1 biogenesis.

Acknowledgement

We thank the Stowers Institute for Medical Research Proteomics, Molecular Biology and Media Preparation facilities as well as the Media Lab and Genomics Core Facility at the Institute of Molecular Biology for excellent service and support; Kristi Jensen and Lisa Lassise for help with preparing the manuscript and all members of the Baumann laboratory for helpful discussions. This work was supported in part by the Howard Hughes Medical

Institute and the Stowers Institute for Medical Research. P.B. is an Alexander von Humboldt Professor at Johannes Gutenberg University.

Chapter IV: Conclusion and Future Directions

IV.1: Contribution to the Telomere Biology field

Telomerase RNA (TR) from all organisms goes through a series of maturation step to ensure the proper quality and assemble to the reverse transcriptase catalytic subunit. During the evolution of telomerase, TR co-opted cellular machinery to optimize its processing which is reflected in the highly diverse sequence and length of the RNA in different clades. The main force driving their biogenesis pathway is the 3', and as such, it contains species-specific sequence and structural elements to accommodate processing factors. For instance, ciliate TR biogenesis relies on the transcription terminal by RNA polymerase III with a short 3' - poly(U) track which binds to p65/43, a La-related protein. The fungal species on the other hand follow the small non-coding RNA biogenesis pathway. Protein complexes like Sm and Lsm2-8 are part of their maturation. Vertebrate TR contains domains similar to the H/ACA small nucleolar and small Cajal body non-coding RNAs which recruits the dyskerin complex. In recent years, common structures in TR have been identified despite the innate diversity of their sequence [72]. The template core domain with a template boundary element and a pseudoknot are examples of the TR conserved structures. I identified Pof8 which the homologs in ciliates also play a prominent role in TR assembly showing conservation of the LARP7 family in telomere biology independent of the biogenesis pathway. Unpublished data from a fellow lab member hint at the involvement of the human LARP7 in TR biogenesis. In conclusion, the findings presented in this dissertation advances our understanding of telomerase RNA (TR) processing in fission yeast and give common principles in the seemly diverse telomerase. Finding these principles can lead to better strategies to manipulate telomerase *in vivo*. Which in turn, it can be critical as a cancer therapy since ~85% of cancer require the re-activation of telomerase for their progression of [241].

In Chapter Two, I described a telomeric protein involved in telomerase RNA biogenesis in fission yeast. I characterized and provided a possible molecular function. Deletion of *pof8* causes an assembly defect by decreasing the affinity of Lsm2-8 complex to TER1. I also showed Pof8 as a member of telomerase holoenzyme, reflecting a deep conserve action role of the LARP7 family in telomerase biogenesis among different clades.

There are three possible non-mutually exclusive mechanisms for how Pof8 favors Lsm2-8 binding to TER1. The first mechanism is that Pof8 directly loads the Lsm2-8 complex via protein-protein interaction which was published during the preparation of this dissertation by Hu et al., 2020. The model proposed that a protein-protein interaction between the Pof8 N-terminal and Lsm8 helps to load the Lsm2-8 complex to the TER1 3' end after spliceosomal cleavage [227]. Hu et al., also suggested that Pof8 RMM1 and xRRM domains preferentially bind to properly folded pseudo-knot (PK) using Pof8 as a quality control sensor. However, all the *in vitro* analyses were done without the context of Thc1 and Bmc1, which I showed affects the Pof8 binding onto TER1. It will be important to understand the behavior of Pof8 in that context. For instance, the addition of both Thc1 and Pof8 on native protein cell extract improves the Lsm2-8 complexly loading and telomerase activity which only adding Pof8 did not in Hu et al., 2020. Hence, I propose a model in which Pof8, as the function of other LARP7 members, is a chaperon ensuring the proper folding of the PK domain, not just properly folded molecules.

The second possible mechanism is that Pof8 partially protects TER1 3' end during Sm to Lsm2-8 transition. TER1 Sm binding site and the 5' splice overlap on *S. pombe* leaving a TER1 molecule containing a U_6G-3' sequence after spliceosomal cleavage. From biochemical and structural analyses a terminal non-uracil, guanine in the case of TER1, diminish Lsm2-8 complex affinity to the 3' end of RNA molecules [119,124]. TER1 species

recovered from 2-8 pull-downs show enrichment of U_{4-6} sequence profile [119]. These results suggest a processing step between Sm to Lsm2-8 complex transition. I propose that Pof8 may temporarily bind to the uridine stretch protecting it from a 3' to 5' exonuclease which trims the terminal guanine increasing the affinity for the Lsm2-8 complex. This particular putative Pof8 function is reminiscent of the Genuine La protein protective mechanism. Supporting the last hypothesis, introducing TER1 mutants that generate a U6-3' or U6G-3' ends via ribozyme cleavage resulted in a clear increase in co-precipitation with Lsm4 and steady-state levels of TER1 in *pof8Δ+* cells without the need to increase transcription (Figure 5.1). The same effect was observed in *pof8Δ* strains resulted in a clear increase in co-precipitation of TER1 with Lsm4 when compared with WT and U6G-3' ribozyme mutant (Figure 5.1). However, there may be more regulatory roles of Pof8 since TER1 U6-3' ribozyme mutant cannot fully rescue the *pof8* deletion phenotype.

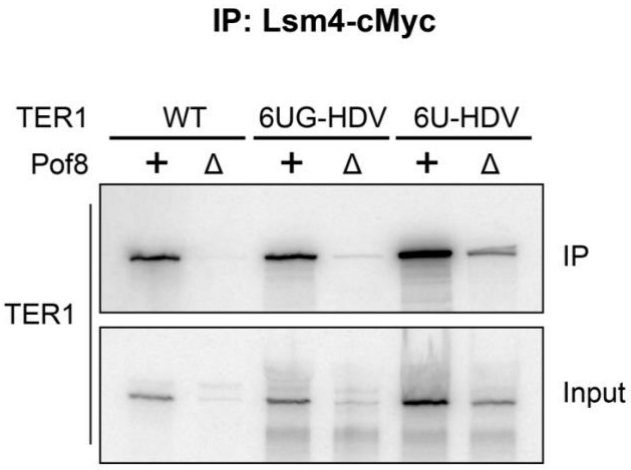

Figure 4.1: Extra non-uridine nucleotides at the TER1 3' affect the loading of the Lsm2-8 complex. Northern blot analysis of TER1 following immunoprecipitation (IP) of cMyc-tagged Lsm4 protein. TER1 WT, TER1 mutant cleaved on U_6G, and U_6 -3' end was expressed on a plasmid under the control of its endogenous promoter. Input represents 10% of the IP samples.

The third model of how Pof8 helps Lsm2-8 complex loading is a destabilization of the Sm ring. The Sm and Lsm binding sites overlap substantially on TER1. It was previously proposed that a temporal separation resolved the potential conflict between these two complexes. Sm is bound to the precursor and pre-mature forms while the Lsm2-8 complex binds exclusively to the mature form [119]. In this scenario, the destabilization of the Sm complex is the result of the change in sequence and structure after spliceosomal cleavage. However, Tgs1 hypermethylates the cap after spliceosomal cleavage via Sm recruitment. The same hypothesis then predicts that Sm binds long enough for Tgs1 catalysis activity but short enough to not interfere with Lsm2-8 complex recruitment. Here, I propose that Pof8 can destabilize the Sm complex on time allowing to load the Lsm ring just after the 5' end modification. Some evidence to support the last hypothesis is the increase in TER1 co-immunoprecipitation with the Sm proteins in a *pof8* and *thc1* deletion (Figure 2.4D). It will be important to create a TER1 mutant where Sm and Lsm2-8 binding site does not overlap allowing us to test this hypothesis.

IV.4: What are the specific functions of telomerase-specific partners?

Chapter Three shows Pof8 forming a complex with at least two subunits, Thc1 and Bmc1. Thc1 is a previously uncharacterized protein that may contain an RNA recognition motif resembling the cap-binding domain of human NCBP3 and PARN. Bmc1 is a capping enzyme whose human homolog associates with the human homolog of Pof8. Deletion of these two proteins results in a defect in Pof8 binding to TER1 reflecting a decrease in telomerase RNA levels and telomere length dysregulation. These results uncover a putative connection between the TER1 biogenesis stages of the 5'and 3' mediated by the Pof8 complex and an unexpected connection between telomerase RNA biogenesis and transcription elongation machinery.

TER1 structure resembles snRNAs containing a 2,3,7 trymethylguanosine (TMG) cap which is added by Tgs1 after spliceosomal cleavage [119,121]. MePCE, the human homolog of Bmc1, transfers a methyl group to the γ-phosphate on the 5' of RNA pol III transcripts [232]. The identification of another cap methyltransferase, Bmc1, came as a surprise. Interestingly, structural and biochemical analyses of MePCE show stable binding to the 5' cap and stem-loop I of 7SK RNA after catalysis [176,179]. It also appears that a change between a methyl γ-phosphate and TMG cap to the 7SK does not interfere with the binding of MePCE [176,178]. We speculate that Bmc1 methyltransferase activity is dispensable in TER1 biogenesis (Figure 3.6) but its binding to the TMG cap 'reads' the biogenesis stage of TER1.

To test this hypothesis, additional Bmc1 separation of function mutants between capping and binding activity will be important to determine what is the function of Bmc1 during TER1 biogenesis. Additionally, the resembles of Thc1 to the cap-binding domains suggest a cooperative binding together with Bmc1. Mutagenesis of Thc1, especially in the putative binding cap domain, will be informative to test the model. Testing the affinity of Bmc1 and Thc1 to the TER1 TMG cap can be done in a Tgs1 deletion or *in vitro* assays using recombinant Thc1 and Bmc1.

LARP7 proteins rely on the xRRM domain for binding affinity and specificity to their target RNAs [80,239]. A prominent structure in the xRRM domain is the α3-helix which extensively interacts with the RNA stabilizing their interaction [80,240]. However, Pof8 α3-helix is considerably shorter when compared to the ones in p65 and LARP7 predicting to affect its binding to TER1 [239]. It was shown that deletion of *thc1* decreased the amount of TER1 co-precipitating with Pof8 (Figure 3.9). I proposed that Thc1, besides the putative cap binding function, can fulfill the function of the C-terminal α3-helix increasing the contact with Pof8 and TER1. It will be important to determine the binding site of Pof8 and Thc1 to TER1, as well as, measuring the binding affinity of Thc1-Pof8 *in vitro* by EMSA. I also purpose to

map the contact between Thc1 and Bmc1 either by mutagenesis follow by yeast two-hybrid or by crosslinking the recombinant proteins followed by mass spectrometry (XL-MS).

Finally, an outstanding question is: how does the Lsm2-8 complex promotes TER1 stability and assembly with TRT1, the fission yeast telomerase catalytic subunit? The most parsimonious hypothesis is that the Lsm2-8 complex form a steric hindrance limiting the access of exonucleases to the TER1 3' end. However, these need to be tested by purifying the putative TER1 exonuclease, Rrp6, and the Lsm2-8 complex. Also, induction of conformational changes of TER1 may favor telomerase catalytic subunit association. This hypothesis can be tested *in vitro* by chemically probing the change of TER1 structure when associating with the Lsm2-8 complex. Another possibility is a direct contact between Lsm2-8 and Trt1 which can be tested by co-immunoprecipitation and Benzonase treatment and/or yeast two-hybrid.

References:

1. Ishikawa, F. & Naito, T. Why do we have linear chromosomes? A matter ofAdam and Eve. *Mutation Research* 99–107 (1999) doi:10.1016/0022-2836(78)90294-2.
2. Lange, T. de. A loopy view of telomere evolution. *Front. Genet.* **6**, (2015).
3. Scully, R., Panday, A., Elango, R. & Willis, N. A. DNA double-strand break repair-pathway choice in somatic mammalian cells. *Nat. Rev. Mol. Cell Biol.* **20**, 698–714 (2019).
4. Su, T. T. Cellular responses to DNA damage: one signal, multiple choices. *Annu. Rev. Genet.* **40**, 187–208 (2006).
5. de Lange, T. How telomeres solve the end-protection problem. *Science* **326**, 948–952 (2009).
6. McClintock, B. The Behavior in Successive Nuclear Divisions of a Chromosome Broken at Meiosis. *Proc. Natl. Acad. Sci. U. S. A.* **25**, 405–416 (1939).
7. McClintock, B. The Stability of Broken Ends of Chromosomes in Zea Mays. *Genetics* **26**, 234–282 (1941).
8. Müller, H. J. The remaking of chromosomes. 181–198 (1938).
9. Blackburn, E. H. & Gall, J. G. A tandemly repeated sequence at the termini of the extrachromosomal ribosomal RNA genes in Tetrahymena. *J. Mol. Biol.* **120**, 33–53 (1978).
10. Szostak, J. W. & Blackburn, E. H. Cloning yeast telomeres on linear plasmid vectors. *Cell* **29**, 245–255 (1982).
11. de Lange, T. Shelterin-Mediated Telomere Protection. *Annu. Rev. Genet.* **52**, 223–247 (2018).
12. Blackburn, E. H. Structure and function of telomeres. *Nature* **350**, 569–573 (1991).

13. Meyne, J., Ratliff, R. L. & Moyzis, R. K. Conservation of the human telomere sequence (TTAGGG)n among vertebrates. *Proc. Natl. Acad. Sci.* **86**, 7049–7053 (1989).
14. Shampay, J., Szostak, J. W. & Blackburn, E. H. DNA sequences of telomeres maintained in yeast. *Nature* **310**, 154–157 (1984).
15. Sugawara, N. DNA sequences at the telomeres of the fission yeast S. pombe. *Harvard University* 464 (1988).
16. Kipling, D. & Cooke, H. J. Hypervariable ultra-long telomeres in mice. *Nature* **347**, 400–402 (1990).
17. Pluta, A. F., Kaine, B. P. & Spear, B. B. The terminal organization of macronuclear DNA IN Oxytricha fallax. *Nucleic Acids Res.* **10**, 8145–8154 (1982).
18. Zakian, V. A. Structure and function of telomeres. *Annu. Rev. Genet.* **23**, 579–604 (1989).
19. Hiraoka, Y., Henderson, E. & Blackburn, E. H. Not so peculiar: fission yeast telomere repeats. *Trends Biochem. Sci.* **23**, 126 (1998).
20. Podlevsky, J. D., Bley, C. J., Omana, R. V., Qi, X. & Chen, J. J.-L. The Telomerase Database. *Nucleic Acids Res.* **36**, D339–D343 (2008).
21. Cooper, J. P., Nimmo, E. R., Allshire, R. C. & Cech, T. R. Regulation of telomere length and function by a Myb-domain protein in fission yeast. *Nature* **385**, 744–747 (1997).
22. Ferreira, M. G. & Cooper, J. P. The Fission Yeast Taz1 Protein Protects Chromosomes from Ku-Dependent End-to-End Fusions. *Mol. Cell* **7**, 55–63 (2001).
23. Kanoh, J. & Ishikawa, F. spRap1 and spRif1, recruited to telomeres by Taz1, are essential for telomere function in fission yeast. *Curr. Biol. CB* **11**, 1624–1630 (2001).
24. Miller, K. M., Rog, O. & Cooper, J. P. Semi-conservative DNA replication through telomeres requires Taz1. *Nature* **440**, 824–828 (2006).

25. Tazumi, A. *et al.* Telomere-binding protein Taz1 controls global replication timing through its localization near late replication origins in fission yeast. *Genes Dev.* **26**, 2050–2062 (2012).

26. Markiewicz-Potoczny, M. *et al.* TRF2-mediated telomere protection is dispensable in pluripotent stem cells. *Nature* **589**, 110–115 (2021).

27. Ruis, P. *et al.* TRF2-independent chromosome end protection during pluripotency. *Nature* **589**, 103–109 (2021).

28. van Steensel, B., Smogorzewska, A. & de Lange, T. TRF2 Protects Human Telomeres from End-to-End Fusions. *Cell* **92**, 401–413 (1998).

29. Sfeir, A. *et al.* Mammalian telomeres resemble fragile sites and require TRF1 for efficient replication. *Cell* **138**, 90–103 (2009).

30. Smogorzewska, A. *et al.* Control of Human Telomere Length by TRF1 and TRF2. *Mol. Cell. Biol.* **20**, 1659–1668 (2000).

31. van Steensel, B. & de Lange, T. Control of telomere length by the human telomeric protein TRF1. *Nature* **385**, 740–743 (1997).

32. Denchi, E. L. & de Lange, T. Protection of telomeres through independent control of ATM and ATR by TRF2 and POT1. *Nature* **448**, 1068–1071 (2007).

33. Griffith, J. D. *et al.* Mammalian Telomeres End in a Large Duplex Loop. *Cell* **97**, 503–514 (1999).

34. Baumann, P. & Cech, T. R. Pot1, the Putative Telomere End-Binding Protein in Fission Yeast and Humans. *Science* **292**, 1171–1175 (2001).

35. Wang, X. & Baumann, P. Chromosome fusions following telomere loss are mediated by single-strand annealing. *Mol. Cell* **31**, 463–473 (2008).

36. Capper, R. *et al.* The nature of telomere fusion and a definition of the critical telomere length in human cells. *Genes Dev.* **21**, 2495–2508 (2007).

37. He, H. *et al.* POT1b protects telomeres from end-to-end chromosomal fusions and aberrant homologous recombination. *EMBO J.* **25**, 5180–5190 (2006).

38. Pitt, C. W. & Cooper, J. P. Pot1 inactivation leads to rampant telomere resection and loss in one cell cycle. *Nucleic Acids Res.* **38**, 6968–6975 (2010).

39. Wu, L. *et al.* Pot1 deficiency initiates DNA damage checkpoint activation and aberrant homologous recombination at telomeres. *Cell* **126**, 49–62 (2006).

40. Miyoshi, T., Kanoh, J., Saito, M. & Ishikawa, F. Fission yeast Pot1-Tpp1 protects telomeres and regulates telomere length. *Science* **320**, 1341–1344 (2008).

41. Chikashige, Y. & Hiraoka, Y. Telomere binding of the Rap1 protein is required for meiosis in fission yeast. *Curr. Biol. CB* **11**, 1618–1623 (2001).

42. Pan, L., Hildebrand, K., Stutz, C., Thomä, N. & Baumann, P. Minishelterins separate telomere length regulation and end protection in fission yeast. *Genes Dev.* **29**, 1164–1174 (2015).

43. Sugiyama, T. *et al.* SHREC, an effector complex for heterochromatic transcriptional silencing. *Cell* **128**, 491–504 (2007).

44. Kim, S. H., Kaminker, P. & Campisi, J. TIN2, a new regulator of telomere length in human cells. *Nat. Genet.* **23**, 405–412 (1999).

45. O'Connor, M. S., Safari, A., Xin, H., Liu, D. & Songyang, Z. A critical role for TPP1 and TIN2 interaction in high-order telomeric complex assembly. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 11874–11879 (2006).

46. Ye, J. Z.-S. *et al.* POT1-interacting protein PIP1: a telomere length regulator that recruits POT1 to the TIN2/TRF1 complex. *Genes Dev.* **18**, 1649–1654 (2004).

47. Ye, J. Z.-S. *et al.* TIN2 binds TRF1 and TRF2 simultaneously and stabilizes the TRF2 complex on telomeres. *J. Biol. Chem.* **279**, 47264–47271 (2004).

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

49. Sarthy, J., Bae, N. S., Scrafford, J. & Baumann, P. Human RAP1 inhibits non-homologous end joining at telomeres. *EMBO J.* **28**, 3390–3399 (2009).

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

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

52. Zhong, F. L. *et al.* TPP1 OB-fold domain controls telomere maintenance by recruiting telomerase to chromosome ends. *Cell* **150**, 481–494 (2012).

53. Watson, J. D. & Crick, F. H. Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. *Nature* **171**, 737–738 (1953).

54. Watson, J. D. & Crick, F. H. Genetical implications of the structure of deoxyribonucleic acid. *Nature* **171**, 964–967 (1953).

55. Bessman, M. J., Lehman, I. R., Simms, E. S. & Kornberg, A. Enzymatic synthesis of deoxyribonucleic acid. II. General properties of the reaction. *J. Biol. Chem.* **233**, 171–177 (1958).

56. Lehman, I. R., Bessman, M. J., Simms, E. S. & Kornberg, A. Enzymatic synthesis of deoxyribonucleic acid. I. Preparation of substrates and partial purification of an enzyme from Escherichia coli. *J. Biol. Chem.* **233**, 163–170 (1958).

57. Alberts, B. *et al. Molecular Biology of the Cell*. (Garland Science, 2002).

58. Olovnikov, A. M. A theory of marginotomy: The incomplete copying of template margin in enzymic synthesis of polynucleotides and biological significance of the phenomenon. *J. Theor. Biol.* **41**, 181–190 (1973).

59. Watson, J. D. Origin of concatemeric T7 DNA. *Nature. New Biol.* **239**, 197–201 (1972).

60. Shay, J. W. & Wright, W. E. Telomeres and telomerase: three decades of progress. *Nat. Rev. Genet.* **20**, 299–309 (2019).

61. Hayflick, L. The limited in vitro lifetime of human diploid cell strains. *Exp. Cell Res.* **37**, 614–636 (1965).

62. Hayflick, L. & Moorhead, P. S. The serial cultivation of human diploid cell strains. *Exp. Cell Res.* **25**, 585–621 (1961).

63. de Lange, T. *et al.* Structure and variability of human chromosome ends. *Mol. Cell. Biol.* **10**, 518–527 (1990).

64. Harley, C. B., Futcher, A. B. & Greider, C. W. Telomeres shorten during ageing of human fibroblasts. *Nature* **345**, 458–460 (1990).

65. Oizumi, Y. *et al.* Complete sequences of Schizosaccharomyces pombe subtelomeres reveal multiple patterns of genome variation. *Nat. Commun.* **12**, 611 (2021).

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

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

68. Lendvay, T. S., Morris, D. K., Sah, J., Balasubramanian, B. & Lundblad, V. Senescence Mutants of Saccharomyces cerevisiae With a Defect in Telomere Replication Identify Three Additional EST Genes. *Genetics* **144**, 1399–1412 (1996).

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

70. Harrington, L. *et al.* A mammalian telomerase-associated protein. *Science* **275**, 973–977 (1997).

71. Nakamura, T. M. *et al.* Telomerase catalytic subunit homologs from fission yeast and human. *Science* **277**, 955–959 (1997).

72. Podlevsky, J. D. & Chen, J. J.-L. Evolutionary perspectives of telomerase RNA structure and function. *RNA Biol.* **13**, 720–732 (2016).

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

74. Wu, R. A., Upton, H. E., Vogan, J. M. & Collins, K. Telomerase Mechanism of Telomere Synthesis. *Annu. Rev. Biochem.* **86**, 439–460 (2017).

75. Schmidt, J. C. & Cech, T. R. Human telomerase: biogenesis, trafficking, recruitment, and activation. *Genes Dev.* **29**, 1095–1105 (2015).

76. Egan, E. D. & Collins, K. Biogenesis of telomerase ribonucleoproteins. *RNA* **18**, 1747–1759 (2012).

77. Aigner, S., Postberg, J., Lipps, H. J. & Cech, T. R. The Euplotes La Motif Protein p43 Has Properties of a Telomerase-Specific Subunit. *Biochemistry* **42**, 5736–5747 (2003).

78. Aigner, S. *et al.* Euplotes telomerase contains an La motif protein produced by apparent translational frameshifting. *EMBO J.* **19**, 6230–6239 (2000).

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

80. Singh, M., Choi, C. P. & Feigon, J. xRRM: a new class of RRM found in the telomerase La family protein p65. *RNA Biol.* **10**, 353–359 (2013).

81. Singh, M. *et al.* Structural Basis for Telomerase RNA Recognition and RNP Assembly by the Holoenzyme La Family Protein p65. *Mol. Cell* **47**, 16–26 (2012).

82. Berman, A. J., Gooding, A. R. & Cech, T. R. Tetrahymena telomerase protein p65 induces conformational changes throughout telomerase RNA (TER) and rescues

telomerase reverse transcriptase and TER assembly mutants. *Mol. Cell. Biol.* **30**, 4965–4976 (2010).

83. Jiang, J. *et al.* The architecture of Tetrahymena telomerase holoenzyme. *Nature* **496**, 187–192 (2013).
84. O'Connor, C. M. & Collins, K. A Novel RNA Binding Domain in Tetrahymena Telomerase p65 Initiates Hierarchical Assembly of Telomerase Holoenzyme. *Mol. Cell. Biol.* **26**, 2029–2036 (2006).
85. Stone, M. D. *et al.* Stepwise protein-mediated RNA folding directs assembly of telomerase ribonucleoprotein. *Nature* **446**, 458–461 (2007).
86. Prathapam, R., Witkin, K. L., O'Connor, C. M. & Collins, K. A telomerase holoenzyme protein enhances telomerase RNA assembly with telomerase reverse transcriptase. *Nat. Struct. Mol. Biol.* **12**, 252–257 (2005).
87. Eckert, B. & Collins, K. Roles of telomerase reverse transcriptase N-terminal domain in assembly and activity of Tetrahymena telomerase holoenzyme. *J. Biol. Chem.* **287**, 12805–12814 (2012).
88. Jiang, J. *et al.* Structure of Telomerase with Telomeric DNA. *Cell* **173**, 1179-1190.e13 (2018).
89. Min, B. & Collins, K. An RPA-related sequence-specific DNA-binding subunit of telomerase holoenzyme is required for elongation processivity and telomere maintenance. *Mol. Cell* **36**, 609–619 (2009).
90. Möllenbeck, M. *et al.* The telomerase-associated protein p43 is involved in anchoring telomerase in the nucleus. *J. Cell Sci.* **116**, 1757–1761 (2003).
91. Feng, J. *et al.* The RNA component of human telomerase. *Science* **269**, 1236–1241 (1995).

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

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

94. Roake, C. M. & Artandi, S. E. Regulation of human telomerase in homeostasis and disease. *Nat. Rev. Mol. Cell Biol.* **21**, 384–397 (2020).

95. Nguyen, D. *et al.* A Polyadenylation-Dependent 3’ End Maturation Pathway Is Required for the Synthesis of the Human Telomerase RNA. *Cell Rep.* **13**, 2244–2257 (2015).

96. Tseng, C.-K. *et al.* Human Telomerase RNA Processing and Quality Control. *Cell Rep.* **13**, 2232–2243 (2015).

97. Boyraz, B. *et al.* Posttranscriptional manipulation of TERC reverses molecular hallmarks of telomere disease. *J. Clin. Invest.* **126**, 3377–3382 (2016).

98. Moon, D. H. *et al.* Poly(A)-specific ribonuclease (PARN) mediates 3’-end maturation of the telomerase RNA component. *Nat. Genet.* **47**, 1482–1488 (2015).

99. Stuart, B. D. *et al.* Exome sequencing links mutations in PARN and RTEL1 with familial pulmonary fibrosis and telomere shortening. *Nat. Genet.* **47**, 512–517 (2015).

100. Tseng, C.-K., Wang, H.-F., Schroeder, M. R. & Baumann, P. The H/ACA complex disrupts triplex in hTR precursor to permit processing by RRP6 and PARN. *Nat. Commun.* **9**, 5430 (2018).

101. Cristofari, G. *et al.* Human telomerase RNA accumulation in Cajal bodies facilitates telomerase recruitment to telomeres and telomere elongation. *Mol. Cell* **27**, 882–889 (2007).

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

103. Chen, L. *et al.* Loss of Human TGS1 Hypermethylase Promotes Increased Telomerase RNA and Telomere Elongation. *Cell Rep.* **30**, 1358-1372.e5 (2020).

104. Venteicher, A. S. *et al.* A human telomerase holoenzyme protein required for Cajal body localization and telomere synthesis. *Science* **323**, 644–648 (2009).

105. Ghanim, G. E. *et al.* Structure of human telomerase holoenzyme with bound telomeric DNA. *Nature* 1–5 (2021) doi:10.1038/s41586-021-03415-4.

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

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

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

109. Bosoy, D., Peng, Y., Mian, I. S. & Lue, N. F. Conserved N-terminal motifs of telomerase reverse transcriptase required for ribonucleoprotein assembly in vivo. *J. Biol. Chem.* **278**, 3882–3890 (2003).

110. Noël, J.-F., Larose, S., Abou Elela, S. & Wellinger, R. J. Budding yeast telomerase RNA transcription termination is dictated by the Nrd1/Nab3 non-coding RNA termination pathway. *Nucleic Acids Res.* **40**, 5625–5636 (2012).

111. Jamonnak, N. *et al.* Yeast Nrd1, Nab3, and Sen1 transcriptome-wide binding maps suggest multiple roles in post-transcriptional RNA processing. *RNA N. Y. N* **17**, 2011–2025 (2011).

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

113. Bajon, E., Laterreur, N. & Wellinger, R. J. A Single Templating RNA in Yeast Telomerase. *Cell Rep.* **12**, 441–448 (2015).

114. Garcia, P. D. *et al.* Stability and nuclear localization of yeast telomerase depend on protein components of RNase P/MRP. *Nat. Commun.* **11**, 2173 (2020).

115. Lemieux, B. *et al.* Active Yeast Telomerase Shares Subunits with Ribonucleoproteins RNase P and RNase MRP. *Cell* **165**, 1171–1181 (2016).

116. Seto, A. G., Zaug, A. J., Sobel, S. G., Wolin, S. L. & Cech, T. R. telomerase is an Sm small nuclear ribonucleoprotein particle. **401**, 4 (1999).

117. Zappulla, D. C. *et al.* Ku can contribute to telomere lengthening in yeast at multiple positions in the telomerase RNP. *RNA N. Y. N* **17**, 298–311 (2011).

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

119. Tang, W., Kannan, R., Blanchette, M. & Baumann, P. Telomerase RNA biogenesis involves sequential binding by Sm and Lsm complexes. *Nature* **484**, 260–264 (2012).

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

121. Box, J. A., Bunch, J. T., Tang, W. & Baumann, P. Spliceosomal cleavage generates the 3′ end of telomerase RNA. *Nature* **456**, 910–914 (2008).

122. Kannan, R. *et al.* Intronic sequence elements impede exon ligation and trigger a discard pathway that yields functional telomerase RNA in fission yeast. *Genes Dev.* **27**, 627–638 (2013).

123. Lee, N. N. *et al.* Mtr4-like protein coordinates nuclear RNA processing for heterochromatin assembly and for telomere maintenance. *Cell* **155**, 1061–1074 (2013).

124. Zhou, L. *et al.* Crystal structures of the Lsm complex bound to the 3′ end sequence of U6 small nuclear RNA. *Nature* **506**, 116–120 (2014).

125. Shchepachev, V., Wischnewski, H., Missiaglia, E., Soneson, C. & Azzalin, C. M. Mpn1, Mutated in Poikiloderma with Neutropenia Protein 1, Is a Conserved 3′-to-5′ RNA Exonuclease Processing U6 Small Nuclear RNA. *Cell Rep.* **2**, 855–865 (2012).

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

127. Séraphin, B. 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 (1995).

128. Veretnik, S., Wills, C., Youkharibache, P., Valas, R. E. & Bourne, P. E. Sm/Lsm Genes Provide a Glimpse into the Early Evolution of the Spliceosome. *PLOS Comput. Biol.* **5**, e1000315 (2009).

129. Wahl, M. C., Will, C. L. & Lührmann, R. The spliceosome: design principles of a dynamic RNP machine. *Cell* **136**, 701–718 (2009).

130. Tarn, W. Y. & Steitz, J. A. A novel spliceosome containing U11, U12, and U5 snRNPs excises a minor class (AT-AC) intron in vitro. *Cell* **84**, 801–811 (1996).

131. Grimm, C., Stefanovic, B. & Schümperli, D. The low abundance of U7 snRNA is partly determined by its Sm binding site. *EMBO J.* **12**, 1229–1238 (1993).

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

133. Branlant, C. *et al.* U2 RNA shares a structural domain with U1, U4, and U5 RNAs. *EMBO J.* **1**, 1259–1265 (1982).

134. Raker, V. A., Hartmuth, K., Kastner, B. & Lührmann, R. Spliceosomal U snRNP core assembly: Sm proteins assemble onto an Sm site RNA nonanucleotide in a specific and thermodynamically stable manner. *Mol. Cell. Biol.* **19**, 6554–6565 (1999).

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

136. Meister, G., Eggert, C. & Fischer, U. SMN-mediated assembly of RNPs: a complex story. *Trends Cell Biol.* **12**, 472–478 (2002).

137. Mouaikel, J., Verheggen, C., Bertrand, E., Tazi, J. & Bordonné, R. 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 (2002).

138. Plessel, G., Fischer, U. & Lührmann, R. 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 (1994).

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

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

141. Zaric, B. *et al.* Reconstitution of two recombinant LSm protein complexes reveals aspects of their architecture, assembly, and function. *J. Biol. Chem.* **280**, 16066–16075 (2005).

142. Wilusz, C. J. & Wilusz, J. Lsm proteins and Hfq. *RNA Biol.* **10**, 592–601 (2013).

143. Achsel, T. *et al.* 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 (1999).

144. Reddy, R., Henning, D., Das, G., Harless, M. & Wright, D. The capped U6 small nuclear RNA is transcribed by RNA polymerase III. *J. Biol. Chem.* **262**, 75–81 (1987).

145. Stefano, J. E. Purified lupus antigen La recognizes an oligouridylate stretch common to the 3' termini of RNA polymerase III transcripts. *Cell* **36**, 145–154 (1984).

146. Didychuk, A. L. *et al.* Usb1 controls U6 snRNP assembly through evolutionarily divergent cyclic phosphodiesterase activities. *Nat. Commun.* **8**, 497 (2017).

147. Lund, E. & Dahlberg, J. E. Cyclic 2',3'-phosphates and nontemplated nucleotides at the 3' end of spliceosomal U6 small nuclear RNA's. *Science* **255**, 327–330 (1992).

148. Trippe, R., Sandrock, B. & Benecke, B. J. A highly specific terminal uridylyl transferase modifies the 3'-end of U6 small nuclear RNA. *Nucleic Acids Res.* **26**, 3119–3126 (1998).

149. Didychuk, A. L., Butcher, S. E. & Brow, D. A. The life of U6 small nuclear RNA, from cradle to grave. *RNA N. Y. N* **24**, 437–460 (2018).

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

151. Salgado-Garrido, J., Bragado-Nilsson, E., Kandels-Lewis, S. & Séraphin, B. Sm and Sm-like proteins assemble in two related complexes of deep evolutionary origin. *EMBO J.* **18**, 3451–3462 (1999).

152. Parker, R. RNA degradation in Saccharomyces cerevisae. *Genetics* **191**, 671–702 (2012).

153. Tharun, S. *et al.* Yeast Sm-like proteins function in mRNA decapping and decay. *Nature* **404**, 515–518 (2000).

154. Chowdhury, A., Mukhopadhyay, J. & Tharun, S. The decapping activator Lsm1p-7p-Pat1p complex has the intrinsic ability to distinguish between oligoadenylated and polyadenylated RNAs. *RNA N. Y. N* **13**, 998–1016 (2007).

155. Sharif, H. & Conti, E. Architecture of the Lsm1-7-Pat1 complex: a conserved assembly in eukaryotic mRNA turnover. *Cell Rep.* **5**, 283–291 (2013).

156. Liu, Q., Greimann, J. C. & Lima, C. D. Reconstitution, activities, and structure of the eukaryotic RNA exosome. *Cell* **127**, 1223–1237 (2006).

157. Lim, J. *et al.* Uridylation by TUT4 and TUT7 Marks mRNA for Degradation. *Cell* **159**, 1365–1376 (2014).

158. Rissland, O. S. & Norbury, C. J. Decapping is preceded by 3' uridylation in a novel pathway of bulk mRNA turnover. *Nat. Struct. Mol. Biol.* **16**, 616–623 (2009).

159. Decker, C. J., Teixeira, D. & Parker, R. Edc3p and a glutamine/asparagine-rich domain of Lsm4p function in processing body assembly in Saccharomyces cerevisiae. *J. Cell Biol.* **179**, 437–449 (2007).

160. Maraia, R. J., Mattijssen, S., Cruz-Gallardo, I. & Conte, M. R. The LARPs, La and related RNA-binding proteins: Structures, functions and evolving perspectives. *Wiley Interdiscip. Rev. RNA* **8**, (2017).

161. Bousquet-Antonelli, C. & Deragon, J.-M. A comprehensive analysis of the La-motif protein superfamily. *RNA* **15**, 750–764 (2009).

162. Merret, R. *et al.* The association of a La module with the PABP-interacting motif PAM2 is a recurrent evolutionary process that led to the neofunctionalization of La-related proteins. *RNA N. Y. N* **19**, 36–50 (2013).

163. Yang, R. *et al.* La-Related Protein 4 Binds Poly(A), Interacts with the Poly(A)-Binding Protein MLLE Domain via a Variant PAM2w Motif, and Can Promote mRNA Stability. *Mol. Cell. Biol.* **31**, 542–556 (2011).

164. Mattioli, M. & Reichlin, M. Heterogeneity of RNA protein antigens reactive with sera of patients with systemic lupus erythematosus. Description of a cytoplasmic nonribosomal antigen. *Arthritis Rheum.* **17**, 421–429 (1974).

165. Lerner, M. R., Boyle, J. A., Hardin, J. A. & Steitz, J. A. Two novel classes of small ribonucleoproteins detected by antibodies associated with lupus erythematosus. *Science* **211**, 400–402 (1981).

166. Rinke, J. & Steitz, J. A. Association of the lupus antigen La with a subset of U6 snRNA molecules. *Nucleic Acids Res.* **13**, 2617–2629 (1985).

167. Rinke, J. & Steitz, J. A. Precursor molecules of both human 5S ribosomal RNA and transfer RNAs are bound by a cellular protein reactive with anti-La lupus antibodies. *Cell* **29**, 149–159 (1982).

168. Fan, H., Goodier, J. L., Chamberlain, J. R., Engelke, D. R. & Maraia, R. J. 5' processing of tRNA precursors can Be modulated by the human La antigen phosphoprotein. *Mol. Cell. Biol.* **18**, 3201–3211 (1998).

169. Van Horn, D. J., Yoo, C. J., Xue, D., Shi, H. & Wolin, S. L. The La protein in Schizosaccharomyces pombe: a conserved yet dispensable phosphoprotein that functions in tRNA maturation. *RNA N. Y. N* **3**, 1434–1443 (1997).

170. Jacks, A. *et al.* Structure of the C-Terminal Domain of Human La Protein Reveals a Novel RNA Recognition Motif Coupled to a Helical Nuclear Retention Element. *Structure* **11**, 833–843 (2003).

171. Liang, C. *et al.* Sjögren Syndrome Antigen B (SSB)/La Promotes Global MicroRNA Expression by Binding MicroRNA Precursors through Stem-Loop Recognition*. *J. Biol. Chem.* **288**, 723–736 (2013).

172. Gurney, T. & Eliceiri, G. L. Intracellular distribution of low molecular weight RNA species in HeLa cells. *J. Cell Biol.* **87**, 398–403 (1980).

173. Nguyen, V. T., Kiss, T., Michels, A. A. & Bensaude, O. 7SK small nuclear RNA binds to and inhibits the activity of CDK9/cyclin T complexes. *Nature* **414**, 322–325 (2001).

174. Wassarman, D. A. & Steitz, J. A. Structural analyses of the 7SK ribonucleoprotein (RNP), the most abundant human small RNP of unknown function. *Mol. Cell. Biol.* **11**, 3432–3445 (1991).

175. Krueger, B. J. *et al.* LARP7 is a stable component of the 7SK snRNP while P-TEFb, HEXIM1 and hnRNP A1 are reversibly associated. *Nucleic Acids Res.* **36**, 2219–2229 (2008).

176. Xue, Y., Yang, Z., Chen, R. & Zhou, Q. A capping-independent function of MePCE in stabilizing 7SK snRNA and facilitating the assembly of 7SK snRNP. *Nucleic Acids Res.* **38**, 360–369 (2010).

177. Jeronimo, C. *et al.* Systematic analysis of the protein interaction network for the human transcription machinery reveals the identity of the 7SK capping enzyme. *Mol. Cell* **27**, 262–274 (2007).

178. Muniz, L., Egloff, S. & Kiss, T. RNA elements directing in vivo assembly of the 7SK/MePCE/Larp7 transcriptional regulatory snRNP. *Nucleic Acids Res.* **41**, 4686–4698 (2013).

179. Yang, Y., Eichhorn, C. D., Wang, Y., Cascio, D. & Feigon, J. Structural basis of 7SK RNA 5'-γ-phosphate methylation and retention by MePCE. *Nat. Chem. Biol.* **15**, 132–140 (2019).

180. Eichhorn, C. D., Yang, Y., Repeta, L. & Feigon, J. Structural basis for recognition of human 7SK long noncoding RNA by the La-related protein Larp7. *Proc. Natl. Acad. Sci. U. S. A.* **115**, E6457–E6466 (2018).

181. Brogie, J. E. & Price, D. H. Reconstitution of a functional 7SK snRNP. *Nucleic Acids Res.* **45**, 6864–6880 (2017).

182. He, N. *et al.* A La-related protein modulates 7SK snRNP integrity to suppress P-TEFb-dependent transcriptional elongation and tumorigenesis. *Mol. Cell* **29**, 588–599 (2008).

183. Eichhorn, C. D., Yang, Y., Repeta, L. & Feigon, J. Structural basis for recognition of human 7SK long noncoding RNA by the La-related protein Larp7. *Proc. Natl. Acad. Sci.* **115**, E6457–E6466 (2018).

184. Yik, J. H. N. *et al.* Inhibition of P-TEFb (CDK9/Cyclin T) kinase and RNA polymerase II transcription by the coordinated actions of HEXIM1 and 7SK snRNA. *Mol. Cell* **12**, 971–982 (2003).

185. Yang, Z., Zhu, Q., Luo, K. & Zhou, Q. The 7SK small nuclear RNA inhibits the CDK9/cyclin T1 kinase to control transcription. *Nature* **414**, 317–322 (2001).

186. Calo, E. *et al.* RNA helicase DDX21 coordinates transcription and ribosomal RNA processing. *Nature* **518**, 249–253 (2015).

187. Cherrier, T. *et al.* CTIP2 is a negative regulator of P-TEFb. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 12655–12660 (2013).

188. Hogg, J. R. & Collins, K. RNA-based affinity purification reveals 7SK RNPs with distinct composition and regulation. *RNA N. Y. N* **13**, 868–880 (2007).

189. Egloff, S., Studniarek, C. & Kiss, T. 7SK small nuclear RNA, a multifunctional transcriptional regulatory RNA with gene-specific features. *Transcription* **9**, 95–101 (2018).

190. Egloff, S. *et al.* The 7SK snRNP associates with the little elongation complex to promote snRNA gene expression. *EMBO J.* **36**, 934–948 (2017).

191. Flynn, R. A. *et al.* 7SK-BAF axis controls pervasive transcription at enhancers. *Nat. Struct. Mol. Biol.* **23**, 231–238 (2016).

192. Páez-Moscoso, D. J. *et al.* Pof8 is a La-related protein and a constitutive component of telomerase in fission yeast. *Nat. Commun.* **9**, 587 (2018).

193. Kannan, R., Helston, R. M., Dannebaum, R. O. & Baumann, P. Diverse mechanisms for spliceosome-mediated 3’ end processing of telomerase RNA. *Nat. Commun.* **6**, 6104 (2015).

194. Qi, X. *et al.* Prevalent and distinct spliceosomal 3’-end processing mechanisms for fungal telomerase RNA. *Nat. Commun.* **6**, 6105 (2015).

195. Gunisova, S. *et al.* Identification and comparative analysis of telomerase RNAs from Candida species reveal conservation of functional elements. *RNA N. Y. N* **15**, 546–559 (2009).

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

197. Aigner, S. & Cech, T. R. The Euplotes telomerase subunit p43 stimulates enzymatic activity and processivity in vitro. *RNA N. Y. N* **10**, 1108–1118 (2004).

198. Baumann, P. & Cech, T. R. Protection of telomeres by the Ku protein in fission yeast. *Mol. Biol. Cell* **11**, 3265–3275 (2000).

199. Bähler, J. *et al.* Heterologous modules for efficient and versatile PCR-based gene targeting in Schizosaccharomyces pombe. *Yeast Chichester Engl.* **14**, 943–951 (1998).

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

201. Brun, C., Dubey, D. D. & Huberman, J. A. pDblet, a stable autonomously replicating shuttle vector for Schizosaccharomyces pombe. *Gene* **164**, 173–177 (1995).

202. Church, G. M. & Gilbert, W. Genomic sequencing. *Proc. Natl. Acad. Sci. U. S. A.* **81**, 1991–1995 (1984).

203. Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinforma. Oxf. Engl.* **29**, 15–21 (2013).

204. Dandekar, T., Ribes, V. & Tollervey, D. Schizosaccharomyces pombe U4 small nuclear RNA closely resembles vertebrate U4 and is required for growth. *J. Mol. Biol.* **208**, 371–379 (1989).

205. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinforma. Oxf. Engl.* **26**, 139–140 (2010).

206. Lehmann, A. *et al.* Molecular interactions of fission yeast Skp1 and its role in the DNA damage checkpoint. *Genes Cells Devoted Mol. Cell. Mech.* **9**, 367–382 (2004).

207. Liu, N.-N., Han, T. X., Du, L.-L. & Zhou, J.-Q. A genome-wide screen for Schizosaccharomyces pombe deletion mutants that affect telomere length. *Cell Res.* **20**, 963–965 (2010).

208. Schmidt, M. W., McQuary, P. R., Wee, S., Hofmann, K. & Wolf, D. A. F-box-directed CRL complex assembly and regulation by the CSN and CAND1. *Mol. Cell* **35**, 586–597 (2009).

209. Hildebrand, A., Remmert, M., Biegert, A. & Söding, J. Fast and accurate automatic structure prediction with HHpred. *Proteins Struct. Funct. Bioinforma.* **77**, 128–132 (2009).

210. Mandell, J. G., Goodrich, K. J., Bähler, J. & Cech, T. R. Expression of a RecQ helicase homolog affects progression through crisis in fission yeast lacking telomerase. *J. Biol. Chem.* **280**, 5249–5257 (2005).

211. Bayfield, M. A., Yang, R. & Maraia, R. J. Conserved and divergent features of the structure and function of La and La-related proteins (LARPs). *Biochim. Biophys. Acta* **1799**, 365–378 (2010).

212. Jiang, J. *et al.* Structure of Tetrahymena telomerase reveals previously unknown subunits, functions, and interactions. *Science* **350**, aab4070 (2015).

213. Shukla, S., Schmidt, J. C., Goldfarb, K. C., Cech, T. R. & Parker, R. Inhibition of telomerase RNA decay rescues telomerase deficiency caused by dyskerin or PARN defects. *Nat. Struct. Mol. Biol.* **23**, 286–292 (2016).

214. Lin, K.-W. *et al.* Proteomics of yeast telomerase identified Cdc48-Npl4-Ufd1 and Ufd4 as regulators of Est1 and telomere length. *Nat. Commun.* **6**, 8290 (2015).

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

216. Kim, N. W. *et al.* Specific association of human telomerase activity with immortal cells and cancer. *Science* **266**, 2011–2015 (1994).

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

218. Pfeiffer, V. & Lingner, J. Replication of telomeres and the regulation of telomerase. *Cold Spring Harb. Perspect. Biol.* **5**, a010405 (2013).

219. Zappulla, D. C. & Cech, T. R. Yeast telomerase RNA: A flexible scaffold for protein subunits. *Proc. Natl. Acad. Sci.* **101**, 10024–10029 (2004).

220. Opresko, P. L. & Shay, J. W. Telomere-associated aging disorders. *Ageing Res. Rev.* **33**, 52–66 (2017).

221. Shay, J. W. Role of Telomeres and Telomerase in Aging and Cancer. *Cancer Discov.* **6**, 584–593 (2016).

222. Coy, S., Volanakis, A., Shah, S. & Vasiljeva, L. The Sm Complex Is Required for the Processing of Non-Coding RNAs by the Exosome. *PLOS ONE* **8**, e65606 (2013).

223. Collopy, L. C. *et al.* LARP7 family proteins have conserved function in telomerase assembly. *Nat. Commun.* **9**, 557 (2018).

224. Mennie, A. K., Moser, B. A. & Nakamura, T. M. LARP7-like protein Pof8 regulates telomerase assembly and poly(A)+TERRA expression in fission yeast. *Nat. Commun.* **9**, 586 (2018).

225. Hasler, D., Meister, G. & Fischer, U. Stabilize and connect: the role of LARP7 in nuclear non-coding RNA metabolism. *RNA Biol.* **18**, 290–303 (2021).

226. Holohan, B. *et al.* Impaired telomere maintenance in Alazami syndrome patients with LARP7 deficiency. *BMC Genomics* **17**, (2016).

227. Hu, X. *et al.* Quality-Control Mechanism for Telomerase RNA Folding in the Cell. *Cell Rep.* **33**, 108568 (2020).

228. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).

229. Lopez-Delisle, L. *et al.* pyGenomeTracks: reproducible plots for multivariate genomic datasets. *Bioinforma. Oxf. Engl.* **37**, 422–423 (2021).

230. Gebhardt, A. *et al.* mRNA export through an additional cap-binding complex consisting of NCBP1 and NCBP3. *Nat. Commun.* **6**, 8192 (2015).

231. Madeira, F. *et al.* The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic Acids Res.* **47**, W636–W641 (2019).

232. Cosgrove, M., Ding, Y., Rennie, W. A., Lane, M. J. & Hanes, S. D. The Bin3 RNA methyltransferase (MePCE) targets 7SK RNA to control transcription and translation. *Wiley Interdiscip. Rev. RNA* **3**, 633–647 (2012).

233. Marz, M. *et al.* Evolution of 7SK RNA and its protein partners in metazoa. *Mol. Biol. Evol.* **26**, 2821–2830 (2009).

234. Peterlin, B. M., Brogie, J. E. & Price, D. H. 7SK snRNA: a noncoding RNA that plays a major role in regulating eukaryotic transcription. *Wiley Interdiscip. Rev. RNA* **3**, 92–103 (2012).

235. Jeronimo, C. *et al.* Systematic analysis of the protein interaction network for the human transcription machinery reveals the identity of the 7SK capping enzyme. *Mol. Cell* **27**, 262–274 (2007).

236. Barboric, M. *et al.* Interplay between 7SK snRNA and oppositely charged regions in HEXIM1 direct the inhibition of P-TEFb. *EMBO J.* **24**, 4291–4303 (2005).

237. Deragon, J.-M. Distribution, organization an evolutionary history of La and LARPs in eukaryotes. *RNA Biol.* **0**, 1–9 (2020).

238. Ji, C. *et al.* Interaction of 7SK with the Smn complex modulates snRNP production. *Nat. Commun.* **12**, 1278 (2021).

239. Basu, R., Eichhorn, C. D., Cheng, R., Peterson, R. D. & Feigon, J. Structure of S. pombe telomerase protein Pof8 C-terminal domain is an xRRM conserved among LARP7 proteins. *RNA Biol.* **0**, 1–12 (2020).

240. Eichhorn, C. D., Chug, R. & Feigon, J. hLARP7 C-terminal domain contains an xRRM that binds the 3' hairpin of 7SK RNA. *Nucleic Acids Res.* **44**, 9977–9989 (2016).

241. Jafri, M. A., Ansari, S. A., Alqahtani, M. H. & Shay, J. W. Roles of telomeres and telomerase in cancer, and advances in telomerase-targeted therapies. *Genome Med.* **8**, (2016).