MECHANISM AND FUNCTION OF SPLICEOSOMAL CLEAVAGE IN FISSION YEAST

BY

Ram Kannan

B. Tech., PSG College Of Technology, 2007

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

_____

Dr. Peter Baumann (Advisor)

Committee members: _____

Dr. Peter Baumann (Co-chairperson)

_____

Dr. Lane Christenson (Co-chairperson)

_____

Dr. Aron Fenton

_____

Dr. Rong Li

_____

Dr. Marco Blanchette

Date Defended: 06/05/2013

The Dissertation Committee for Ram Kannan

certifies that this is the approved version of the following dissertation:

MECHANISM AND FUNCTION OF SPLICEOSOMAL CLEAVAGE IN FISSION YEAST

_____

Dr. Peter Baumann (Advisor)

_____

Dr. Lane Christenson (Co-chairperson)

_____

Dr. Aron Fenton

_____

Dr. Rong Li

_____

Dr. Marco Blanchette

Date approved: 06/06/13

Abstract

Telomerase is the ribonucleoprotein complex that replenishes lost DNA sequences at the ends of chromosomes. At its core, telomerase consists of an RNA subunit (TERC) that provides the template and a catalytic protein component (TERT). Insufficient telomerase activity leads to various disorders like dyskeratosis congenita, aplastic anemia and idiopathic pulmonary fibrosis. How different mutations in the same gene lead to disparate symptoms and disorders is not clear. The overall objective of my project is to understand the biogenesis of telomerase in the genetically tractable eukaryote *S. pombe,* whose telomere maintenance machinery closely resembles that of humans. Our laboratory has previously shown that the mature 3' end of *S. pombe* telomerase RNA (TER1) is generated by the first step of spliceosomal splicing. The cis- and trans- acting factors that distinguish the single step spliceosomal cleavage in TER1 from the two-step splicing reaction that removes introns in other genes are being investigated. We now demonstrate that a strong branch site (BS), a long distance to the 3' splice site (SS) and a weak polypyrimidine tract (Py) tract act synergistically to attenuate the transition from the first to the second step of splicing. The observation that a strong BS antagonizes the second step of splicing in the context of TER1 suggests that the BS-U2 snRNA interactions are disrupted after the first step and thus earlier than previously thought. The slow transition from first to second step triggers the Prp22 DExD/H-box helicase-dependent rejection of the cleaved products and Prp43-dependent discard of the splicing intermediates. Related to this work, we have established that the spliceosome generates the 3' ends of telomerase RNA in *S. cryophilus* and *S. octosporus* albeit via a different

mechanism involving U6 snRNA hyperstabilization at the 5'ss. Our findings explain how the spliceosome can function in 3' end processing and provide new insights into the mechanism of splicing.

Acknowledgements

This dissertation would not have been possible without the guidance and essential contributions of several individuals. Though, I can not include every single individual who has made a difference, the following people deserve a special mention:

First, I must thank my parents, Kannan and Vasu for their unconditional love and support which was imperative for completion of my graduate research. In addition, I would like to thank my sister, Akila for her constant guidance and help. I hope at some point in life I can repay these three individuals for all the things they have done for me.

I am indebted to my thesis advisor, Peter for providing a stimulating environment to perform research over the past 4 years. I will always cherish the fruitful discussions I had with Peter regarding my thesis project. Even more, Peter spent quality time in improving my soft skills that aided my professional development. It is hard to imagine a better place for a graduate student to complete their dissertation research.

My committee members Lane Christenson, Marco Blanchette, Rong Li, Aron Fenton and Gregory Vanden Heuvel gave many valuable advice that not only shaped my research but also steered me towards graduation on time. It was an amazing experience to be advised by such outstanding scientists and mentors.

To be successful in any career, it is important to have the right work environment. I was extremely lucky that Aracely, Elena, Evan, Guan, Jeremy, Jessica, Lili, Li, Martina,

Rachel, Richard, Saori, Sean,Veronica, Wen created an atmosphere that few can match. Thank for all the 'TNTL' moments (too numerous to list). You guys will be missed.

Life as a PhD student can be difficult. Add in a new country and culture, you begin to wonder what in the world have you gotten into. My task was simplified a great deal by a select few who made the transition as smooth as possible. Felcy, Naveen, Aritra, Hemant and Shachi were instrumental in helping me get settled in Kansas City. Their company made even the dreaded first year of IGPBS fun. Thanks for everything guys.

Special thanks to Aashwin, BBU, Vasanth, Vijay, Kavitha, Kokila, Shobhana, Shwetha for the fun moments that kept me in good stead always.

I am extremely thankful to the American Heart Association for awarding me a predoctoral fellowship that funded my stipend over the past two years. Getting a fellowship at a very early stage in my career gave me the confidence to succeed in graduate school.

Finally, I would like to thank my friends and fellow graduate students in Stowers and KUMC. Thank you for everything you have done.

Table of Contents

List of Figures

List of Appendices

Chapter I:

Introduction to telomerase and significance of research

I.1:Introduction to telomeres and telomerase

The genetic information that dictates the functions of a cell is packed in a thread like structure called chromosome. Within the nucleus, chromosomes are maintained in either circular or linear state. Though, circular chromosomes predominate in bacteria and archaea [1], eukaryotes exclusively maintain linear chromosomes due to their pivotal role in productive meiosis and for accommodating their genomic complexity [2]. In spite of these advantages, evolution of linear chromosomes in eukaryotes poses two perennial problems:

A. End protection problem

B. End replication problem

I.1.a: The first problem with linear ends: 'End protection' problem

Cells encounter wide variety of stress and mutagenic insults like UV light, genotoxic chemicals, radiation and reactive oxygen species. Each of these insults can alter DNA sequences and generate double strand break (DSB) thereby affecting the output of the genetic code, namely RNA and protein. To avoid accumulation of hazardous mutations, cells have developed numerous DNA repair pathways [3, 4]. As the mechanisms for fixing DSB were being deciphered, it became clear that the repair pathways recognize linear DNA molecules as a signal for broken DNA ends. In light of this finding, an

important question arose - Are linear chromosome ends sensed as DSB within the cell, or if not, how do they evade these repair pathways [5]?

I.1.b: Linear chromosome ends are shielded from DNA repair

Pioneering studies from McClintock and Muller established critical differences between behavior of linear chromosome ends and that of broken DNA. McClintock documented that, although double strand breaks are efficiently repaired by the cell, naturally occurring chromosome termini are immune to this process [6]. Muller reported similar findings and coined the term 'telomere' from the Greek words telos (end) and meros (part)[7].

One of the first insights regarding the function of telomeres came from Szostak and colleagues who demonstrated that linear DNA introduced into eukaryotes is unstable, because it recombines with the genome [8]. We now know that linear DNA is subject to two major repair pathways that mend DSB: Homology Directed Repair (HDR) and Non Homologous End Joining (NHEJ) [4, 9, 10].

A second breakthrough came from the work of Hartwell and colleagues who established that budding yeast cells lacking RAD9 failed to arrest the cell cycle in response to DSB [11]. Similar studies carried out in fission yeast and mammals further reinforced the idea that broken chromosomes by themselves are not triggering cell cycle arrest [12]. Rather, there are molecular pathways that sense double strand breaks and arrest cell cycle. Work over the past two decades has established that two conserved independent signaling pathways are activated by double-strand breaks: (i) the ATM (ataxia telangiectasia mutated) kinase pathway, which is directly activated by DNA ends, and (ii) the ATR

(ataxia telangiectasia and Rad3 related) kinase pathway, which senses the single-stranded DNA formed as a result of 5′ end resection of a double-strand break [5]. Based on these observations, a critical question arose: how chromosome ends inhibit these molecular pathways?

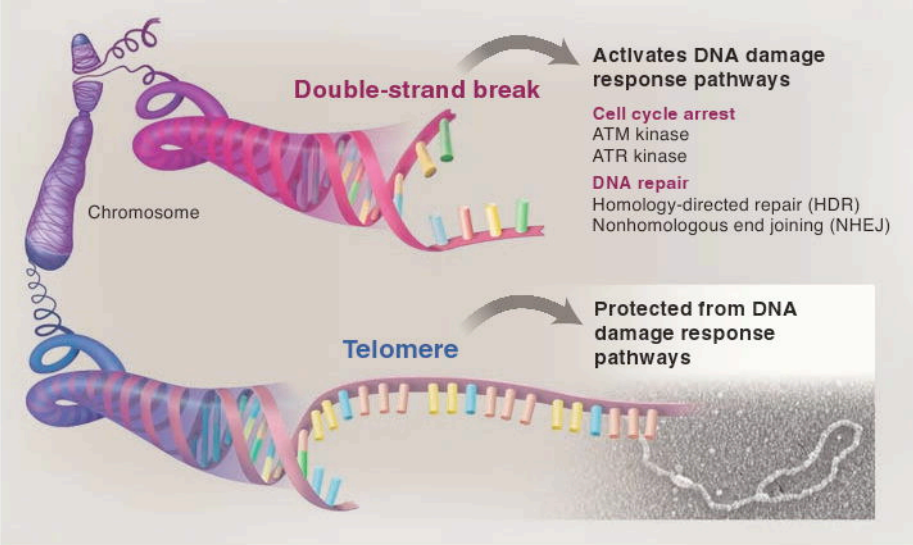


Figure 1.1: How telomeres solve 'The end-protection problem'. Adapted from [5]. A DNA break in mammalian chromosome (top) generates exposed DNA ends that can activate two signaling pathways (the ATM and ATR kinase pathways). These distinct pathways arrest the cell division cycle and can induce cell death. Either NHEJ or HDR pathways repair DSB, allowing normal progression of cell cycle. The presence of DNA damage response pathways is a perennial problem for the linear ends of chromosome (telomeres, bottom) because activation of DNA damage signaling or DNA repair at telomeres would be disastrous. Mammalian telomeres solve this end-protection problem through the use of a telomere-specific protein complex (shelterin) and an altered structure (the t-loop, pictured in the EM) that together safeguard linear chromosome ends.

I.1.c: Telomeres protect chromosome ends from DNA repair & degradation

Seminal studies from Blackburn & Szostak demonstrated that telomeres from the ciliate *Tetrahymena* can stabilize linear ends of yeast artificial chromosomes illustrating that the protective function of telomere is conserved across diverse species [13]. We now know 'telomeres' are nucleo-protein complexes that safeguard chromosome termini against erosion and illegitimate recombination events. Mammalian chromosomes end in tandem repeats of TTAGGG that ranges in size from 10-15 kilobases in humans to 20-60 kilobases in mice [14, 15]. A unique requirement of telomeric DNA is the presence of a single stranded 3' overhang which has been proposed to invade the double stranded region, thereby sequestering it from exposure to exonucleases [16]. In addition, specialized proteins recognize the single-stranded and double-stranded portions, coating telomeres to form a cap like structure. Telomeric Repeat Binding Factor 1 and 2 (TRF1 & TRF2) bind the double stranded part in a sequence specific fashion [17, 18]. TRF1 and TRF2 recruit TIN2 and RAP1 respectively [19, 20]. The single stranded region of telomeres is bound by POT1/TPP1 complex [21-25]. Together, these six proteins are referred to as 'Shelterin' [26]. Shelterin components have specific roles in chromosomal end protection and maintenance.

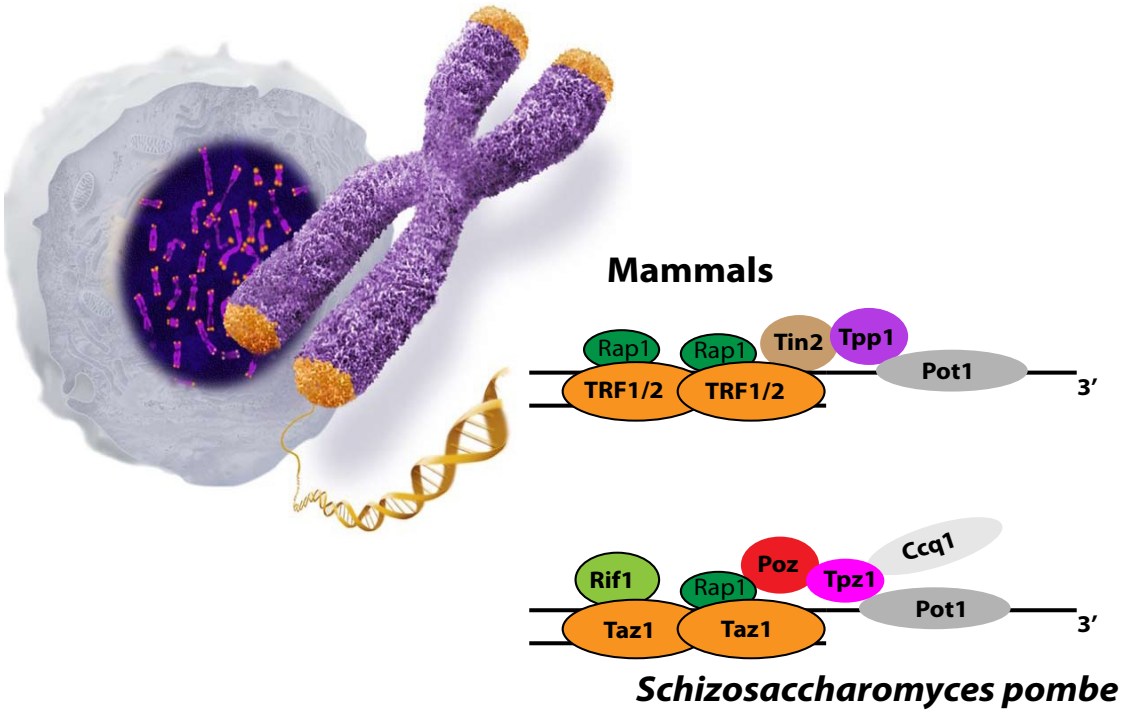


Figure 1.2: Schematic of human chromosome termini with telomere binding proteins. Human chromosome is shown in purple with telomeres highlighted in orange (left), adapted from Nobel Prize press release, 2009. The shelterin components in humans and *S. pombe* are depicted (right).

Repression of the ATM pathway is largely mediated by TRF2 [27]. Loss of TRF2 leads to ATM activation at natural chromosome ends, which can be visualized by telomeric localization of DNA Damage response proteins like 53BP1 and H2AX [28, 29]. Similarly, removal of POT1 triggers ATR signaling pathway, which results in phosphorylation of the downstream effectors Chk1 and Chk2 [30, 31]. Although, TRF2 is sufficient for mitigating the threat provided by NHEJ in mouse [32], studies in humans have implicated a role for both TRF2 and RAP1 [33, 34].

Apart from chromosome end protection, shelterin components have additional telomere related functions. TRF1 has been shown to promote efficient replication of repeats and prevents replicative fork stalling [35]. In light of recent studies, it is evident that TPP1 is essential for recruiting telomerase to chromosome ends in addition to its canonical function in end protection [24, 36, 37]. In summary, by recruiting shelterin proteins to the ends of all chromosomes, eukaryotes have developed an ingenious method to block the DNA repair machinery and safeguard the genome [5].

I.1.d: The second problem with linear ends: End replication problem

A second issue with linear DNA stems from inherent limitations of the DNA replication machinery. Semi-conservative replication can synthesize DNA only in the 5'-3' direction and the DNA polymerase requires binding of an RNA primer (Fig 1.3a). Thus, as the replication fork moves along the chromosome, there is a discord in the synthesis of the two strands. While one of the two daughter strands is synthesized continuously, the other daughter strand, known as the lagging strand, is synthesized discontinuously as short DNA fragments known as Okazaki fragments, each of which has its own RNA primer. The RNA primers are subsequently degraded, and the gaps between the Okazaki fragments are then filled in. A problem arises at the end of the chromosome, because the DNA repair machinery is unable to repair the gap left by the terminal RNA primer. Consequently, the new DNA molecule is shorter than the parent DNA molecule by at least the length of one RNA primer. In the absence of any solution, telomere shortening leads to replicative senescence, a physiological state in which no further cell division can

occur. This phenomenon predicted independently by James Watson & Alexander Olovnikov, is referred to as the 'end replication problem' [38-40], and is widely considered as a potent tumor suppressive mechanism for somatic cells. On the other hand, stem cells and rapidly proliferating cells require continued cell division and a strategy to replenish telomere length.

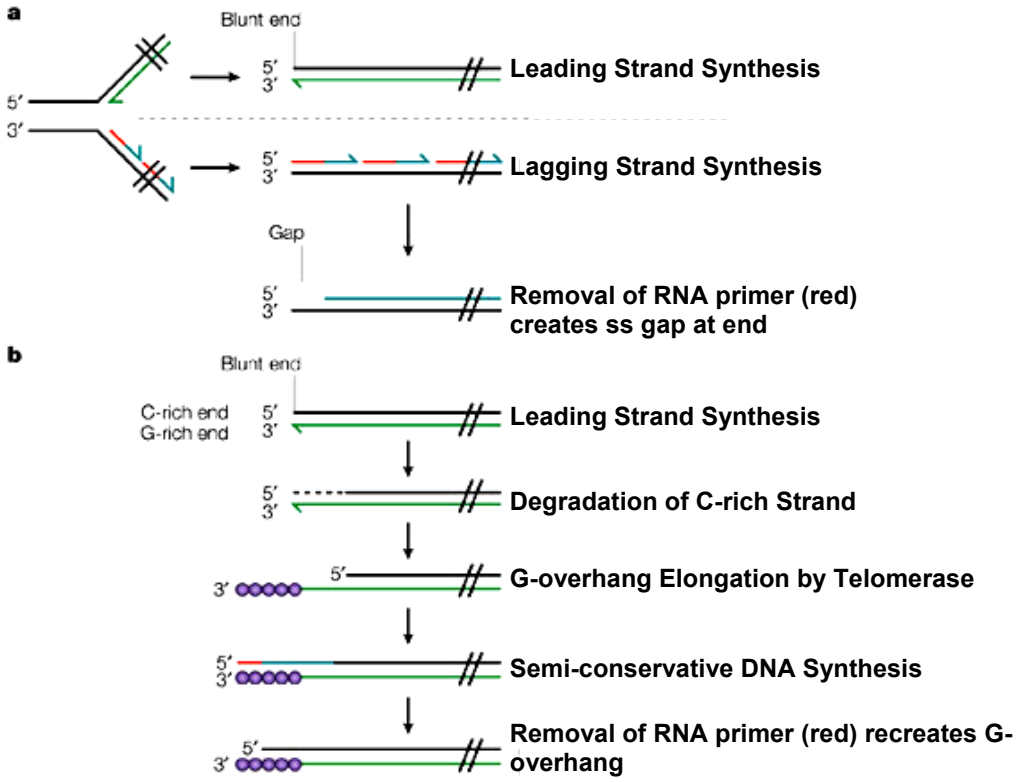


Figure 1.3: Telomerase counteracts telomere shortening in most of the eukaryotes. a) Conventional DNA polymerases synthesize DNA in the 5'→3' direction and require a 3'OH group provided by an 8–12-base segment of RNA as a primer (red). The leading strand is synthesized continuously (green). In contrast, the lagging strand is synthesized in short, RNA-primed OKAZAKI FRAGMENTS (blue). After extension, the primers are removed and the gaps filled in by DNA polymerase priming from upstream DNA 3' ends. Removal of the 5'-most RNA primer generates an 8–12-base gap. Failure to fill in this gap leads to a small loss of DNA in each round of DNA replication. b) How telomerase solves the end replication problem.Telomere shortening after each round of replication

and resection of lagging strand by exonucleases, generates a 3' overhang which serves as primer for telomerase mediated telomere extension of the leading strand.This allows DNA replication machinery to fill in the lagging strand and overcome telomere shortening. Adapted from [41].

I.1.e: Telomerase solves the end replication problem

Pioneering experiments from Greider and Blackburn established that *Tetrahymena* cell extract could extend short telomeric DNA sequences *in vitro,* proving the existence of a special polymerase within the cells which was subsequently named 'Telomerase' [42].

We now know that telomerase is a large ribonucleoprotein (RNP) complex that synthesizes and maintains telomeres at the ends of linear chromosomes. The catalytic core of this complex contains an essential RNA component that serves as template for telomere synthesis (telomerase RNA – TER) and a unique reverse transcriptase protein (telomerase reverse transcriptase – Tert)[43, 44]. The widespread adoption of telomerase as the main strategy to counter telomere shortening illustrates the pivotal role of the enzyme in maintaining genomic integrity and cellular proliferation. Indeed, abnormal expression of telomerase components leads to cancer and wide variety of degenerative disorders as mentioned below.

I.1.f: Telomeres and telomerase associated human diseases

An early indicator that telomeres were crucial for cancer progression came from examining primary human fibroblasts grown in cell culture. With completion of ~60–80 population doublings these cells entered into senescence [45]. In contrast, established cancer lines divided indefinitely with passage in culture. A vital clue came with the

observation that telomeres shorten progressively during division of normal human fibroblasts in culture, yet are maintained in cancer cells [46]. Consequently, these shortened telomeres activate a 'senescence' program; however, this growth arrest phenotype can be neutralized through inactivation of p53 and Rb (the Retinoblastoma tumor suppressor protein) [47, 48]. Senescence in these fibroblasts is effectively averted by overexpression of telomerase reverse transcriptase (TERT) [49]. Considering these results, it is not surprising that highly recurrent mutations in TERT promoter, up-regulating telomerase levels, have been detected in wide variety of skin cancers [50, 51]. In essence, these observations suggest that high amounts of telomerase are essential for cancer development and progression, making it an attractive target for pharmaceutical manipulations and cancer treatment [52].

Equally important, shortening of telomeres has been shown to contribute broadly to aging & senescence [53]. In patients with germline mutations in genes controlling telomere maintenance, telomere shortening caused a range of disease phenotypes including dyskeratosis congenita (DC), pulmonary fibrosis and aplastic anemia. Dyskeratosis congenita is the most severe form of the 'telomere diseases' and is characterized by a classic triad of symptoms including abnormal skin pigmentation; nail dystrophy and oral leukoplakia [54-56]. DC causing mutations have been identified not only in the catalytic components TER [56] and TERT [57], but also other components of telomerase critical for stability [55, 58] and localization [59, 60]. To elucidate the mechanism behind dyskeratosis congenita, patient specific induced pluripotent stem cells (IPSc) have been generated. Unlike control samples, patient iPS cells displayed higher tendency to

differentiate and lesser ability to self renew [61]. Thus a possible mechanism of dyskeratosis congenita could be exhaustion of stem cells and their inability to divide due to short telomeres.

Although we are beginning to conceive how telomerase mutations causes degenerative disorders, it is not clear why different mutations in the same gene lead to disparate symptoms and disorders. For a complete understanding of telomere and telomerase associated disorders, it is imperative to elucidate how the correct amount of telomerase is made.

Although the catalytic activity of telomerase requires only TER and TERT, other proteins in the holoenzyme complex regulate complex assembly, trafficking, and stability. Consequently, cellular telomerase holoenzymes are multi-subunit complexes with an apparent mass of 500 kDa or more as estimated by gel fractionation studies[62, 63]. Though the telomerase holoenzyme proteins and their exact mechanisms of function differ between organisms, some of their biochemical roles bear significant similarities[62]. Next, I will describe the similarities and differences of telomerase RNA biogenesis and RNP assembly from ciliates, yeasts, and vertebrates.

I.2.a: Telomerase RNA biogenesis in Ciliates

Blackburn and Greider identified telomerase activity using extracts prepared from ciliates [42]. In addition, telomerase was first purified from *Euplotes aediculatus* by affinity chromatography which led to the identification of the catalytic subunit TERT [44] and a La-motif protein p43[43]. Though the function of p43 has been characterized only

biochemically [64-66] and its *in vivo* function is still unclear, it was crucial in identification of its homolog, p65 in *Tetrahymena thermophila,* which is necessary for TERT and TER accumulation *in vivo* [67].

Ciliate telomerase RNA is transcribed by RNA polymerase III [68] and mature TER terminates with a 3′ uridine-rich sequence. In *T. thermophila*, the TER 3′ polyuridine tail along with preceding RNA stems, provide a binding site for the telomerase-specific p65 protein. In addition, binding of p65 to telomerase RNA changes RNA conformation to promote TERT assembly and catalytic activity [69].

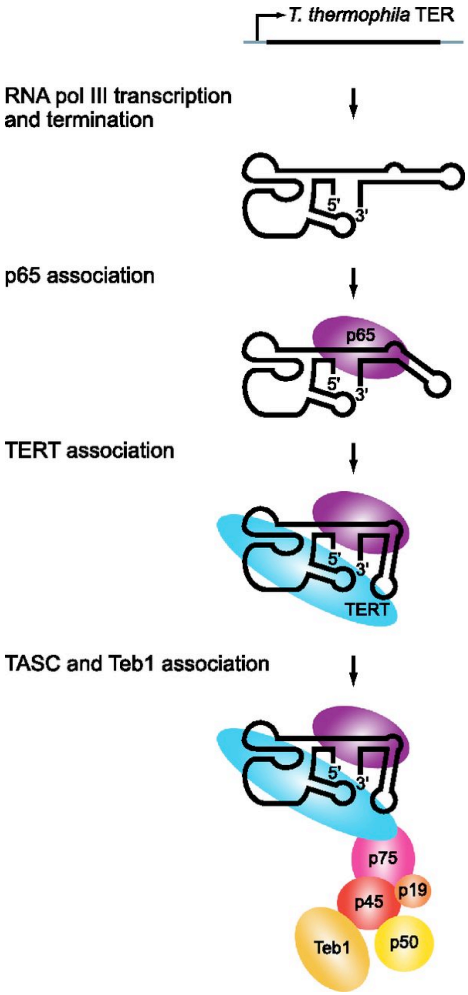


Figure 1.4: Schematic of telomerase RNA (TER) biogenesis in *T. Thermophila.* TER is transcribed by RNA polymerase III. The binding of p65 alters conformation of stem IV that enhances TERT binding. The p65-TER-TERT ternary complex then interacts with a complex of p75, p50, p45 and p19 that recruit the single-stranded telomeric DNA-binding protein Teb1, which provides a stable grip on DNA and promotes telomerase activity. Adapted from [62].

I.2.b: Telomerase RNA biogenesis in Yeast

In contrast to ciliates, budding yeast telomerase RNA (TLC1) is transcribed by RNA pol II and the mature form is non-polyadenylated [70]. Although, longer forms of TLC1 that are polyadenylated have been identified, their function remains elusive [70].

Surprisingly, the 3′ end of functional TLC1 is generated independently of the longer transcript. In this case, Nrd1–Nab3–Sen1 pathway is required for transcription termination and 3’ end formation of TLC1[71].

In general, the transcription termination pathway mediated by Nrd1-Nab3-Sen1 has been well studied in budding yeast[72]. Binding to the C-terminal domain (CTD) of RNA Pol II, Nrd1 and Nab3 recognize sequences in the precursor transcript downstream from the mature RNA 3′ end. They interact with the Sen1 helicase, which is thought to terminate transcription by unwinding the RNA–DNA hybrid in the RNA polymerase active site. The TRAMP (Trf4, Air2, Mtr4) complex then adds a short polyadenosine tail to the 3′ end, which is ultimately removed by the nuclear exosome.

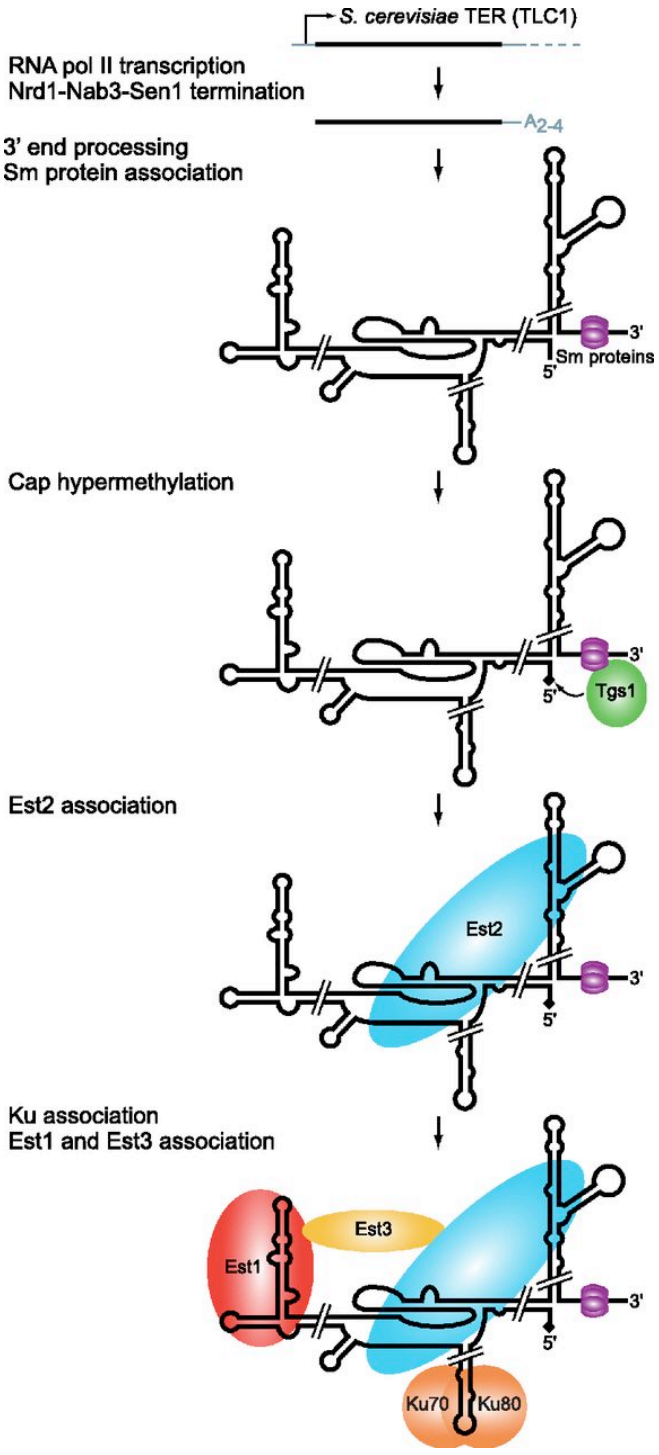


Figure 1.5: Representation of telomerase RNA biogenesis in *S. cervisiae.* Adapted from [62]. *S. cervisiae* telomerase RNA (TLC1) is transcribed by RNA polymerase II, and the Nrd1-Nab3-Sen1 pathway terminates transcription. Binding of Sm proteins stabilize the

3'end and recruits Tgs1 which hypermethylates the cap structure. Est2 (TERT) binds directly to the cap structure. In addition, Ku heterodimer is required for TLC1 stability and nuclear import. Along with the cell cycle dependent Est1, Est3 is required for telomerase function at telomeres.

Similar to small nuclear RNA (snRNA), TLC1 is bound by Sm proteins (SmB, SmD1, SmD2, SmD3, SmE, SmF, and SmG), which forms a heterohepatmeric ring essential for TLC1 stability. Additionally, Sm proteins recruit Tgs1, the methyl transferase, hypermethylating the cap [73].

Studying TLC1 trafficking has provided critical insights regarding the dynamics of RNP assembly and its role in telomere length regulation. Unlike humans, where TMG capping has been established to occur in the cytoplasm for small nuclear RNAs[60], TLC1 hypermethylation occurs in the nucleolus [74]. During S phase, TLC1 is localized to a few clustered telomeres which are speculated to be elongated [75].

I.2.c: Telomerase RNA Biogenesis in Humans

Unlike yeast and ciliate TER, human telomerase RNA (hTR) comprises of a H/ACA box at its 3' end which is common feature of guide RNAs that modify snRNA and snoRNA [60]. In humans, H/ACA RNAs are processed from introns, which are excised from protein coding genes during precursor splicing [76]. In contrast, hTR is transcribed from its own promoter by RNA pol II. Similar to telomerase RNA in other species, mature hTR is non-polyadenylated and TMG capped [62, 77]. Although mature hTR is only 451nt in length, longer polyadenylated forms have been detected but their function in hTR processing has remained elusive (Gaspari M & Baumann P, unpublished). The presence of a downstream polyadenylation signal or U1 snRNA 3′ box/transcription

terminator antagonizes hTR accumulation [77], leaving the actual mechanism of nascent transcript termination and/or endonucleolytic cleavage an open question. The importance of deciphering human telomerase RNA processing is highlighted by the observation that certain mutations in hTR, which lead to accumulation of precursor, causes dyskeratosis congenita in patients (Gaspari M and Baumann P, unpublished data). In other words, elucidating hTR processing has critical implication for DC treatment and therapy.

Though not much progress has been made in the field of hTR processing, focusing on hTR dynamics has yielded fruitful results. Human telomerase RNA localizes to Cajal bodies via its association with TCAB1/WDR79, which recognizes a Cajal body localization signal in hTR called the CAB box [78]. Cellular depletion of TCAB1/WDR79 does not affect hTR accumulation but does reduce RNP association with Cajal bodies and telomeres and results in telomere shortening. Above all, the significance of TCAB1/WDR79 for telomere maintenance is highlighted by the discovery of TCAB1/WDR79 gene mutations that cause dyskeratosis congenital [59].

Although the different model organisms have been extremely useful in elucidating telomerase function in telomere repeat addition, we still know little about telomerase RNA biogenesis in humans. Since the small nuclear RNAs display remarkable conservation between *S. pombe* and humans, we decided to study telomerase RNA maturation in *S. pombe*. Before, we discuss this further; it is necessary to understand spliceosome assembly and function.

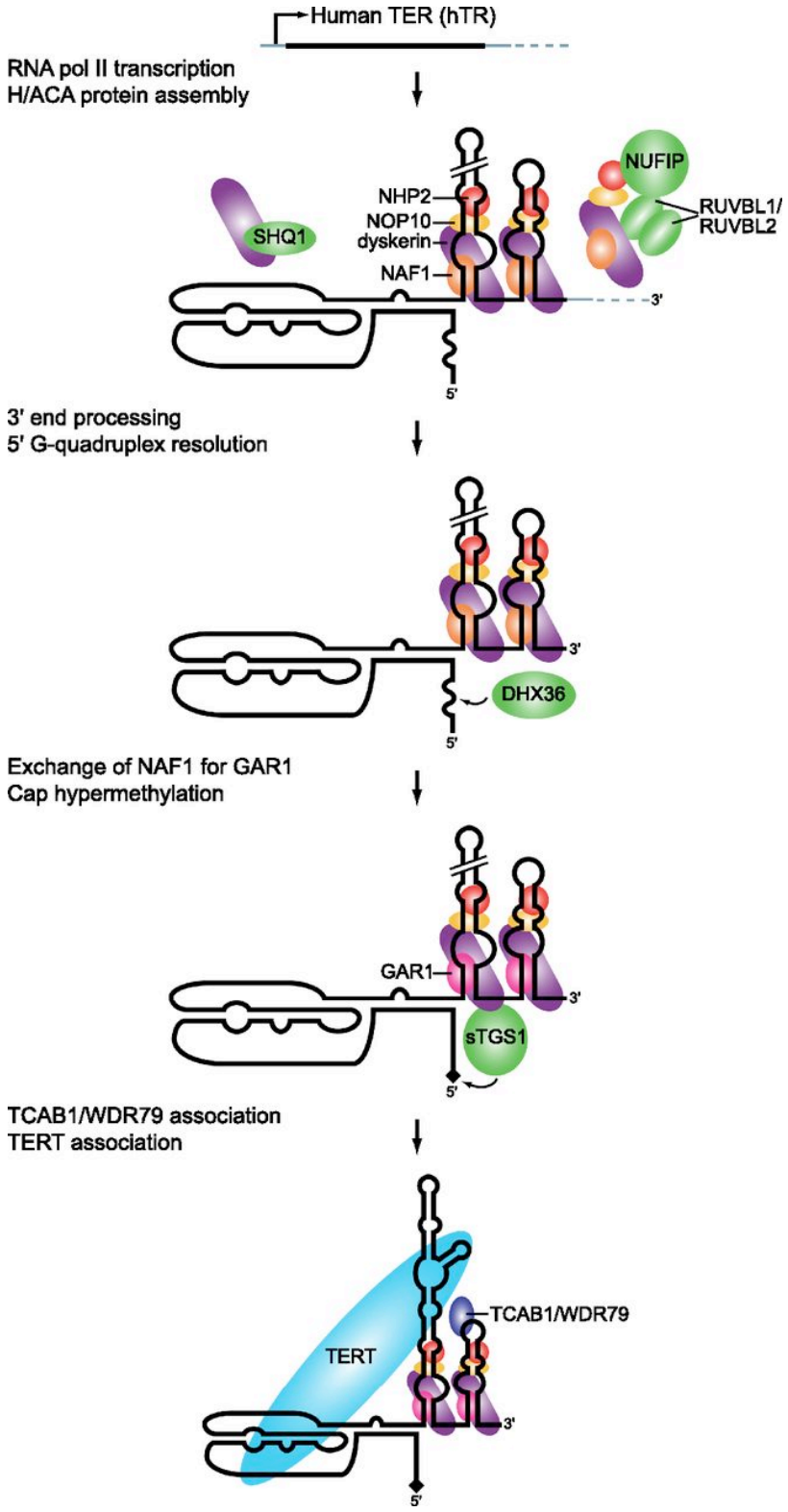

Human TER (hTR)
RNA pol II transcription
H/ACA protein assembly
SHQ1
NHP2
NOP10
dyskerin
NAF1
NUFIP
RUVBL1/
RUVBL2
3'
5'
3' end processing
5' G-quadruplex resolution
DHX36
Exchange of NAF1 for GAR1
Cap hypermethylation
GAR1
sTGS1
TCAB1/WDR79 association
TERT association
TERT
TCAB1/WDR79

Figure 1.6: Schematic of human telomerase RNA biogenesis. Adapted from[62]. 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 poorly understood processing mechanism. Dyskerin, NAF1, NHP2, NOP10 bind hTR and are required for its accumulation *in vivo*.

I.3: Introduction to splicing

Appreciation of the central role that RNA plays in controlling gene expression began with the advent of new methodologies and biological insights in the early 1970s and 1980s [79]. The initial transcripts generated by RNA polymerase II are long, inactive, heterogeneous nuclear RNAs (hnRNAs, now termed pre-mRNAs) that are precursors for functional shorter mRNAs which are exported to cytoplasm for translation [80]. Seminal studies established that processing of adenovirus pre-mRNA to mature mRNA involves removal of non-coding sequences (introns) and splicing together of the coding sequences (exons) [81, 82]. Although splicing was initially discovered in adenovirus, we now know that all eukaryotes from yeast to human display this conserved mode of RNA processing. In case of higher eukaryotes, a diverse array of proteins is generated from the same gene by alternative splicing of exons [83]. Unraveling splicing at the molecular level is important not only for understanding gene regulation but also has a medical impetus, as many human diseases are caused by defective pre-mRNA splicing [84].

Nuclear pre-mRNA splicing is catalyzed by a multi-mega dalton ribonucleoprotein (RNP) complex called the spliceosome. Two unique spliceosomes co-exist in most eukaryotes: the U2-dependent spliceosome (major) which catalyzes the removal of U2-type introns and the less abundant U12-dependent spliceosome which is present in only a subset of eukaryotes and splices the rare U12-type class of introns [85]. For the

remainder of the introduction, I will discuss the literature on the U2-dependent spliceosome as this is most relevant to the thesis.

I.3.a: Chemistry of splicing

The spliceosome removes an intron in two consecutive reactions. In the first reaction, the 2′ hydroxyl of a conserved, intronic adenosine attacks the phosphate at the 5′ splice site, producing a free 5′ exon and a branched species, termed the lariat intermediate. In the second reaction, the 3′ hydroxyl of the 5′ exon attacks the phosphate at the 3′ splice site, yielding ligated mRNA and a lariat intron [86, 87]. The intermediates and products of pre-mRNA splicing are similar to those generated during group II intron splicing [87].

I.3.b: Sequence elements required for splicing

Three core elements namely the 5' splice site, branch site and the 3' splice site defines the intron (Fig. 1.7A). The consensus 5'ss is **GU**AUGU in budding yeast, **GU**A (A/U) GU in fission yeast and **GU**RAGY in mammals [88]. The branch site consensus (UACUA**A**C) is highly conserved in budding yeast and any deviations from this site lead to defects in splicing. In contrast, fission yeast introns have a degenerate consensus branch site (CUR**A**Y), similar to mammalian introns [89]. Despite the varying length of introns between yeast to humans, position of the branch site seems to be under strict evolutionary pressure with the branch point adenosine located 18-40 nt upstream of the 3'ss [90]. Although, increasing the distance between branch point and 3'ss impairs splicing, the significance of the distance constraint is not clear [91]. Apart from the conserved 5'ss and branch site which are necessary for the first catalytic step in all

eukaryotes, the polypyrimidine tract is essential for splicing in metazoans [92]. In contrast, a functional 3'ss is dispensable for the first catalytic step in most introns with certain exceptions [92, 93]. For the second step of splicing, the conserved AG dinucleotide is the only essential feature that is conserved in all species. The nucleotide preceding AG occurs in the frequency of U/C >A > G[94].

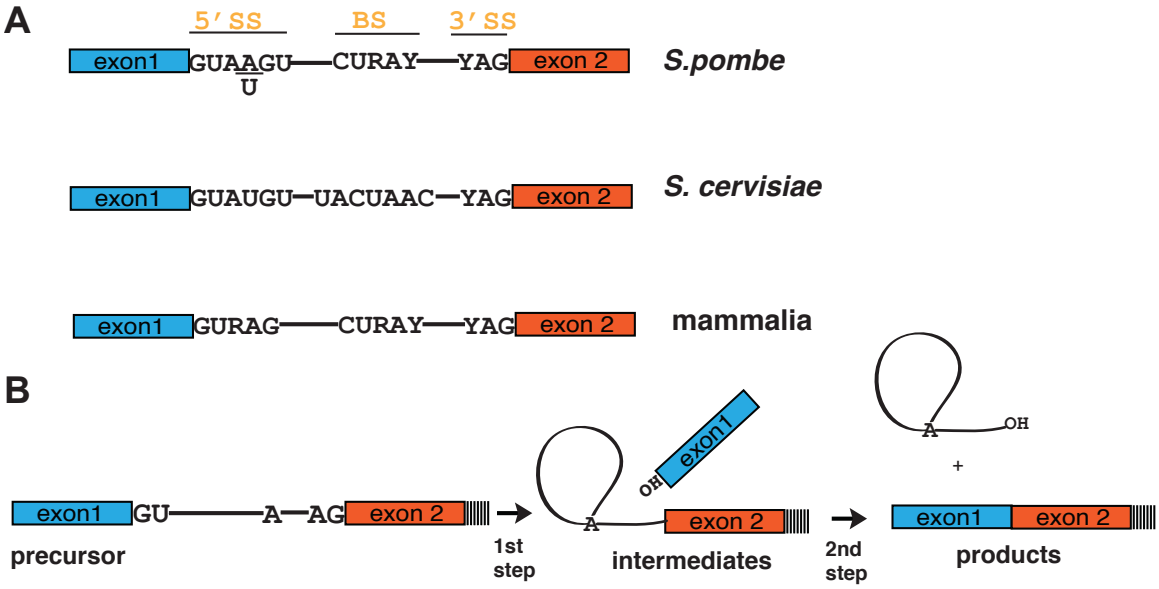


Figure 1.7: Depiction of the sequence motifs required for splicing. A) The schematic shows the conserved motifs of 5'ss, branch site and 3'ss present in *S. pombe, S. cerevisiae* and mammalian introns respectively. (Exons- blue and red boxes and lines-introns; R and Y represent purines and pyrimdines respectively. B) The two steps of splicing are shown. Adapted from [88].

Additional splicing signals include short and diverse sequences called enhancers or silencers which are located in exons and introns. These recruit trans-acting factors to promote or repress splicing [90]. Although a lot of cis-elements that regulate splicing of an intron have been characterized, little is known about sequences that promote the first

step over the second. Therefore, one of the main objectives of my thesis was to characterize such sequences.

I.3.c: Spliceosome Assembly and Rearrangement

To compensate for the limited information present in the precursor, a large number of trans-acting factors that constitute the spliceosome interact with the sequences in the intron to spatially arrange the reactive chemical groups required for splicing in an optimum position. Although there are species-specific variations, the core machinery consists of five snRNPs namely U1,U2,U5 and U4/U6, which are conserved from yeast to mammals. With the exception of U6, each snRNP is composed of a snRNA that is transcribed by RNA pol II and a group of seven Sm proteins (B/B',D3,D2,D1,E,F,and G) which are critical for snRNA stability & for modifying the cap structure of snRNA [60]. Surprisingly, U6 snRNA is different from the other spliceosomal snRNAs as it is transcribed by RNA pol II, is not TMG capped and is bound by Lsm2-8 complex [60]. Why U6 snRNP biogenesis is distinct from other snRNPs is not yet known.

Since the sequence information present in the precursor is limited, the spliceosome has to differentiate real splice sites from decoy sites. Accuracy of splicing is ensured by stepwise recognition of the intron in a mutually exclusive manner. Spliceosome assembly begins with U1snRNP recognizing and base pairing with 5'ss. Following U1, U2snRNP binds to the intron by base pairing with the sequences that flank the branch point (bp) adenosine. However, the adenosine residue itself doesn't base pair with U2, but instead

bulges out of the U2 helix, and is used as a nucleophile for the first catalytic step. After U2 snRNP binding, U4/U6/U5 bind the precursor RNA as a tri-snRNP.

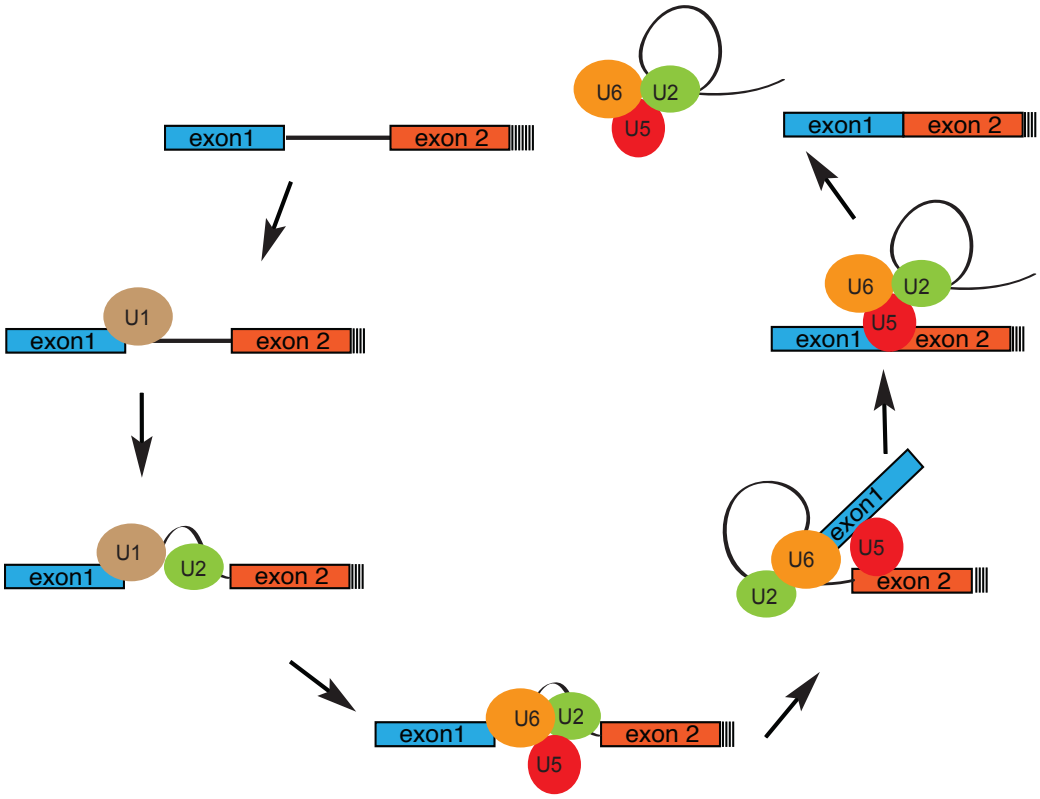


Figure 1.8: Stepwise assembly of spliceosomal snRNAs. Note that only the snRNA binding and transition is shown here. Adapted from [95]

U4/U6 are extensively base paired, but binding of the tri-snRNP to the RNA triggers the release of U6 from U4. As a result, U6 displaces U1 from the 5' ss and the spliceosome is said to be catalytically active for the first step. This ordered assembly of snRNPs is conserved from yeast to humans and prevents the spliceosome from cleaving precursor RNAs prematurely.

I.3.d: snRNAs required for catalysis

Despite the pivotal role of U1 snRNP in early spliceosome assembly, only U2, U5 and U6 are required for the two catalytic steps of the spliceosome. In fact, splicing in the absence of U1 and U4 has been observed [96-98]. Additionally, U2 and U6 base pair with each other and these interactions are important for the first catalytic step. U6, which is at the heart of the splicing reaction, interacts with the last 3 nucleotides of most 5'splice sites (GUA**UGU**), through the first 3 nucleotides of its conserved **ACA**GAG box. In contrast to the U6 and U2 binding regions, the 5'ss and branch site are poorly conserved among higher eukaryotes. Why this fervent anti-correlation exists between the spliceosomal snRNAs and its binding sites has remained elusive.

As the chemical steps of splicing occur, the spliceosome changes its conformation from the first to the second step of splicing to position the reaction intermediates optimally for the 2nd transesterification. However, the precise nature and timing of these events is not completely understood. Recent data have demonstrated that base pairing of the 5'ss with U6 must be disrupted prior to the second step of splicing [99]. Hence, this indicates a requirement for U6 to shift its position after the first step, which produces a free exon and lariat intermediate. Further, it has been demonstrated that the U2-branch site base pairing can be disrupted and is not essential for the 2nd step of splicing [100]. Whether U2-BS is actually detrimental to the 2nd step of splicing has not been elucidated.

Unlike U6 and U2, U5 snRNA does not bind any sequence motif in the intron. In contrast, U5 plays an important role in aligning the two exons for the second catalytic

step. Furthermore, probing the requirements for U5 snRNA binding has revealed optimal sequences at the intron-exon boundaries that enhance splicing [101].Though splicing has been very well studied, not much is known about sequences and trans-factors required for the second step of splicing. After the second step of catalysis, the spliceosome releases the ligated RNA and lariat intron.

I.3.e: Kinetic proofreading hypothesis

Faithful expression of genes requires selection of cellular machineries that have high substrate specificity. This paradigm is inherently true for pre-mRNA splicing, where the spliceosome has to distinguish the real splice site from a myriad of choices [102].Since introns predominate over exons in eukaryotes, it is imperative to ensure splicing fidelity. Indeed, recent estimates of splicing errors range from 1 in 100 to 1 in 10000 [103, 104].

Forty years ago, Hopfield and Nino proposed a model to explain how specificity is achieved in translation and termed it 'kinetic proofreading' [105, 106]. This concept has been validated in translation and other systems [107, 108] including splicing [109, 110]. A critical feature of the kinetic proofreading hypothesis is a branch in the pathway that competes with a productive step and specifically antagonizes suboptimal substrates.

Christine Guthrie and colleagues extrapolated the kinetic proofreading hypothesis to explain the role of DExH/D ATPases in spliceosome function and splicing fidelity. From a genetic screen to identify splicing fidelity factors in budding yeast, Prp16 emerged as an ATPase that facilitates the transition of the spliceosome from the first- to the second-step conformation [109]. Mutations in Prp16 improve the first step of splicing for

suboptimal substrates by lowering the rate of ATP hydrolysis by Prp16, altering its binding to the spliceosome, or both. Although the way in which Prp16 affects the first-to-second step transition is not fully understood, mutations in the ATPase domain are thought to slow the rate of exit from the first-step conformation. Kinetic competition between catalysis of the first step and substrate rejection serves as the basis for the control of fidelity by Prp16-mediated ATP hydrolysis, also known as the kinetic-proofreading hypothesis (Fig. 1.9).

Since then numerous DExH/D-box ATPase/helicases have been identified that participate in the splicing process, and each is thought to facilitate a structural transition by coupling ATP hydrolysis to an interdomain movement that results in remodeled RNA-RNA and/or RNA-protein interactions. These structural transitions occur throughout the assembly of the spliceosome, the two consecutive transesterification reactions and the product release and disassembly phase [102].

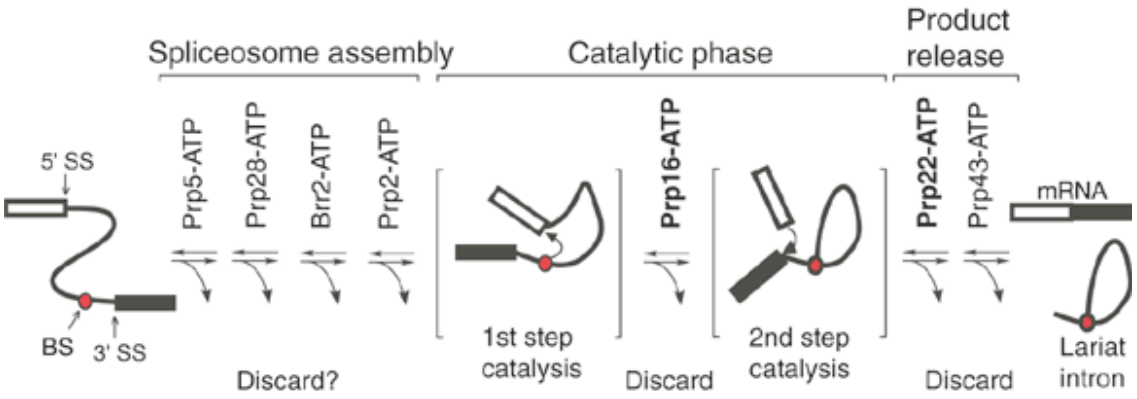


Figure 1.9: Illustration of kinetic proof reading by various DExH box helicases at each step of splicing. Adapted from [111]

1.3.f: A 'discard pathway' for splicing

Prp22 is the DExH/D box helicase that uses its ATPase activity to alter conformation of the spliceosome that allows release of the spliced form. Usually, the 2nd step of splicing is efficient and occurs before the ATP hydrolysis activity of Prp22. After exon ligation, Prp22 changes conformation of the spliceosome in a ATP dependent manner and releases the spliced form in a Prp43 dependent manner [112]. In contrast, 3'ss mutations are suboptimal for the 2nd step of splicing. In this scenario, Prp22 ATP hydrolysis occurs prior to the 2nd step of splicing, releasing the 5' exon and lariat intermediate.

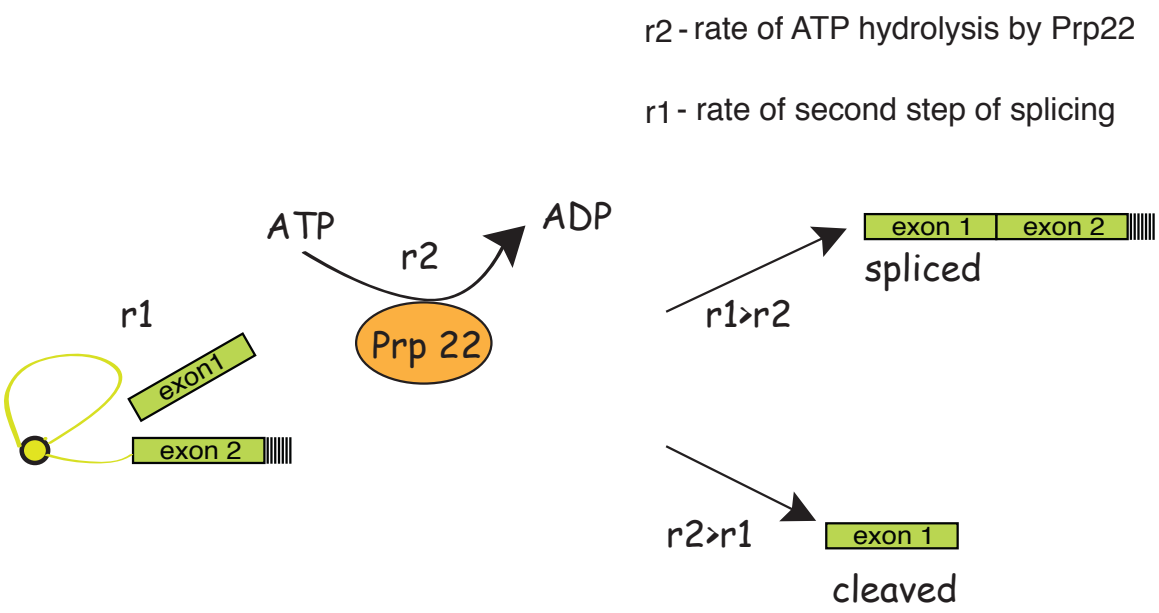


Figure 1.10: The discard pathway of splicing. If the rate of ATP hydrolysis by Prp22 is slower than the first step, then Prp22 releases ligated exons in a Prp43-dependent manner. However, if the rate of ATP hydrolysis by Prp22 is faster than the first step, then Prp22 releases the intermediates, i.e. the free 5' exon and intron lariat, in a Prp43-dependent manner.

Surprisingly, purified spliceosomes from budding yeast, in the absence of ATP no longer exhibit this inhibition. Similarly, Prp22 mutants that slow down ATP hydrolysis allow splicing of suboptimal substrates [113]. Likewise, Prp43 has been implicated in the discard of rejected second step intermediates [112]. In conclusion, it is evident that Prp22 and Prp43 prevent splicing of suboptimal 3'ss using ATP hydrolysis as a timer.

I.3.g: Limitations of splicing studies

The discovery that genes are not continuous but rather split by intervening sequences, which must be removed by splicing was first discovered in viral mRNAs. Consequently, the majority of studies have focused on understanding intron removal in the context of protein coding genes, where release of the 5' exon after the first step of splicing might lead to degradation of splicing intermediates and hinder gene expression. For this reason, it is not surprising that in protein coding genes, the two steps of splicing are coupled [114].

I.4: Telomerase RNA biogenesis in *S. pombe*

I.4.a: Spliceosomal cleavage generates mature TER1

Telomerase RNA (TER1) in *S. pombe* is initially transcribed as a precursor containing an intron flanked by two exons [115, 116]. Similar longer forms have been identified in budding yeast and humans, but their function remains unclear [70, 117]. However, the mature form of TER1 that predominantly associates with active telomerase ends precisely at the 3' end of the 5'exon. Surprisingly, functional telomerase RNA in *S. pombe* (TER1) is generated by only the first step of splicing [116]. I refer to this single step cleavage

event as ‘spliceosomal cleavage’. Consistent with the first step of splicing being required for TER1 processing, mutations in the 5’ss and branch site blocked TER1 accumulation. In contrast, 3’ss mutations did not alter TER1 levels. A direct role for the spliceosome in TER1 processing became clear when mutations in 5’ss were rescued by complementary changes in the U1 binding site that rescues base pairing. The idea that the second step of splicing is detrimental for TER1 maturation was evident when TER1 intron was replaced with a heterologous intron from protein coding gene, which compromised TER1 accumulation. In contrast, mutations in the 3’ss of the heterologous intron rescued TER1 levels. It is unknown as to how TER1 could utilize the first step of splicing but escape the second step for its 3’ end processing. As the TER1 intron has a very good 5’ splice site (ss) and branch point (bp), it undergoes an efficient first step. An unusually long distance of 22 nt between the branch site and 3’ss may lead to an inefficient second step. For example, it has been shown that increasing the distance between the branch site and the 3’ss of the cdc2 intron leads to a significant decrease in accumulation of the spliced form [91]. Indeed, shortening this distance by 14 nt in the TER1 intron favors the accumulation of spliced form over the cleaved form and telomeres shorten. These results demonstrated that unlike canonical splicing, telomerase RNA maturation requires only the first step of splicing (Fig 1.11).

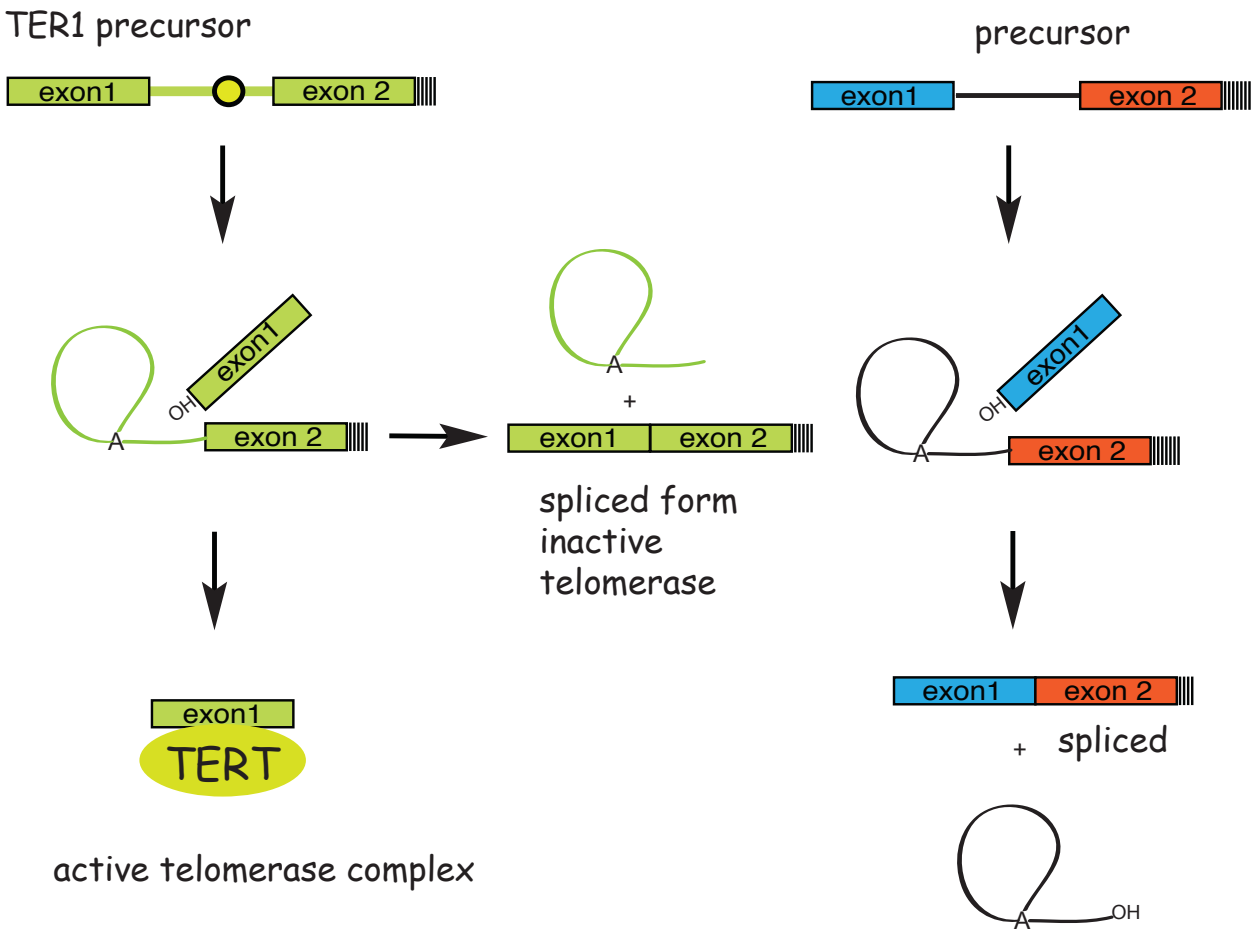


Figure 1.11: Representation of the difference between spliceosomal cleavage and canonical splicing. Canonical splicing involves tight coupling of the two transesterification reactions (Right). Spliceosomal cleavage generates only the first step of splicing followed by release of 5' exon that forms the active telomerase complex in fission yeast (Left).

I.4.2: Sm proteins and Lsm proteins are required for telomerase RNA biogenesis

Directly upstream and partly overlapping with the spliceosomal cleavage site is a putative binding site for Sm proteins. Sm and like-Sm (LSm) proteins belong to an ancient family of RNA-binding proteins represented in all three domains of life. Members of this family form ring complexes on specific sets of target RNAs and have critical roles in their

biogenesis, function and turnover. In a manuscript that I co-authored with my colleagues we demonstrated that the canonical Sm ring and the Lsm2–8 complexes sequentially associate with fission yeast TER1 [118]. The Sm ring binds to the TER1 precursor and promotes the hypermethylation of the 5′-cap by Tgs1. Sm proteins are then replaced by the Lsm2–8 complex, which promotes the association with the catalytic subunit and protects the mature 3′-end of TER1 from exonucleolytic degradation. Our findings define the sequence of events that occur during telomerase biogenesis and characterize roles for Sm and Lsm complexes as well as for the methylase Tgs1.

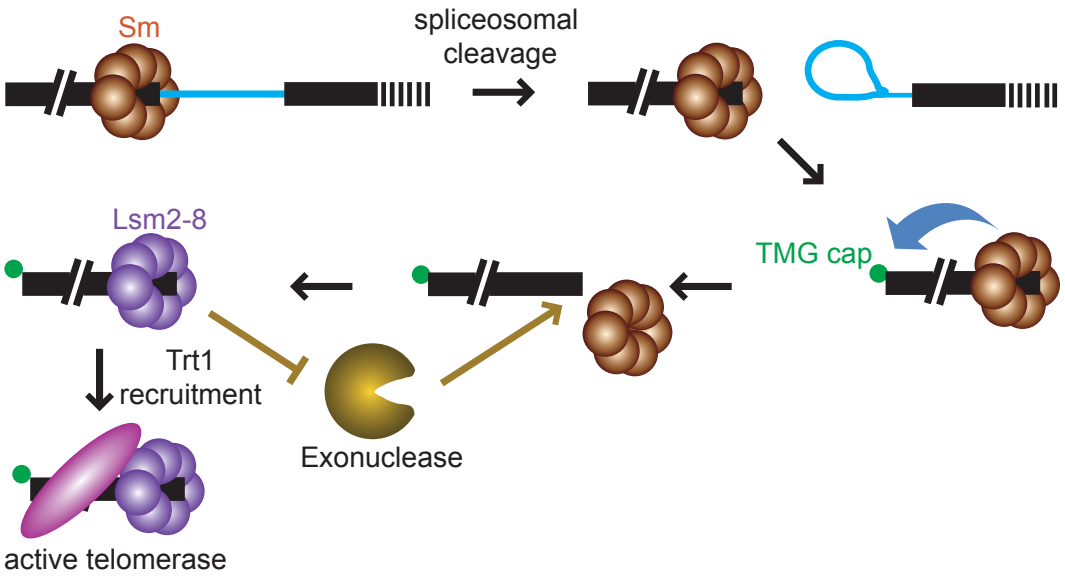


Figure 1.12: Sm and Lsm proteins are required for telomerase RNA Biogenesis in *S. pombe.* (adapted from **[118]**).

I.5: Scope of dissertation

The primary goal of my thesis project was to elucidate the mechanism and function of spliceosomal cleavage. Specifically, my aims were to determine the cis-acting elements

and trans-acting factors that inhibit the 2nd step of splicing. Based on these findings, I wanted to identify other RNAs processed by spliceosomal cleavage in fission yeast.

By precisely dissecting out the function of several spliceosomal components and their associated cis-acting elements, we have uncovered novel roles for U2 snRNP and U2AF59 in regulating the second step of splicing. In addition, our results invoke a critical role for the Prp22 and Prp43 mediated 'discard' pathway in TER1 biogenesis. Results of this aim are summarized in Chapter Three of my thesis.

Furthermore, during the course of investigating the function of uncoupling elements identified in Chapter Three, we performed a global analysis of *S. pombe* introns to identify other introns that could be cleaved. Our analysis has provided new insights on the mechanism of splicing and indicates a possible role for the spliceosome in regulating gene expression. Pursuing this further, we established that 3' end of telomerase RNA in other fission yeast species are also generated by spliceosomal cleavage. The significance and implications of these findings are presented in Chapter Four.

Chapter Two of this dissertation consists of a comprehensive list of materials and methods used to execute the research described in Chapters Three and Four. Chapter Five provides a synthesis of the information presented in Chapters Three and Four with current literature. In conclusion, Chapter Five discusses the numerous future directions that are opened up in the field due to the core results obtained in Chapter Three and Four.

Chapter Two:

Materials and Methods

II.1: Yeast strains and constructs

Strain Number	Genotype	Source
PP138	h^- ade6-M216 leu1-32 ura4-D18 his3-D1	Lab stock
PP265	h^-	Lab stock
PP407	h^+/h- leu1-32/leu1-32 ura4-D18/ura4-D18 his3-D1/ his3-D1 ade6-M210/ade6-M216 ter1+/ ter1::kanMX6	Box etal, Nature 2008
PP433	h+/h- leu1-32/leu1-32 ura4-D18/ura4-D18 his3-D1/ his3-D1 ade6-M210/ade6-M216 ter1+/ ter1::ura4	Wen et al, Nature 2011.
PP491	h+, leu1-32, ura4, ade6-M216, rrp6::kanMX6	Lab stock
PP550	h^-, leu1-32, ura4-D18, ade6-M210	Haraguchi et al.; 2007
FP953	h^- ade6-M216 leu1-32 ura4-D18 his3-D1, prp22::kan (pSH7)	This study
S. cryophilus		Lab stock
S. octosporus		Lab stock

Genotypes of the strains used in this study are listed in the table above. The basic reporter was generated by replacing the *ura4* open reading frame in pDblet [119] with a synthetic DNA fragment composed of the *ura4* open reading frame followed by 14 nucleotides of

TER1 exon 1 sequence (gggcccauuuuuug), the TER1 intron and the TER1 second exon using AvrII and StuI restriction sites. Derivatives of this construct were generated by site-directed mutagenesis or subcloning. Plasmids were introduced into *S. pombe* strains PP138 and PP407 by electroporation and transformants were selected on Edinburgh Minimal Media (EMM) lacking uracil. For genomic integration, *ter1::ura4* spores were generated from PP433 [118], germinated on YEA and selected on EMM lacking uracil. A DNA fragment containing the mutations in the context of a genomic copy of TER1 was amplified from plasmid DNA by PCR using primers PBoli923 (gtaaacggaatatccgcgatg) and BLoli1122 (ttccatatagtcgatgctcg). The PCR product was purified and 1µg was used for lithium acetate transformations as described [120]. Transformants were selected on media containing 5' FOA and strains containing the correct integration were identified by colony PCR and sequencing.

Replacing one copy of prp22+ in a diploid strain with the kanamycin resistance marker generated the prp22 knockout strain. The heterozygous *prp22*$^{+}$ / *prp22::*kan^{R} strain was transformed with plasmid pSH3 containing wild type *prp22*$^{+}$ and the *ura4* gene. The diploid transformants were sporulated and *prp22::*kan^{R} cells containing pSH3 were selected. The resultant strain was then transformed with a plasmid derived from pBG1 [121] containing *prp22*$^{+}$ or point mutants and the *his3* gene as a selection marker. Media containing 5-fluoroorotic acid and lacking histidine was used to select for cells that have lost pSH3 and now harbor a *prp22* mutant or *prp22*+ on the pBG1-derived plasmid. The presence of each *prp22* mutant allele was confirmed by PCR and sequencing. Temperature and cold sensitivity were examined by incubation at 36°C and 22°C

respectively. The same procedure was employed to generate temperature sensitive *prp43* strains which contained *prp43* mutants on a plasmid derived from pBG1 in a *prp43::*kan^R background.

II.2: RNA analysis

Cultures (400 ml) were grown to a density of 5 x 10^6 cells/ml and cells were collected by centrifugation in JA 10 rotor (Beckmann) for 4 min at 4000g [116]. Cell pellets were washed once with cold water and quick frozen as beads by dripping into liquid nitrogen. Cells were lysed in a 6850 Freezer mill (SPEX SamplePrep) using 8 cycles (2 min) at a rate of 10 per second with 2 min cooling time between cycles. After grinding, the lysed powder was transferred into 50 ml falcon tubes containing 10 ml of phenol/chloroform/isoamyl alcohol (25:24:1) and 10 ml sodium acetate (50 mM), 1% (w/v) sodium dodecyl sulphate preheated to 65 °C. The purity of RNA was increased by extracting RNA at least 5 times with phenol/chloroform/isoamyl alcohol and once with and once with chloroform/isoamyl alcohol. The supernatant containing RNA was subjected to ethanol precipitation and resuspended in 50 mM Sodium Acetate (pH 5.2)

The following modifications were made for temperature sensitive strains. Cultures (400 ml) of U2AF59 ts strains were grown at 25°C to a density of 5 x 10^6 cells / ml. The culture was shifted rapidly to 36°C by immersion and agitation in warm water. Once the culture medium had reached 36°C, a 100 ml sample was taken for RNA isolation. Cultures were then transferred into a shaker at 36°C and additional samples were harvested at the indicated time points. Similarly, *prp22* and *prp43* mutant strains were grown into log phase at 32°C and shifted rapidly to 22°C. The point at which the culture

medium reached 22°C is referred to as 0 minute. To isolate RNA, 3 or 5 rounds of phenol chloroform extraction were performed depending on whether the sample volume was 100 ml or 400 ml. After the final chloroform isoamyl alcohol extraction, the supernatant containing RNA was ethanol precipitated at -20°C for at least one hour. Total RNA was precipitated by centrifugation and the pellet was dissolved in 50 mM sodium acetate at pH 5.2. Except for the northern shown in Fig. 3.1B, all samples were subjected to RNaseH cleavage prior to northern analysis to improve resolution of precursor, spliced and cleaved forms. Briefly, 15 µg of DNase treated RNA were combined with two DNA oligos (600 pmol) complementary to sites in the first and second exon, respectively. BLoli2269 (AACATCCAAGCCGATACCAG) and BLoli2326 (GCAAACAAGGCATCGACTTTTTCAATAACCAACCAAAA) were used for constructs containing the ORF and 3' UTR of ura4. BLoli2326 was replaced with BLoli1275 (CGGAAACGGAATTCAGCATGT) for constructs containing TER1 sequences downstream of the intron. Oligonucleotides BLoli3695 (TGGGCAATTGTATTCTCTGC), BLoli3563 (AAAAGCCAACGTAAATGCTG) and oligo-dT were used for RNaseH cleavage of ATP23 mRNA. The mix was heated to 65°C in a heat block and a second heat block at 75°C was placed on top of the tubes to reduce condensation. After 5 min, the heat block sandwich was transferred to a Styrofoam box to allow slow cooling to room temperature. After 45 min of slow cooling, RNaseH buffer (NEB, final concentration 1 X) and RNaseH enzyme (5 units) were added and samples were incubated at 37°C for 30 min. RNaseH treated samples were ethanol precipitated for at least one hour at -20°C. RNA pellets were recovered by centrifugation, dissolved in 1X

formamide loading buffer and run on a polyacrylamide gel for northern analysis as described below.

II.3: Northern Blot

Samples were run on 4% polyacrylamide gels in Tris-borate-EDTA (TBE) containing 7M Urea and transferred to a Biodyne Nylon transfer membrane (Pall Corporation) in TBE buffer at 0.4 A for 1 h in 0.5× TBE buffer. RNA was ultraviolet cross-linked (254 nm, 120 mJ) in a Stratalinker (Stratagene). Hybridizations with radiolabelled probes were carried out in Church-Gilbert buffer at 55 °C (TER1 probe) or 42 °C (reporter and sn101 probes). A TER1 probe recognizing the RNaseH cleavage products was generated by nick-translation of a PCR fragment (nucleotides 536 to 998 of TER1) in the presence of ^{32}P-α-dCTP. The ura4 open reading frame was detected using BLoli2072 (GTCTTTGCTGATATGCCTTCCAACCAGCTTC) and the 3' UTR with BLoli2318 (GAAACCGGAAACGGAATTCAGCATGTTTTAATAAAAAGAT), each labeled with polynucleotide kinase in the presence of ^{32}P-γ-ATP. BLoli1136 (CGCTATTGTATGGGGCCTTTAGATTCTTA) complementary to snRNA101 was used to visualize a loading control. ATP23 mRNA was detected on northern blots using a nick-translated PCR fragment generated with primers BLoli3693) and BLoli3592 (AGGAATTGAACACTTCTTCAACGAT) as probe and a hybridization temperature of 55 °C.

II.4: RT–PCR

DNase-treated RNA samples (2.5 µg) were incubated with oligonucleotide BLoli1275 (CGGAAACGGAATTCAGCATGT, 10 pmol) and dNTP mix (10 nmol). The total volume was adjusted to 13 µl with distilled water and samples were heated to 65 °C for 5 min and allowed to reach room temperature by slow cooling. The volume was increased to 20 µl by the addition of RNasin (40 U, Promega), dithiothreitol (5 mM final), 5× first strand buffer, and Superscript III reverse transcriptase (200 U, Invitrogen). Control reactions did not include SuperscriptIII to monitor for DNA contamination. After incubation at 55 °C for 60 min, RNaseH (5 U, NEB) was added and incubation was continued at 37 °C for 20 min. Aliquots (2 µl) of the RT reactions were used for PCR amplification with Taq polymerase (NEB) and oligonucleotide primers BLoli1020 (CAAACAATAATGAACGTCCTG) and BLoli1275 under the following conditions: 5 min at 94 °C followed by 30 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. PCR products were analysed by electrophoresis on 1.8% agarose gels. For splicing analysis of FCP1 intron1, Bloli 2149 (ATTGAAATCGCGAGCTGTCT) and Bloli 2150 (TTCCACAAATTGCACAGAGG) were used.

II.5: Determination of 3' end sequences

DNase-treated total RNA samples (2.5 mg) were treated with poly (A) polymerase (US Biologicals), RNase inhibitor (RNasin, 40 U), and ATP (0.5 mM) in the buffer provided by the manufacturer in 20 µl reactions and incubated at 30 °C for 30 min. The reaction volume was increased to 35.5 µl by the addition of the DNA oligonucleotide PBoli560 (GCGGAATTCT18, 125 pmol), dNTP mix (25 nmol), and the reactions were incubated

at 65 °C for 3 min followed by slow cooling to room temperature (20 °C). The reaction volume was then adjusted to 50 ml with first strand buffer (Invitrogen), dithiothreitol (5 mM), RNasin (40 U) and Superscript III reverse transcriptase (200 U, Invitrogen), and reactions were incubated at 50 °C for 60 min. This was followed by addition of RNaseH (5 U, NEB) and incubation was continued at 37 °C for 20 min. Aliquots (2.5 ml) of this reaction were used for PCR amplification with Taq polymerase (5 U, NEB), and oligonucleotide PBoli517 (GCTTTTGGCTGAAATGTCTTCC) specific for *ura4* and PBoli560 (200 nM each) under the following conditions: 3 min at 94 °C followed by 30 cycles of 30 s at 94 °C, 45 s at 55 °C and 120 s at 72 °C, followed by 7 min at 72 °C. PCR products were separated by electrophoresis on 0.8% agarose gels, and bands of the correct size were excised, purified and cloned into the TOPO TA cloning system (Invitrogen) for sequence analysis.

II.6: Spotting Assay

Cultures were grown to log phase ($5x10^6$ cells/ml) in EMM supplemented with uracil (150 mg/L), leucine (150 mg/L), and adenine (75 mg/L). 3-fold dilutions of samples were plated on PMG (Pombe Minimal media with Glutamate) in triplicate using an epMotion automated pipetting system (Eppendorf) and plates were incubated at 22°C, 32°C and 36°C, respectively. Plates were scanned after 3 days at 32°C and 36°C and after 7 days at 22°C.

II.7: Computational analysis

To create a database of *S. pombe* introns we first compiled the sequences corresponding to the 4862 introns based on a 2007 release of the *S. pombe* genome. The

sequence corresponding to the TER1 intron was added manually. We then applied a filter to remove duplicate sequences and annotated introns that did not begin with '$G^{T}/_{C}$' and end with 'AG', which reduced the number to 4824 putative intronic sequences.

A Gibb's sampler (http://bayesweb.wadsworth.org/gibbs/gibbs.html) was then used to construct position weight matrixes (PWM). The first 7 and last 3 nucleotides were excluded as they represent the 5' and 3' splice sites, respectively. The recursive sampling algorithm was used with the default settings for eukaryotes. The top 5 motifs of lengths 7, 9, and 10 were evaluated. All PWMs generated by this analysis resembled branch sites. The Gibb's sampler was rerun to construct the best PWM found at least once in all of the sequences. The final branch site PWM (represented as a sequence logo in Fig. 5.1A) was 10 nucleotides long.

BS scores were calculated as follows. At each position a BS score S_{BS} was calculated by applying equation 1 to a 10-nucleotide window starting at that position. In this equation f_i is the frequency for the nucleotide found in the PWM at position i, and I_i is the Shannon-information (in bits) for position i as calculated from the PWM for that position (the Shannon-information was calculated as described [122]. The score S_{BS} is therefore the total number of bits for the sequence contained in the window as defined by the branch site PWM. Using this scoring method the top scoring sequence is 'TTACTAACTT' which has a $S_{BS} = 7.856$. .

Equation 1: $$S_{BS} = \sum_{i=1}^{10} f_i I_i$$

The average S_{BS} for all positions was 3.1, but the average score for the highest scoring position in each intron was 7.1 (Fig. 5.2D). Approximately 96% of all introns had at least one site with a score >6 (Fig. 5.2C), this was therefore used as the cutoff for predicting putative branch sites in each intron. Nearly half (42%) of the introns had only a single site with a score >6 (Fig. 5.2D). In the cases where there was more than 1 site with a score >6 the site that was closest to the acceptor site, but was greater than 5 nucleotides from the 3' ss, was chosen as the putative BS.

Chapter Three:

Intronic sequence elements impede exon ligation and trigger a discard pathway that yields functional telomerase RNA in fission yeast

III.1: Abstract

The fission yeast telomerase RNA (TER1) precursor harbors an intron immediately downstream of its mature 3' end. Unlike most introns, which are removed from precursor RNAs by the spliceosome in two sequential but tightly coupled transesterification reactions, TER1 only undergoes the first cleavage reaction during telomerase RNA maturation. The mechanism underlying spliceosome-mediated 3' end processing has remained unclear. We now demonstrate that a strong branch site (BS), a long distance to the 3' splice site (3' ss) and a weak polypyrimidine (Py) tract act synergistically to attenuate the transition from the first to second step of splicing. The observation that a strong BS antagonizes the second step of splicing in the context of TER1 suggests that the BS – U2 snRNA interaction is disrupted after the first step and thus much earlier than previously thought. The slow transition from first to second step triggers the prp22 DExD box helicase dependent rejection of the cleaved products and prp43 dependent 'discard' of the splicing intermediates. Our findings explain how the spliceosome can function in 3' end processing and provide new insights into the mechanism of splicing.

III.2: Introduction

The enzyme telomerase replenishes repetitive DNA sequences at the ends of eukaryotic chromosomes that otherwise shorten with each round of replication [123]. At its core, telomerase is comprised of a catalytic protein subunit and a non-coding RNA component. The RNA serves as a scaffold for the assembly of the holo-enzyme and provides the template for telomeric repeat synthesis by the catalytic subunit telomerase reverse transcriptase (TERT).

The gene encoding the telomerase RNA subunit from the fission yeast *Schizosaccharomyces pombe* has been identified [115, 124] and examination of the maturation pathway revealed that the RNA is first transcribed as a longer poly-adenylated precursor containing an intron immediately downstream of the 3' end of the mature form [116]. Interestingly, the mature 3' end of TER1 is generated by a single cleavage reaction at the 5' splice site akin to the first step of splicing (Fig 5.1A).

Prior to cleavage, a ring of seven Sm proteins binds the TER1 precursor directly upstream of the intron and recruits the methylase Tgs1, which converts the monomethyl guanosine cap at the 5' end of TER1 into the trimethyl guanosine form [118]. After cleavage by the spliceosome has occurred, Sm proteins dissociate from the RNA and are replaced by a related protein complex comprised of seven Like-Sm proteins, Lsm2 to 8 [125]. The Lsm complex protects the 3' end of telomerase RNA against degradation and promotes assembly of the RNA with the catalytic subunit of telomerase [118]. The identification of several key events that convert the telomerase RNA precursor into the mature form has laid the foundation for mechanistic studies. A question of particular

interest is how the spliceosome carries out a single cleavage reaction in 3' end processing when its normal role is to remove introns in two tightly coupled steps?

In the first step of intron excision, the 2′ hydroxyl of the branch point (BP) adenosine, located within the branch site (BS) sequence, attacks the sugar-phosphate bond at the 5′ splice site (5' ss), forming a 2'-5' linkage and producing a free 5′ exon and a branched species, termed the lariat intermediate [114]. In the second reaction, the now exposed 3′ hydroxyl of the upstream exon attacks the phosphate at the 3′ splice site (3' ss), yielding ligated mRNA and the lariat form of the intron. Although the TER1 intron contains all RNA elements required for complete splicing, the level of spliced TER1 accounts for less than 1% of TER1 isoforms in actively growing cells [116].

Consistent with telomerase RNA 3' end processing involving only the first step of splicing, nucleotide changes at the 5' ss and the BP adenosine blocked TER1 maturation [116]. In contrast, mutations at the 3' ss did not impair the accumulation of the cleaved form, but eliminated the small amount of spliced TER1 that is normally observed [116]. Direct evidence for spliceosome-mediated 3' end processing of TER1 came from experiments in which 5' ss mutations were rescued by compensatory changes in the U1 snRNA which base pairs with the 5' ss during spliceosome assembly. Furthermore, replacement of the TER1 intron with a heterologous intron from a protein-encoding gene resulted in the production of spliced TER1 RNA, which failed to support telomerase activity. When the 3' ss was mutated in this context, some cleaved TER1 was produced

and telomere maintenance was partially restored [116]. Blocking the completion of splicing is thus of pivotal importance for producing functional telomerase enzyme.

The mechanism of splicing has been studied primarily in the context of protein encoding genes. There, release of the 5' exon after the first cleavage reaction would block gene expression as released splicing intermediates will be degraded rather than translated. It is thus no surprise that the two steps of splicing are tightly coupled, and at least for some introns the first step of splicing is dependent on the presence of an intact 3' ss [91, 92, 126]. It is presently unclear whether fission yeast telomerase RNA is unique in utilizing the first step of splicing for 3' end processing or whether it is the first example of a class of non-coding RNAs that are processed in this manner. The observation that introns are found near the 3' ends of telomerase RNAs from several other fungi indicates that spliceosomal 3' end processing is at least not limited to *S. pombe* [127, 128], P.B., unpublished data).

Here we show that the ability of the spliceosome to function in 3' end processing is intrinsic to the kinetic quality control mechanism that permits discard of suboptimal intermediates [112, 113, 126, 129]. The uncoupling of the first and second step of splicing that occurs naturally in the context of TER1 processing provides new insights into the roles of the branch site in splicing.

III.3: Results

III.3.a: RNA elements within the intron inhibit the second step of splicing.

Many RNA processing events are initiated by the recruitment of processing factors to the carboxy-terminal domain (CTD) of RNA pol II [130-132]. These factors travel with the polymerase during transcription and act on the nascent RNA transcript upon recognition of a processing signal [133]. To test if the *ter1* promoter and first exon are required for 3' end processing by the spliceosome, we replaced them with the promoter and open reading frame of the *ura4* gene. As the binding of Sm proteins directly upstream of the 5' splice site may be responsible for inhibiting the completion of splicing, 14 nucleotides of TER1 sequence including the Sm site were included in the chimeric construct. A radiolabelled oligonucleotide probe complementary to TER1 exon 2 was used to detect the precursor and spliced forms (Fig. 3.1B, upper panel). Reprobing the northern with a probe complementary to the first exon (Ura4 ORF) revealed an additional faster migrating form corresponding in size to the first exon, the expected product of the first step of splicing (middle panel). This form was insensitive to a mutation of the 3' ss (lane 4), but was abolished when the 5' ss was mutated (lane 5). Its identity as the product of spliceosomal cleavage was further confirmed by cloning and sequencing its 3' terminus (Fig. 3.2) Consistent with previous findings in the context of endogenous TER1 [116], reducing the distance between the BP and 3' ss to nine nucleotides promoted complete splicing (Fig. 3.1, lane 6).

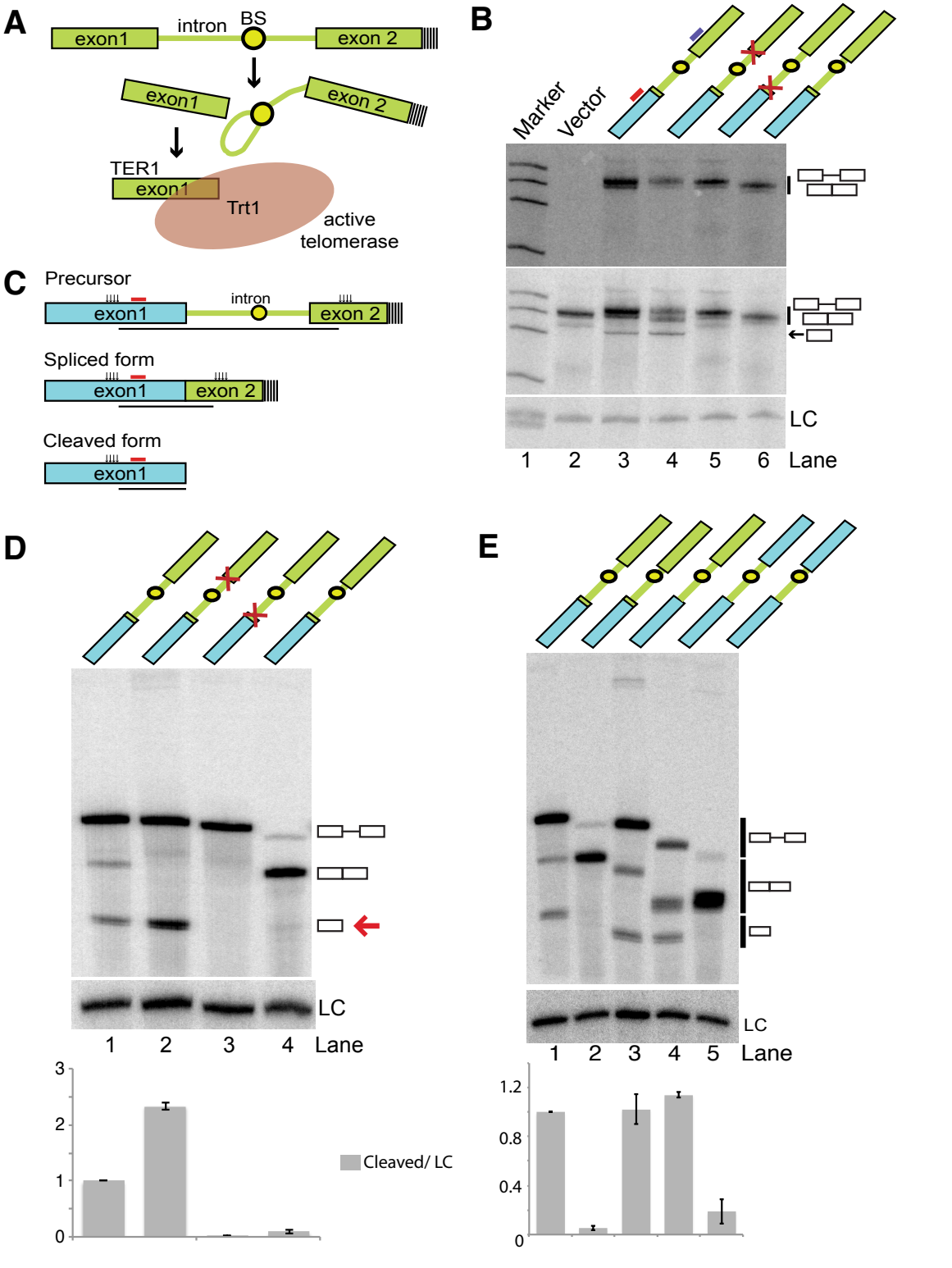
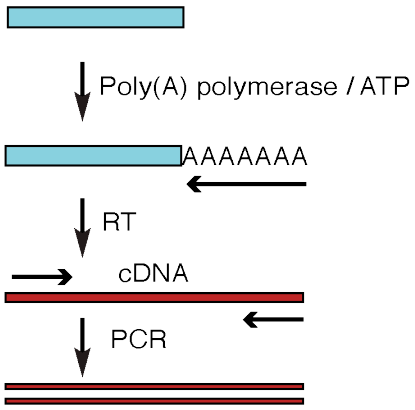

A
exon1
intron
BS
exon 2
exon1
exon 2
TER1
exon1
Trt1
active
telomerase
B
Marker
Vector
LC
1 2 3 4 5 6 Lane
C
Precursor
exon1
intron
exon 2
Spliced form
exon1
exon 2
Cleaved form
exon1
D
LC
1 2 3 4 Lane
3
2
1
0
Cleaved/ LC
E
LC
1 2 3 4 5 Lane
1.2
0.8
0.4
0

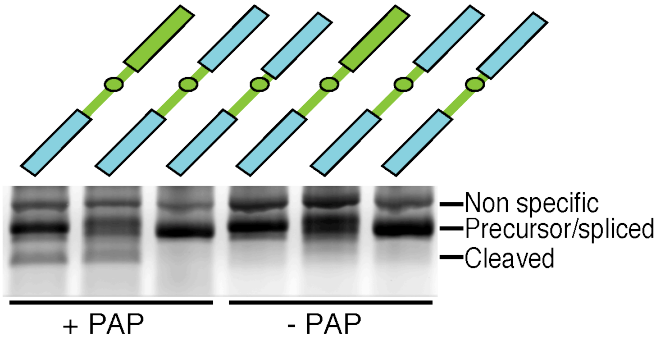
Figure 3.1: The TER1 intron contains all elements required for spliceosomal cleavage. **(A)** Schematic of TER1 3' end processing by the first step of splicing. The branch site (BS) is shown as a circle and poly(A) tails are indicated by dashed lines. **(B)** Northern blot on RNA isolated from cells expressing reporter constructs as indicated. Ura4 sequence is shown in blue, TER1 sequence in green; mutations at the 5' ss and 3' ss are symbolized by red crosses. A ^{32}P-labeled oligonucleotide complementary to the second exon (position indicated by purple line above lane 3) was used for detection of precursor and spliced forms (top panel). The northern was reprobed with a radiolabeled oligo complementary to the first exon to also visualize the cleaved form (middle panel, position of probe is indicated by red line above lane 3). A probe against snoRNA snR101 was used as loading control (LC). **(C)** Schematic of RNaseH cleavage assay. Cleavage sites are indicated by clusters of vertical arrows. Those cleavage products that are homogenous in size and visualized on northern blots are underlined. The probe used for detection of all three forms is indicated by a red line. **(D)** RNaseH cleavage followed by northern blot analysis of the indicated constructs. TER1 sequences in green, URA4 sequences in blue, a red arrow points to the cleaved form. Products were detected using a probe complementary to the first exon as indicated in (C). The ratios of cleaved form to LC for each lane were normalized to wild type (lane1) with the data being represented as mean +/- SEM from three independent experiments. **(E)** As in (D) for additional constructs including one that places the TER1 intron in an entirely heterologous context (lane 4). Differences in mobility for each form between lanes reflect differences in the length and sequence of constructs used. Quantification as described in (E) with n=2.

The similarity in sizes and length heterogeneity arising from multiple poly-adenylation sites and variable polyA tail length complicated quantification of the different RNA forms on northern blots. This issue was resolved by treating the RNA with RNaseH in the presence of two DNA oligonucleotides complementary to sites in exons one and two, respectively (Fig. 3.1C). Following RNaseH cleavage, precursor, spliced and cleaved forms resolved into three well-separated bands (Fig. 3.1D, lane 1). A 2.3 fold increase in the cleaved product was observed when the second step of splicing was blocked by a 3' ss mutation (lane 2). A 5' ss mutation reduced the first step by more than 10-fold (lane 3) and shortening the BP to 3' ss distance promoted complete splicing (lane 4).

A


Poly(A) polymerase / ATP
AAAAAAA
RT
cDNA
PCR

B


Non specific
Precursor/spliced
Cleaved
+ PAP
- PAP

C

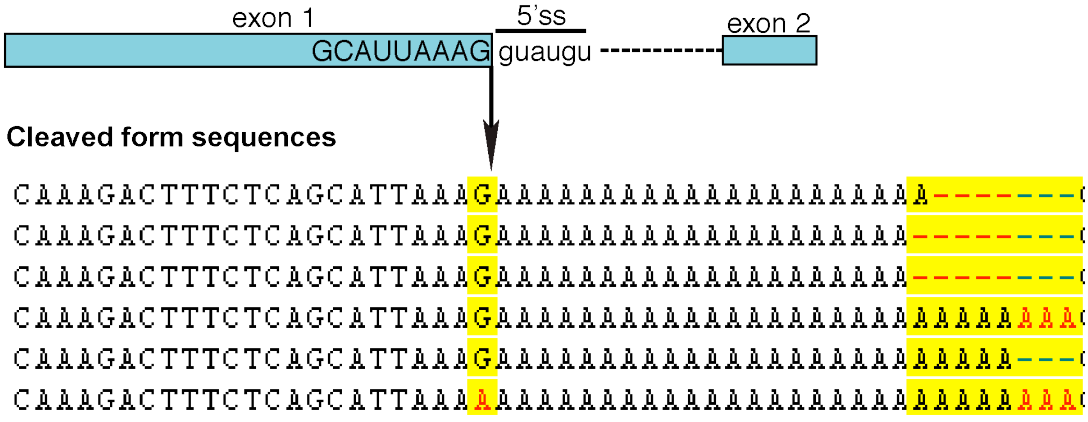

exon 1
5'ss
exon 2
GCAUUAAAG guaugu
Cleaved form sequences
CAAAGACTTTCTCAGCATTAAAGAAAAAAAAAAAAAAAAAAAAA--------
CAAAGACTTTCTCAGCATTAAAGAAAAAAAAAAAAAAAAAAAA---------
CAAAGACTTTCTCAGCATTAAAGAAAAAAAAAAAAAAAAAAAA---------
CAAAGACTTTCTCAGCATTAAAGAAAAAAAAAAAAAAAAAAAAAAAAAAA
CAAAGACTTTCTCAGCATTAAAGAAAAAAAAAAAAAAAAAAAAAAAAA---
CAAAGACTTTCTCAGCATTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

Figure 3.2: Cloning and sequencing of 3' end processing products. **(A)** Schematic of protocol used to capture the 3' ends of the spliceosomal cleavage products. RNA (shown in blue) was tailed by treatment with poly (A) polymerase in the presence of ATP and reverse transcribed using an oligo-dT primer. The first-strand cDNA (shown in red) was amplified by polymerase chain reaction (PCR) using oligo-dT and the gene specific primer (PBoli517; GCTTTTGGCTGAAATGTCTTCC). **(B)** The PCR products of 3' end cloning reaction were separated by agarose gel electrophoresis and visualized using ethidium bromide staining. The appearance of a band corresponding to the first exon released after spliceosomal cleavage is dependent on the reporter undergoing spliceosomal cleavage rather that complete splicing and on PAP treatment. The second exon of TER1 is shown in green, URA4 sequences in blue, the TER1 intron as a black line and the branch site as a red circle. The construct used in lane 3 has a shortened distance between the BP and 3'ss and is thus efficiently spliced. Poly (A) polymerase was omitted from the samples in lanes 3 to 6. As the cleaved form is not polyadenylated *in vivo*, the lowest band is absent in these lanes. **(C)** Six representative sequences corresponding to the fast migrating band in lane 1 in (B). The URA4 reporter sequence ends precisely upstream of the 5' splice site (indicated by arrow) confirming the identity of this band as the product of spliceosomal cleavage. The run of adenosines reflects nucleotides added by PAP.

The observed cleavage of a reporter containing a heterologous promoter and first exon indicated that the TER1 Sm site, intron and second exon contain all elements required to inhibit the second step of splicing. To further define the essential elements, we deleted the Sm site (Fig. 3.1E, lane 3) and replaced the TER1 second exon with the 3'UTR of *ura4* (lane 4), neither of which reduced the levels of cleaved form. The spliced form was more abundant in the context of the ura4 3'UTR than in the presence of the TER1 second exon. This may reflect differences in stability between the two products or a context dependency of splicing efficiency. In either case, these results show that the TER1 intron is sufficient to promote spliceosomal cleavage of a heterologous RNA and thus harbors the elements required to uncouple the two transesterification reactions.

III.3.b: The strength and spacing of the 3' intronic elements control uncoupling.

Since deletion of 14 nucleotides between the BP and 3' ss promoted complete splicing, we asked whether inhibition of the second step of splicing was due to the long distance or a specific sequence motif. Starting with the *ter1-55* mutant (Fig. 3.3A) we increased the distance between the BP and 3' ss in a step-wise manner by insertion of adenines. Adenines were chosen as increasing pyrimidine content between the BP and 3' ss has been shown to promote splicing in fission yeast [91] and metazoans [92]. As the BP to 3' ss distance was increased from nine to 12 and 16 nucleotides, we observed a stepwise increase in the cleaved form (Fig. 3.3B). However, the spliced product was also readily detected by northern blot for all three constructs and dominated over the precursor in RT-PCR analysis (Fig. 3.3C). Further increasing the distance from 16 to 20 nucleotides resulted in an 8-fold increase in the cleaved product, indicating that the second step of splicing was specifically inhibited by the increased distance between BP and 3' ss. The first step appeared unaffected by the replacement of endogenous TER1 sequence with adenosines as the 9+11A and 9+14A constructs yielded as much or more cleaved product as wild type TER1 (Fig. 3.3B, compare lanes 1 with 4 and 5).

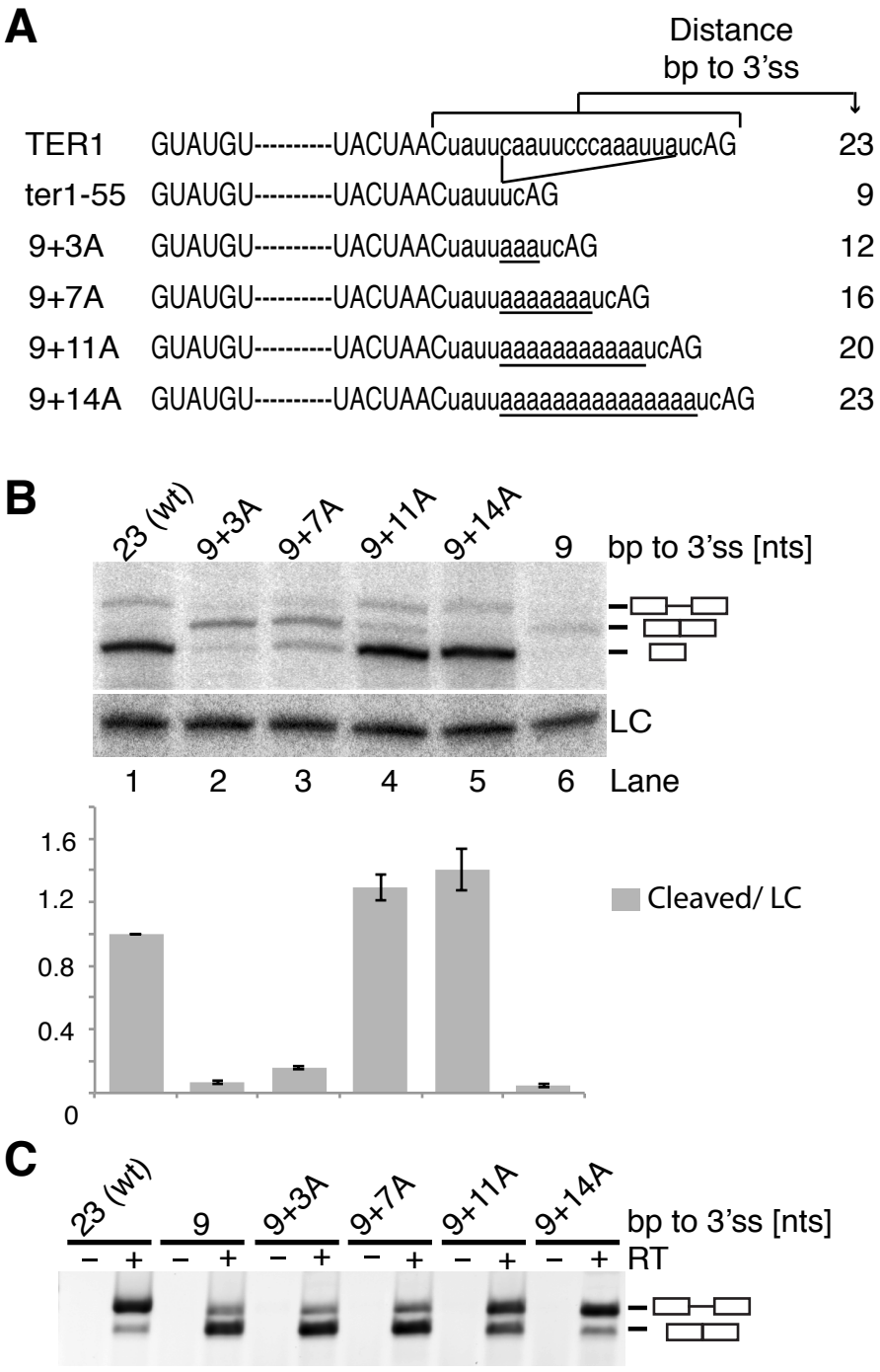


Figure 3.3: A long distance between the BP and 3' ss inhibits the 2nd step of splicing. **(A)** Intronic sequences of TER1 constructs used in (B) and (C). Splice and branch sites are in capital letters for emphasis. Inserted adenosines are underlined. **(B)** Northern blot analysis for *ter1* constructs expressed from plasmids under the control of the *ter1* promoter. RNaseH cleavage was used to improve resolution of precursor, spliced and

cleaved forms. The cleaved form was quantified relative to loading control snRNA101 (LC) and normalized to wild type levels. The data were represented as mean +/- SEM with n= 2 for '9+3A' construct and n=3 for all other constructs. An arrow points to the cleaved form. **(C)** RT-PCR across the TER1 intron to visualize the relative abundance of precursor and spliced forms.

The observation that the first step but not the second step of splicing proceeded efficiently in the 9+14A mutant was surprising as the replacement of endogenous TER1 sequence with 14 adenosines eliminated the Py-tract. Biochemical analysis in fission yeast cell-free extracts has identified the Py-tract as an important element in spliceosome assembly at a step preceding U2 recruitment [134]. In humans, Py-tract strength predominantly affects spliceosome assembly and efficiency of the first step [92, 135], but effects of distance and sequence on the second step have been observed as well [92, 136, 137]. To further investigate the effect of pyrimidine content in the context of telomerase RNA 3' end processing, additional constructs with weak and strong Py-tracts were synthesized and integrated at the TER1 locus (Fig. 3.4A). TER1 itself has a moderately

strong Py-tract with an uninterrupted run of five pyrimidines and an overall pyrimidine content of 65% between the BS and 3' ss. Increasing the pyrimidine content favored completion of the second step as indicated by a 2-fold reduction in the cleaved form and now readily detectable levels of fully spliced product for the Py+ construct (Fig. 3.4B). In contrast, loss of the Py-tract in the context of the TER1 intron had little effect on the first or second step of splicing (lane 3).

III.3.c: A role for U2AF in promoting the second step of splicing.

During spliceosome assembly the Py-tract is bound by the large subunit of U2AF (U2AF65 in mammals, U2AF59 in fission yeast; [138-140]. Py-tract strength correlates with U2AF65 binding affinity and selection experiments have shown that U2AF65 prefers binding sites containing uridines over cytosines [141]. Replacement of the five cytosines in the TER1 Py-tract with uridines (C2U mutant) resulted in an increase in spliced form at the expense of first step product (Fig. 3.4B, lane 4). This result further supports that a strong Py-tract favors completion of splicing and indicates that the effect may be mediated by U2AF. We thus examined processing of the TER reporter construct in the context of a temperature sensitive allele of U2AF59 [142]. When RNA was isolated from cells immediately after shifting to the restrictive temperature, a 2.5-fold lower level of cleaved TER1 was observed in U2AF59^{ts} compared to wild type cells (Fig. 3.4C). This is consistent with a modest assembly and first step defect even at the permissive temperature. Interestingly, no further reduction in the first-step product was observed in the mutant after 30 minutes at the restrictive temperature. However, the second step of splicing was now impaired as evidenced by a 2-fold reduction in the level of spliced form in the mutant. Taken together, these observations reveal that in the context of the TER1 intron the first step of splicing occurs largely independent of the presence of a Py-tract. Notably though, completion of the second step of TER1 splicing was favored by a strong Py-tract and functional U2AF59.

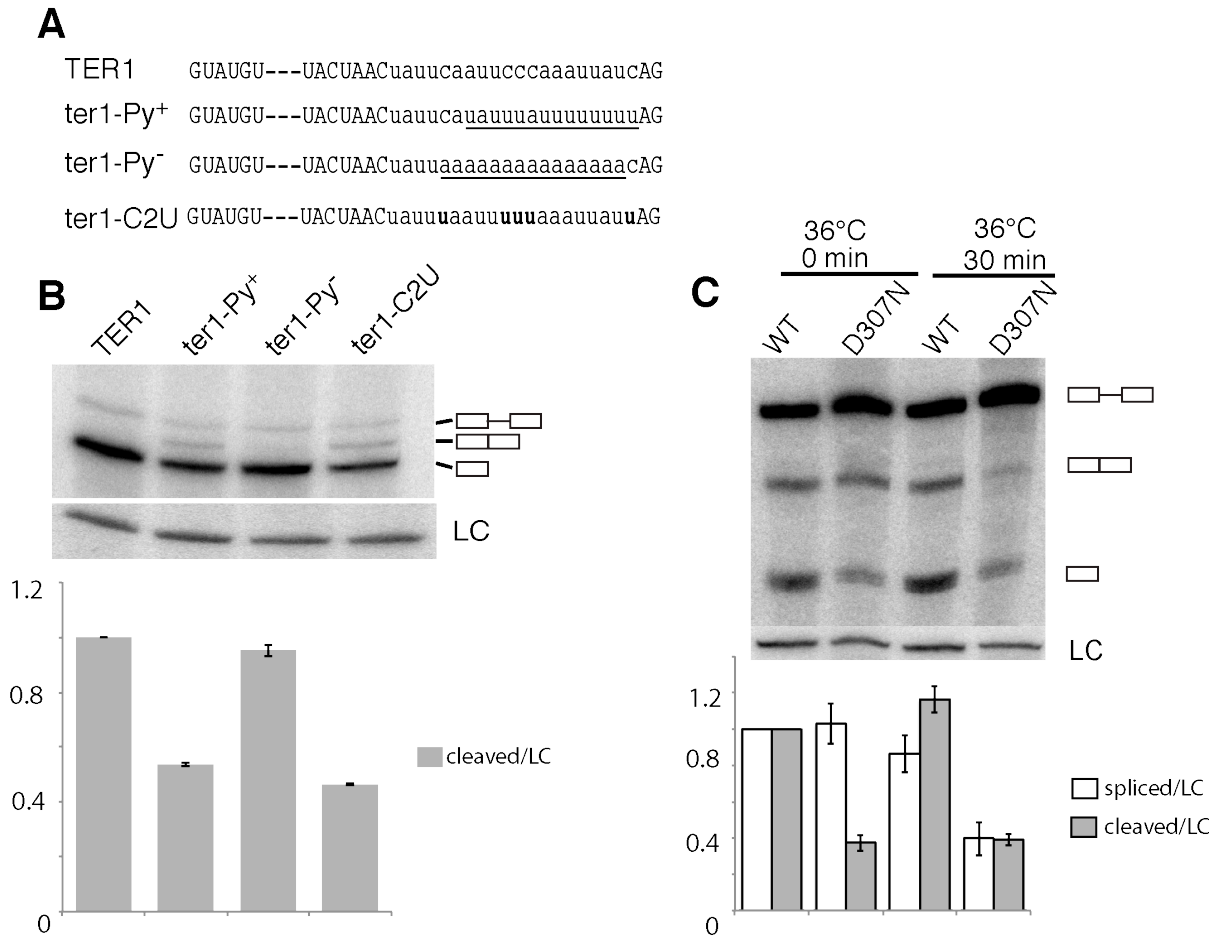


Figure 3.4: A strong Py-tract and U2AF59 favor completion of the 2nd step of splicing. **(A)** TER1 mutants with strengthened (Py+) or weakened (Py-) pyrimidine tract or cytosines replaced with uridines (C2U) were integrated at the *ter1* locus. **(B)** Detection of precursor, spliced and cleaved forms by northern blot. LC and quantification as in 2B with n=2. **(C)** Northern analysis of RNA from cells containing the *ura4* reporter with TER1 intron in the context of $prp2^{+}$ ($U2AF^{59}$) and $prp2^{D307N}$ mutant, respectively. Cells were grown at 25°C to 5 x 10^6 cells/ml, shifted to 36°C and harvested after 0 or 30 min. Quantification of cleaved and spliced forms relative to LC (snR101) for each lane and normalized to wild type (lane 1). Quantification were performed as in Fig. 2B with n=2.

III.3.d: Unexpected role for the branch site in inhibiting completion of splicing.

Although the BP to 3' ss distance is about twice as long in the TER1 intron than the median distance for all annotated introns in the *S. pombe* genome (see below), well over 100 introns in protein encoding genes have a similar or greater BP to 3' ss distance. We thus wondered whether these introns are efficiently spliced. The first intron in the *fcp1* gene encoding for a serine phosphatase [143] resembles the TER1 intron in overall length and architecture with a BP to 3' ss distance of 24 nucleotides (Fig. 3.6A, B). However, when this intron was inserted into the reporter robust completion of splicing was observed (Fig. 3.5A, lane 2). The intron was also efficiently spliced in the context of the endogenous *fcp1* gene (Fig. 3.6C). A longer distance between BP and 3' ss is therefore not sufficient to account for the uncoupling of the first and second step of splicing in TER1.

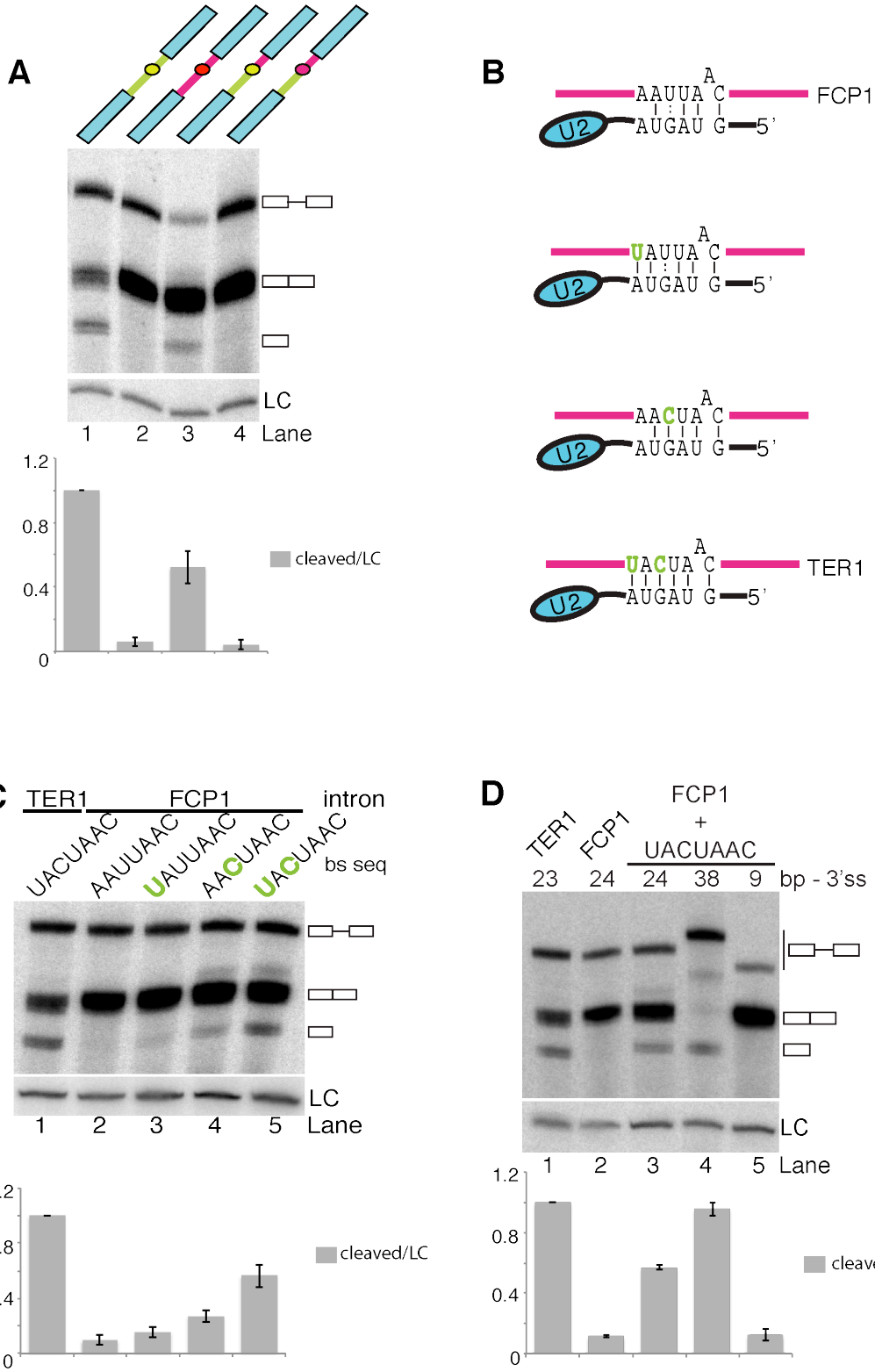

A
LC
1 2 3 4 Lane
1.2
0.8
0.4
0
cleaved/LC
B
A
AAUUA C
AUGAU G 5'
U2
FCP1
UAUUA C
AACUA C
UACUA C
TER1
C
TER1
FCP1
intron
UACUAAC
AAUUAAC
UAUUAAC
AACUAAC
UACUAAC
bs seq
LC
1 2 3 4 5 Lane
cleaved/LC
D
FCP1
+
UACUAAC
TER1
FCP1
23 24 24 38 9 bp - 3'ss
LC
1 2 3 4 5 Lane
cleaved/LC

Figure 3.5: A strong BS inhibits the second step of splicing. **(A)** Northern blot analysis of reporter constructs that contain the TER1 intron (green) or FCP1 intron (pink) between *ura4* ORF and *3'* UTR (blue). Chimeric reporters contain TER1 or FCP1 BS indicated by green or pink filled circles, respectively. A probe against snR101 was used as loading control (LC) for normalization of cleaved form relative to lane 1 with the data being represented as mean +/- SEM (n=2). **(B)** Schematic of base pair interactions between U2 snRNA and branch sites of FCP1 and variations including the TER1 BS sequence. A vertical line is used to represent Watson Crick base pairing, two dots represent G:U interactions. The bulged adenosine at the branch point is drawn above the paired sequence. Nucleotide changes from the wild type FCP1 sequence are highlighted in green. **(C)** Northern blot analysis for reporter constructs with FCP1 intron containing the indicated branch site sequences. Quantification as in (A) with n=2 for UAUUAAC, AACUAAC mutants and n=3 for other mutants. **(D)** Distance between BP and 3' ss was varied in the context of the FCP1 intron with a strong BS (for schematic see SF 2E). The cleaved forms were quantified relative to loading control (LC) and normalized to lane 1 (n=2).

To identify sequence elements responsible for spliceosomal cleavage several chimeric introns were generated and tested in the context of the reporter assay. Chimeras containing TER1 sequences upstream of the branch site and FCP1-derived sequences between the BS and 3' ss were spliced well, whereas reciprocal chimeras were processed poorly (Fig. 3.5A and Fig. 3.6D). Most notable, in otherwise identical constructs, spliceosomal cleavage was observed when the BS sequence was derived from TER1, but not with the BS sequence from FCP1 (Fig. 3.5A, lanes 3 and 4 and Fig. 3.6D). This observation was surprising, as cross-linking between U2 and the BS in the pre-mRNA and in the excised intron lariat had led to the view that the U2 – BS association is maintained through both steps of splicing [144]. In contrast, our observation suggests that a tight U2 – BS interaction promotes accumulation of the 5' cleavage product, although in the context of the reporter the spliced form still predominates.

The TER1 branch site is fully complementary to the BS binding sequence in U2 snRNA (Fig. 3.5B), whereas the BS in FCP1 intron 1 deviates by two nucleotides from the U2 complement, resulting in a mismatch and the replacement of a G-C base pair with a G:U wobble (Fig. 3.5B). These two nucleotide differences alone reduced the level of cleavage product by 10-fold (Fig. 3.5C). Restoring one or the other of the Watson-Crick base pairs resulted in spliceosomal cleavage at 16 and 27% of the level observed with the TER1 intron, respectively (lanes 3 and 4). When both nucleotides were changed together, spliceosomal cleavage products accumulated to 56% of the TER1 control (lane 5). Further increasing the distance between BS and 3' ss to 38 nucleotides stimulated spliceosomal cleavage at the expense of splicing, whereas reducing the distance promoted the 2nd step of splicing (Fig. 3.5D).

These observations demonstrate that, at least in the context examined here, the U2 – BS interaction has opposing effects on the first and second step of splicing. High U2 complementarity promotes spliceosome assembly and the first cleavage reaction, but is inhibitory to the completion of splicing. The most parsimonious explanation for these findings is that U2 must dissociate at least partially from the BS for splicing to proceed beyond the first step.

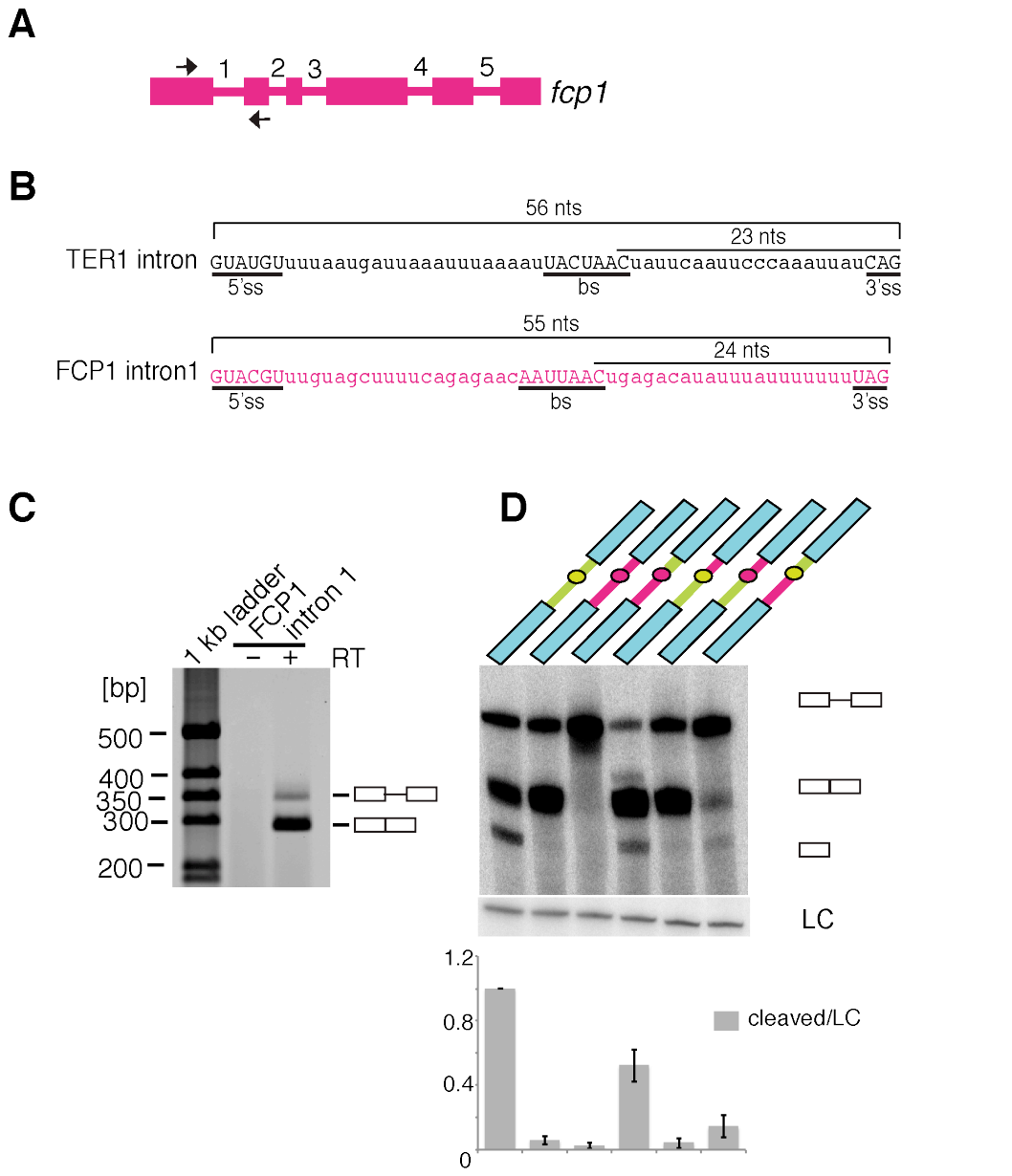

A
1
2
3
4
5
fcp1
B
56 nts
23 nts
TER1 intron
GUAUGUuuuaaugauuaaauuuaaaauUACUAACuauucaauucccaaauuauCAG
5'ss
bs
3'ss
55 nts
24 nts
FCP1 intron1
GUACGUuuguagcuuuucagagaacAAUUAACugagacauauuuauuuuuuuUAG
5'ss
bs
3'ss
C
1 kb ladder
FCP1 intron 1
–
+
RT
[bp]
500
400
350
300
200
D
LC
1.2
0.8
0.4
0
cleaved/LC
E
distance between bp and 3'ss
FCP1 intron
GUACGU---------AAUUAACugagacauauuuauuuuuuuUAG
5'ss
bs
3'ss
+ UACUAAC BS
GUACGU---------UACUAACugagacauauuuauuuuuuuUAG
38 nts bp - 3'ss
GUACGU---------UACUAACugagacauauuuauuuuuuu(a)14UAG
9 nts bp - 3'ss
GUACGU---------UACUAACuuuuuUAG

Figure 3.6: Chimeric introns identify the branch site as a key determinant for the inhibition of the second step of splicing. **(A)** Schematic of *fcp1* gene structure comprised of 6 exons (rectangles) and 5 introns (lines). Introns are numbered and arrows mark the positions of primers used to examine the efficiency of splicing for intron 1. **(B)** Comparison of intronic sequences and distances for TER1 (black) and FCP1 intron 1 (pink). The 5' and 3' splice sites and the branch site (BS) are underlined and capitalized. **(C)** RT-PCR across FCP1 intron1 using primers BLoli2149 (ATTGAAATCGCGAGCTGTCT) and BLoli2150 (TTCCACAAATTGCACAGAG). RT=reverse transcriptase was omitted from the control reaction. **(D)** Northern blot analysis for reporter constructs containing the *ura4* ORF and 3'UTR (blue), and chimeric introns (TER1 sequence in green, FCP1 sequence in pink). Branch sites are depicted by circles filed in green or pink depending on the origin of the BS sequence. **(E)** Partial sequence of constructs used in Figure 4D. Mutations in the BS sequence that convert the FCP1 BS into the TER1 BS are shown in green. An insertion of 12 adenosines increases the BP to 3'ss distance to 38 nucleotides, whereas a deletion of 15 nucleotides reduces the distance to 9 nucleotides.

III.3.e: The 'discard' pathway promotes spliceosomal cleavage.

The transition from the first to the second step of splicing involves major conformational changes within the spliceosome. Biochemical experiments using budding yeast cell-free extracts have revealed a quality control mechanism that monitors the success of exon ligation [112, 113]. When the DExD box helicase Prp22 hydrolyses ATP prior to the exon ligation step, the splicing intermediates are rejected. This process has been suggested to permit recycling of spliceosomes bound to mutant RNA substrates that arise due to RNA polymerase errors or mutations in the DNA template [102]. It also allows recovery of spliceosomes that have selected suboptimal 3' splice sites.

Here we have identified several elements in the TER1 intron that encumber the first to second step transition and may thus trigger Prp22-dependent 'discard' of the first-step products. Based on sequence similarity with *S. cerevisiae PRP22* and other DExD-box helicases [145], we designed point mutations in *S. pombe prp22*. T665A and H634A displayed weak temperature sensitivity, whereas S663A was strongly cold sensitive (Fig.

3.8A). As *prp22* promotes the release of first and second step products in an ATPase-dependent manner and also stabilize the exon ligation conformation in an ATPase-independent manner, it was important to distinguish between these effects. Mutations that affect the ATPase activity and thereby impair 'rejection' are expected to permit increased completion of splicing, whereas mutations that compromise the prp22 function in exon ligation will reduce the amount of spliced product. We therefore calculated what fraction of RNA had undergone the first and second step respectively and normalized to the wild type sample at 32 °C, for which both values were set to 100. Although such quantification does not directly measure the actual percentage of RNA molecules that are released after the first step, it is nevertheless an established measure for first and second step efficiencies [99]. All three mutations in the Prp22 helicase domain shifted the ratio of first to second step product in favor of spliced RNA supporting a role for the 'discard pathway' in TER1 biogenesis (Fig. 3.7A, B and Fig. 3.8B).

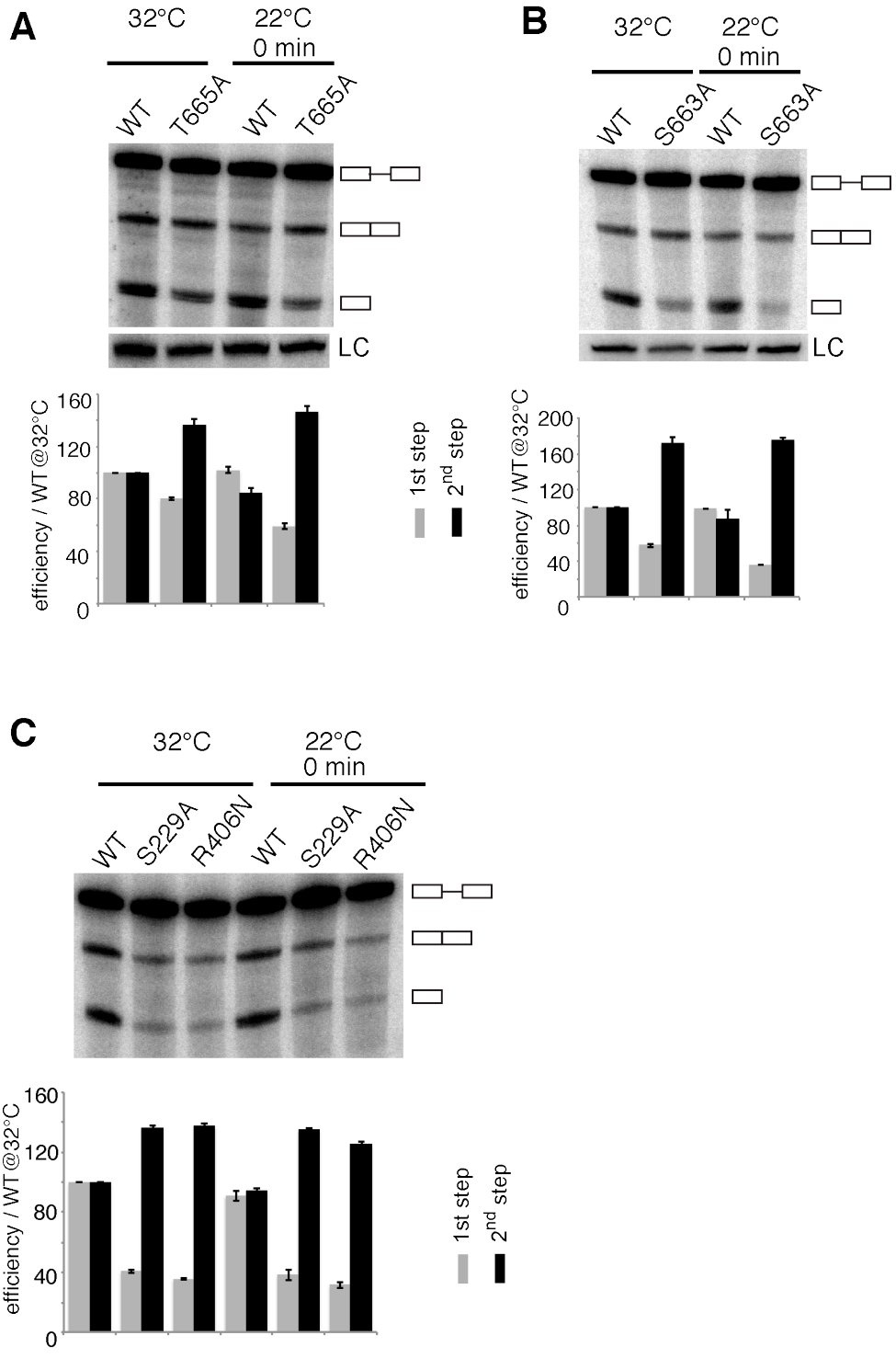

A
32°C
22°C
0 min
WT
T665A
WT
T665A
LC
efficiency / WT@32°C
160
120
80
40
0
1st step
2nd step
B
32°C
22°C
0 min
WT
S663A
WT
S663A
LC
efficiency / WT@32°C
200
160
120
80
40
0
C
32°C
22°C
0 min
WT
S229A
R406N
WT
S229A
R406N
efficiency / WT@32°C
160
120
80
40
0
1st step
2nd step

Figure 3.7: Prp22 and Prp43 promote spliceosomal cleavage. **(A)** Northern Blot analysis of reporter containing wild type TER1 intron. RNA was isolated from *prp22*$^{+}$ and *prp22*^{T665A} cells initially grown at 32°C to 5 x 10^6 cells/ml, shifted to 22°C and harvested immediately. 1st step efficiencies were calculated as cleaved plus spliced form over all forms normalized to wild type at 32 °C multiplied by 100. 2nd step efficiencies were calculated as spliced form over cleaved plus spliced form normalized to wild type at 32 °C times 100. Data represented as mean +/- SEM (n=3). **(B)** Northern blot analysis of a second Prp22 mutant *prp22*^{S663A}. Analysis of TER1 processing as in (A) with N=3. **(C)** Analysis for *prp43*^{S229A} and *prp43*^{R406N} by northern blotting. Quantification as described in (A) with n=2.

The Prp43 helicase functions downstream of Prp22 in the release of spliced mRNA as well as the discard of suboptimal splicing intermediates [112]. A mutation in budding yeast *PRP43* further revealed a role in enhancing the specificity of exon ligation by repressing the use of a cryptic 3' splice site. Consistent with a role for *S. pombe prp43* in TER1 processing, two mutations (Fig. 3.7C) also reduced first step and increased second step efficiencies for the TER1 reporter (Fig. 3.7C).

In summary, several RNA elements in the TER1 intron act cooperatively to attenuate the second step of splicing, thus creating an opportunity for the rejection and discard of the splicing intermediates in a Prp22 and Prp43-dependent manner.

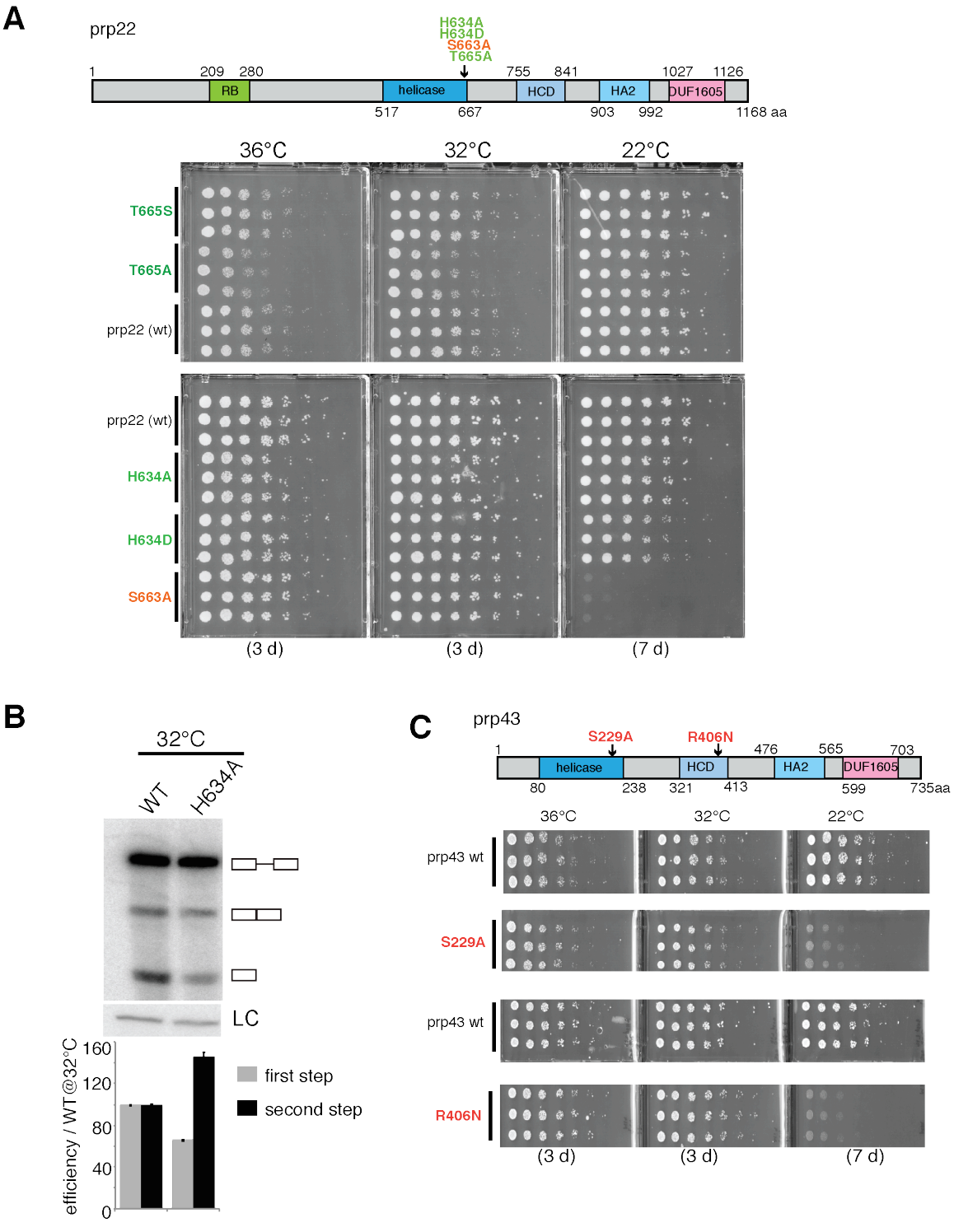


Figure 3.8: Characterization of *prp22* and *prp43* mutant alleles. **(A)** Schematic of Prp22 domain structure. Amino acid change and position of mutations used in this study are

indicated above the bar. Start and end positions of known domains are shown as colored boxes and by amino acid position. Abbreviations: RNA binding domain (RB), helicase conserved C-terminal domain (HCD), helicase associated domain (HA2) and domain of unknown function 1605 (DUF1605). Spotting assay for strains harboring *prp22*$^{+}$ or mutant alleles. Five fold dilutions of cultures were spotted in triplicates and plates were incubated at 22°C, 32°C and 36°C, respectively. Pictures were taken after the indicated number of days. **(B)** Northern blot analysis of the reporter containing the wild type TER1 intron in the context of *prp22*$^{+}$ and *prp22*^{H634A} cells initially grown at 32°C to 5 x 10^{6} cells/ml and then shifted to 22°C. Cells were harvested and RNA was isolated as soon as the culture reached 22°C. **(C)** 1st step efficiencies were calculated as cleaved plus spliced form over all forms normalized to wild type at 32°C times 100. 2nd step efficiencies were calculated as spliced form over cleaved plus spliced form normalized to wild type at 32°C times 100.

III.4: Discussion

The realization that fission yeast telomerase RNA 3' end processing is carried out by the first step of splicing was surprising at first. The need to accurately remove thousands of introns from pre-mRNAs has resulted in tight coupling of the two steps of splicing, as to not inadvertently release the intermediates. Utilization of the first step of splicing as a means for 3' end formation necessitates that splicing is aborted after the first cleavage reaction. Here we have shown how the spliceosome can be converted into a 3' end processing machinery in the context of TER1 RNA: A combination of RNA elements that promote spliceosome assembly and the first step, but encumber the transition to the second step is sufficient to uncouple the two reactions and promote the release of the free 5' exon via the discard pathway. In addition to providing further insights into the mechanism of telomerase biogenesis, these studies shed new light onto aspects of the splicing reaction that are still only poorly understood.

III.4.a: A conserved role for distance and Py-tract in regulating the second step.

The most obvious feature that distinguishes the TER1 intron from the majority of introns in fission yeast is the distance between the BP and 3' ss. At 23 nucleotides, TER1 is in the top three percentile of introns with respect to BP to 3' ss distance. Although, the BP to 3' ss distance varies considerably between introns, there is a clear species-specific optimum into which the majority of introns fall. In budding yeast, the mean distance between BP and 3' ss is around 39 [143, 146]. Increasing the distance in the context of a specific intron hampers progression through the second step [147, 148]. In humans, BP to 3' ss distances also cluster in a fairly narrow window around 25 nucleotides [149, 150], with the exception of a small number of introns that contain distant branch sites that can be greater than 150 nucleotides from the 3' ss [128]. Altering the distance between BP and 3' ss in the context of in vitro splicing substrates results in accumulation of splicing intermediates [92, 151]. So despite differences during spliceosome assembly between yeasts and metazoans, uncoupling of the two steps of splicing as a consequence of a long BP to 3' ss distance has been observed in diverse systems.

In vertebrates, binding of U2AF to the polypyrimidine tract promotes U2 recruitment [139, 140] and mutations that shorten the Py-tract or replace pyrimidines with purines interfere with spliceosome assembly [135, 152]. Similarly in fission yeast, the Py-tract has been shown to function prior to the first step of splicing [91]. Here, a preformed complex of SF1 and the small and large subunits of U2AF bind cooperatively to the BS, Py-tract and 3' ss [134]. The prominent role of a Py-tract during the early stages of

splicing has complicated the analysis of its roles during the second step. However, synthetic splicing substrates containing competing splice sites revealed a function of the Py-tract in 3' ss choice and second step kinetics [92]. Notably, completion of the second step was only dependent on a Py-tract in the context of a long distance between the BP and 3' ss. Similarly, in budding yeast, where introns were long thought not to contain functionally important Py-tracts, examination of model substrates revealed a critical function for uridine-rich motif in the identification and utilization of distant splice acceptor sites [153]. A role for Py-tracts in promoting the second step therefore appears to be widely conserved as well.

What is more surprising is that a temperature sensitive allele of U2AF59 would compromise completion of the second step at the restrictive temperature (Fig. 3.4C). At least in mammalian splicing extracts, the large subunit of U2AF is replaced by 3 proteins of the U5 snRNP prior to the second step of splicing [154] and is not detected in the C complex [155]. Fission yeast may differ in this regard or U2AF may remain transiently associated with the spliceosome and promote the second step even after direct interactions with the Py-tract has ceased.

III.4.b: A strong branch site can impede completion of splicing.

Binding of the U2 snRNP to the BS is one of the earliest, and in some cases the earliest, events in spliceosome assembly [156], and the extent of sequence complementarity correlates positively with the rate of assembly and first cleavage reaction [157].

Pioneering cross-linking studies in the early 1990s suggested that the U2 – BS association is maintained through both steps of splicing [144] and consistent with this assertion, the U2 snRNP remains part of the spliceosome after the first step [158]. What has remained puzzling in this regard is that BS mutations to either G or U inhibit the second step of splicing in yeast and human cell extract [159-162]. Using an orthogonal splicing system involving a highly substituted and reporter-specific second copy of U2, it has been shown that a change in the position of the bulged nucleophile imposes a block on the second step [163]. One interpretation of these results is that the nucleotide change affects the flexibility of the RNA and this impairs the conformational rearrangements that position the RNA correctly for the second transesterification reaction.

An indication that the U2 – BS interaction may unwind prior to 5' splice site cleavage came from the observation of trans-splicing products between the 5' end of U2 and the 3' exon of a reporter gene in budding yeast [164]. While those results indicated that U2 – BS base pairing is more dynamic than previously thought, it did not suggest that maintenance of the U2 – BS interaction is detrimental to the completion of splicing. Our analysis now showed that in the appropriate context sequence complementarity between the BS and U2 antagonizes the completion of splicing and favors uncoupling of the two steps. In light of these observations, it seems likely that at least partial or transient disruption of U2 – BS base pairing is part of the transition from the first to the second step of a normal splicing reaction. The molecular function of this rearrangement will need to be defined by future studies.

III.4.c: A 'proofreading pathway' generates functional telomerase RNA

Our results show that a combination of intronic elements hamper the transition from the first to the second step of splicing and are critical for the release of the TER1 5' exon after the first step of splicing. The implication of Prp22 and Prp43 in TER1 processing further suggests that the 'discard pathway' is not merely acting in the recovery of spliceosomes from defective or suboptimal substrates, but in the case of TER1 promotes the release of a functional product. As the mechanism for attenuating the transition from the first to the second step as well as the molecular machinery are widely conserved, additional targets for this pathway are likely to be discovered.

Chapter Four:

U6 hyper-stabilization with 5'ss generates the 3' end of telomerase RNA in *S. cryo*philus and *S. octosporus*

IV.1: Abstract

The mature 3' end of fission yeast telomerase RNA (TER1) is generated by spliceosomal cleavage akin to the first step of splicing. In the case of TER1, uncoupling the two steps of splicing is achieved via a combination of strong branch site interactions with U2 snRNA and a long distance between branch point and 3'ss, which stimulates cleavage. Whether TER1 is unique in employing spliceosomal cleavage or other examples exist is unknown. We now demonstrate that the majority of uncoupling elements are selected against introns in protein coding genes. In addition, we have determined that telomerase RNA in *S. cryophilus* and *S. octosporus* are also cleaved. Unlike TER1, hyper-stabilization of U6 snRNA with 5'ss impedes the 2nd step of splicing allowing cleaved form accumulation in *S. cryophilus* and *S. octosporus*. In summary, our studies indicate a general role for the spliceosome in telomerase RNA processing in fission yeast.

IV.2: Introduction

Fine tuning RNA expression is a critical step in controlling gene expression [79]. The generation of mature RNA inside the cells is one of the most complicated events in all cell biology and biochemistry [80]. Not only is there an elaborate repertoire of proteins that direct RNA polymerase II to begin synthesis of precursor RNA, but in addition accurate processing and correct localization of mature RNA is equally, if not, more complex [80].

One of the pivotal events in RNA processing is precursor RNA splicing, which involves tight coupling of intron removal and exon ligation. In the first step of intron excision, the 2′ hydroxyl of the branch point (BP) adenosine, located within the branch site (BS) sequence, attacks the sugar-phosphate bond at the 5′ splice site (5’ ss), forming a 2’-5’ linkage and producing a free 5′ exon and a branched species, termed the lariat intermediate [114]. In the second reaction, the now exposed 3′ hydroxyl of the upstream exon attacks the phosphate at the 3′ splice site (3’ ss), yielding ligated mRNA and the lariat form of the intron. The importance of studying splicing at a molecular level is illustrated by the finding that more than 30% of human diseases are caused due to mutation in splice sites or the spliceosome components [84].

Intron removal from precursor RNA is catalyzed by the spliceosome, a mega dalton complex, whose assembly and function require transition between multiple conformations [90]. Based on accumulating genetic and biochemical data, we now understand that 5’ ss is positioned for first step catalysis via base pairing with the

conserved ACAGAG region of U6 snRNA while the region surrounding the branch point is paired with the conserved GUAGUA sequence of U2 snRNA. In addition, U5 snRNA interacts with both exons, maintaining optimum position required for catalysis.

Notably, splicing studies have focused on protein encoding genes. There, a strong bias exists to select for tight coupling of the two transesterification reactions, as release of splicing intermediates after the first cleavage reaction would result in degradation and hence block gene expression [165]. Predictably, examining current splicing literature has led to the deduction that the two steps of splicing are normally linked to one another [95].

Telomerase is the ribonucleoprotein complex that synthesizes and maintains telomeres. The catalytic core of telomerase is composed of a non-coding RNA (TER) and protein subunit (TERT) which is the reverse transcriptase. Although many insights have been gained regarding telomerase function in telomere repeat addition and processivity, we know little about the biogenesis of telomerase in eukaryotes [62].

The fission yeast telomerase RNA harbors an intron immediately downstream of its mature 3' end. Unlike canonical splicing, only the first step cleavage is required to generate mature TER1 [116]. Particularly, TER1 intron undergoes cleavage irrespective of the identity of the first and second exons [116, 165]. Conversely, replacement of TER1intron with a heterologous intron from a protein-coding gene leads to completion of splicing. Notably, the TER1 branch site is fully complementary to the BS binding sequence in U2 snRNA, whereas the BS in FCP1 intron 1 deviates by two nucleotides from the U2 complement, resulting in a mismatch and the replacement of a G-C base pair

with a G: U wobble. Indeed, restoring the two base pair interactions, promotes spliceosomal cleavage of the heterologous intron. Furthermore, increasing the distance between branch point and 3' ss to 38 nucleotides stimulated spliceosomal cleavage at the expense of splicing, whereas reducing the distance promoted the 2nd step of splicing. In essence, a strong branch site in the context of a long distance between branch point and 3'ss impedes the 2nd step of splicing, triggering release of 5' exon via Prp22 and Prp43 mediated discard pathway [165]. The observation that intronic elements can prohibit the transition to the 2nd step of splicing to generate the 3' end of a non-coding RNA suggests that other examples of cleaved RNAs could exist. However, the identity of these RNAs has been elusive.

Here, we demonstrate that the majority of the introns present in protein coding genes in *S. pombe* have selected against uncoupling elements that promote spliceosomal cleavage. We demonstrate one example where an intron in a protein-coding gene is cleaved implicating a function for spliceosomal cleavage in regulating gene expression. Further, we establish that telomerase RNA in *S. octosporus* and *S. cryophilus* is processed by spliceosomal cleavage albeit via a different mechanism. Our studies bring out new insights as to how the spliceosome can function in 3' end processing.

IV.3: Results

IV.3.a: Global selection against spliceosomal cleavage.

Sequence elements that hinder transition of the spliceosome during the second step of splicing will be detrimental for protein coding genes. To confirm the above hypothesis, we examined those fraction of annotated fission yeast introns that share the combination of a strong branch site, long distance and weak Py-tract and are therefore likely candidates for spliceosomal cleavage [165]. While the use of the spliceosome in 3' end processing may be limited to non-coding RNAs, spliceosomal cleavage could also serve a function in regulating gene expression by reducing the amount of spliced mRNA that is produced [129].

To identify putative branch sites we first used a Gibb's sampler algorithm to identify motifs that are over-represented in *S. pombe* introns (excluding the 5' and 3' splice sites, see Experimental Procedures for details). The positional weight matrix (PWM) generated by this approach strongly resembled known branch sites (Fig. 4.1A). We then used the PWM to score every position in the intron data set. This score (referred to as S_{BS}) ranges from 1.0 to 7.86. Plotting the S_{BS} values versus their positions relative to the acceptor site revealed that sites with scores > 6 clustered near the 3' splice site (Fig. 4.2A) Only 189 annotated introns (4%) returned no match for a site with $S_{BS} > 6.0$ and were excluded from further analysis. 42% had exactly one site with a score ≥ 6 and the remaining 54% had more than one such site (Fig. 4.2B). For introns with more than one site ≥ 6, the site closest to the 3' ss was chosen as the putative BS.

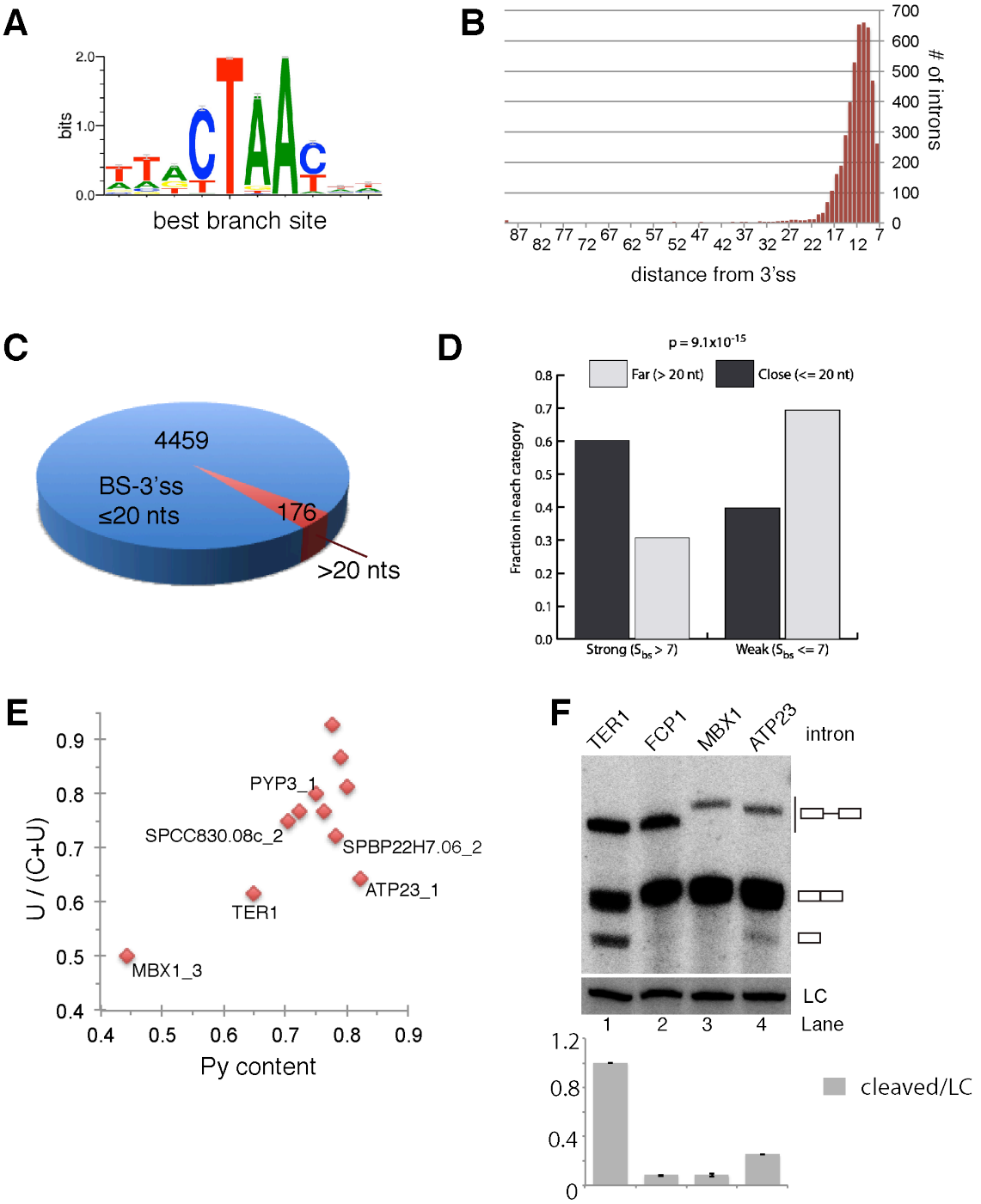


Figure 4.1: Global analysis indicates selection against the intron structure that favors release after the first step of splicing. **(A)** Sequence logo representation of the *S. pombe* branch site PWM constructed as described in Experimental Procedures. **(B)** Histogram

showing the locations of branch sites relative to the 3' splice site. A vertical arrow marks the position for TER1. **(C)** Pie chart representation of introns with long and short distance between BP and 3' ss, respectively. **(D)** Graphical representation of the relationship between BS strength and distance to 3' ss. **(E)** Scatter plot illustrating Py-tract strength for introns that resemble TER1 in BS sequence and distance to 3'ss. **(F)** Northern blot for the reporter RNA containing the indicated introns. A probe against snR101 was used as loading control (LC). Data represented as mean +/- SEM with n=2.

With putative branch sites mapped in this manner, the mean distance from BP to 3' ss was 11 nucleotides (Fig. 4.1B), a finding that is in good agreement with the positions of predicted branch sites based on a subset of fission yeast introns [89]. For 96% of introns the distance was below 20 nucleotides and only 176 introns had a distance between the predicted BP and 3' ss in the range of where we detect uncoupling of splicing for TER1 (Fig. 4.1C). We next compared the strength of the branch sites between these introns and those with a shorter distance. Whereas 60% of introns with a BP located within 20 nucleotides from the acceptor have strong predicted branch sites (defined as a score above 7.0), only 30% of introns with a distance greater than 20 nucleotides have strong branch sites (Fig. 4.1D). This statistically significant difference (χ^2 p-value= 9.1×10^{-15}) between these categories demonstrates that the combination of a strong BS and a long BP to 3' ss distance is underrepresented in introns of protein encoding genes.

We next plotted the uridine to pyrimidine ratio over the pyrimidine content as a measure of Py-tract strength for the 10 introns that share the consensus BS sequence with TER1 and have a BP to 3' ss distance of 20 or more nucleotides (Fig. 4.1E). With the exception of MBX1 intron 3, all candidates have stronger Py-tracts than TER1. The two introns that resembled TER1 most closely were cloned into the reporter construct and examined by northern. No cleavage was observed for MBX1, but the free 5' exon was readily

detected for ATP23 (Fig. 4.1F). To assess whether the cleavage of the ATP23 RNA was a consequence of placing the intron into a heterologous context, we examined the endogenous message in the context of wild type fission yeast and a *rrp6* deletion. The latter compromises the activity of the nuclear exosome and is thus expected to stabilize a released 5' exon. We indeed observed the spliceosomal cleavage product for ATP23 in both strain backgrounds (Fig. 4.2C). However, unlike TER1, the spliced form of ATP23 RNA dominated over the cleaved form in steady state. In summary, the combination of RNA elements that promote the uncoupling of the two steps of splicing is largely absent from the vast majority of annotated introns, but at least in one case cleaved pre-mRNA is detected. Importantly, the global analysis exposed a fervent anticorrelation between a strong branch site and a long BS – 3' ss distance.

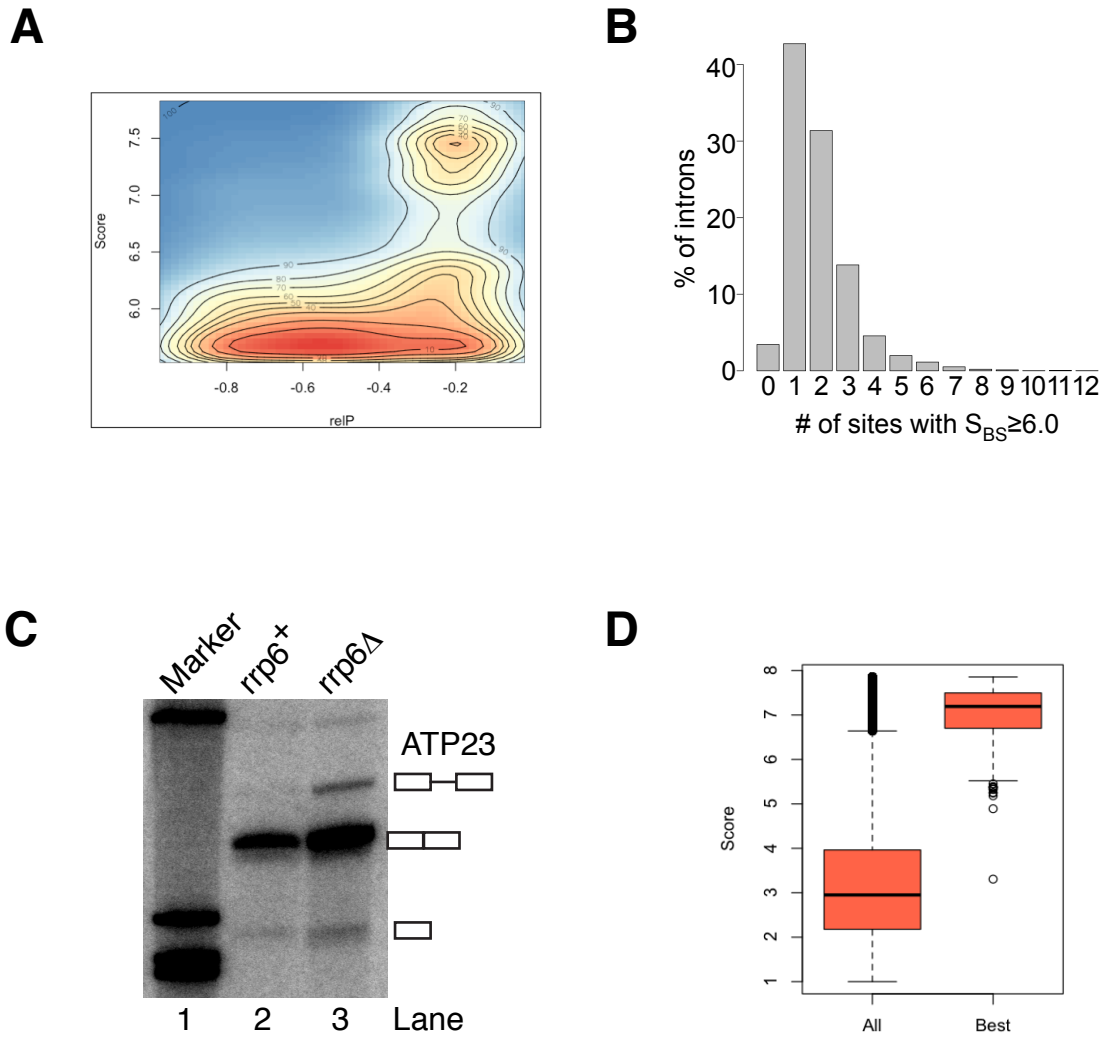
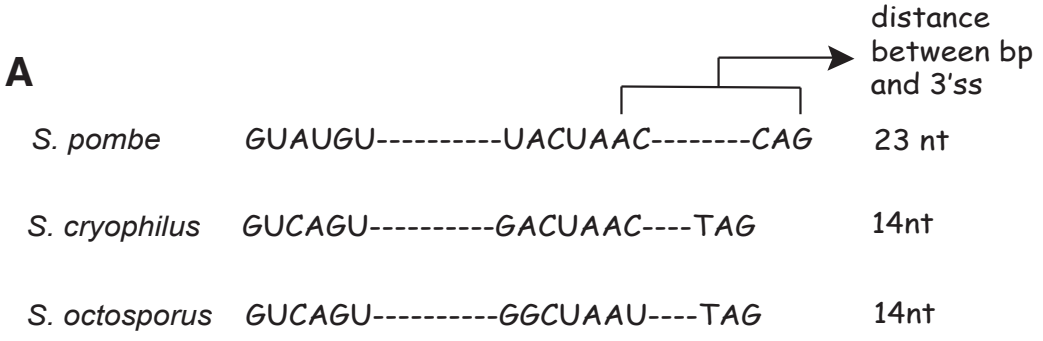
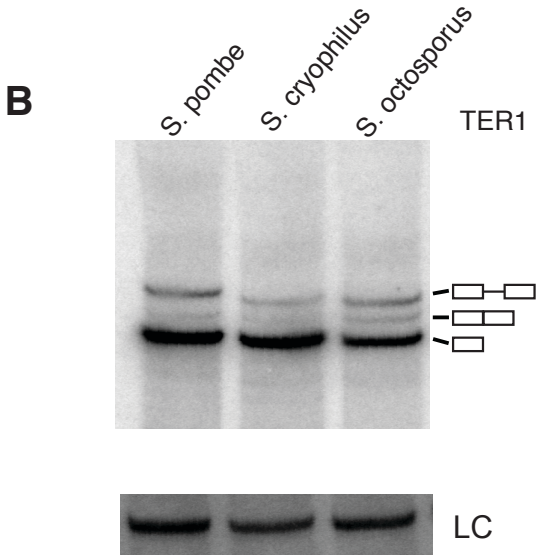


Figure 4.2: Branch site prediction and analysis of Atp23. **(A)** Heatmap showing the S_{BS} score (y-axis) for each intronic 10-base k-mer versus its relative position in the intron (x-axis). Only kmers with scores > 5.5 are shown. The relative position (relP) is the normalized position relative to the 3'ss (i.e. the number of nucleotides from the acceptor site divided by the length of the intron). **(B)** Bar graph based on the number of branch sites with S_{BS}>6 found in each intron expressed as percentage of all introns. **(C)** Northern blot on total RNA isolated from wild type and rrp6D strains, subjected to RNaseH cleavage and probed for ATP23. **(D)** Comparison of BS strength between all sequences that match the PWM and the best match for each intron.

Considering that majority of introns in protein coding genes are selecting against uncoupling elements, we reasoned that if spliceosomal cleavage is essential for telomerase RNA processing, the mechanism should be conserved in closely related fission yeast species. With this in mind, we identified the telomerase RNA subunits in *S. cryophilus* and *S. octosporus* by sequence homology to TER1. Notably, both TERs contained intron like sequences beyond their 3' end (Fig. 4.3A). To determine if these introns can promote spliceosomal cleavage, we replaced the *S. pombe* TER1 intron with introns from either *S. cryophilus* or *S. octosporus*. Despite displaying limited sequence homology, both introns promote spliceosomal cleavage indicating that the function for spliceosome in 3' end processing is conserved in fission yeast (Fig. 4.3 B and C).


A
distance between bp and 3'ss
S. pombe GUAUGU----------UACUAAC--------CAG 23 nt
S. cryophilus GUCAGU----------GACUAAC----TAG 14nt
S. octosporus GUCAGU----------GGCUAAU----TAG 14nt


B
S. pombe
S. cryophilus
S. octosporus
TER1
LC

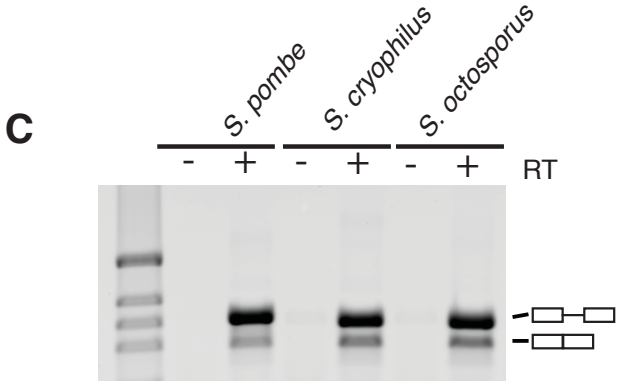

C
S. pombe
S. cryophilus
S. octosporus
- + - + - +
RT

Figure 4.3: 3' end processing of telomerase RNA is conserved among fission yeast species. A) Schematic of the intron sequences in the 3 different species. B) Northern blot analysis of ter1 constructs containing the different introns and the loading control (LC). B) RT-PCR across the intron to assess the abundance of the spliced form.

IV.3.b: Hyper-stabilizing U6 snRNP with 5'ss promotes spliceosomal cleavage

Next, we wanted to examine whether mechanism of telomerase RNA processing is conserved in the different species of fission yeast. Unlike TER1, both *S. cryophilus* and *S. octosporus* introns contain a short distance between the bp and 3'ss which will promote completion of splicing irrespective of the branch site strength [165]. We wondered if the requirement for a long distance between the bp-3'ss is offset by the presence of other uncoupling elements. Particularly, the unique aspect of both introns is the 5'ss GUCAGU, where the weak A-U base pair interaction at +3 position in the 5'ss is replaced with a strong C-G base pair interaction (Fig. 4.4B). Elegant studies in budding yeast has shown that increasing base pair interactions of 5'ss with U6 traps the spliceosome in the first step conformation, thereby inhibiting the 2nd step of splicing [166]. To test if U6 stabilization at 5'ss is essential for spliceosomal cleavage of *S. cryophilus* and *S.octosporus* TER, we changed the 5'ss to GU**A**AGU. Indeed, a single nucleotide change at the 5'ss, promotes potent completion of the 2nd step leading to accumulation of the spliced form at the expense of cleaved form (Fig. 4.4C and D).

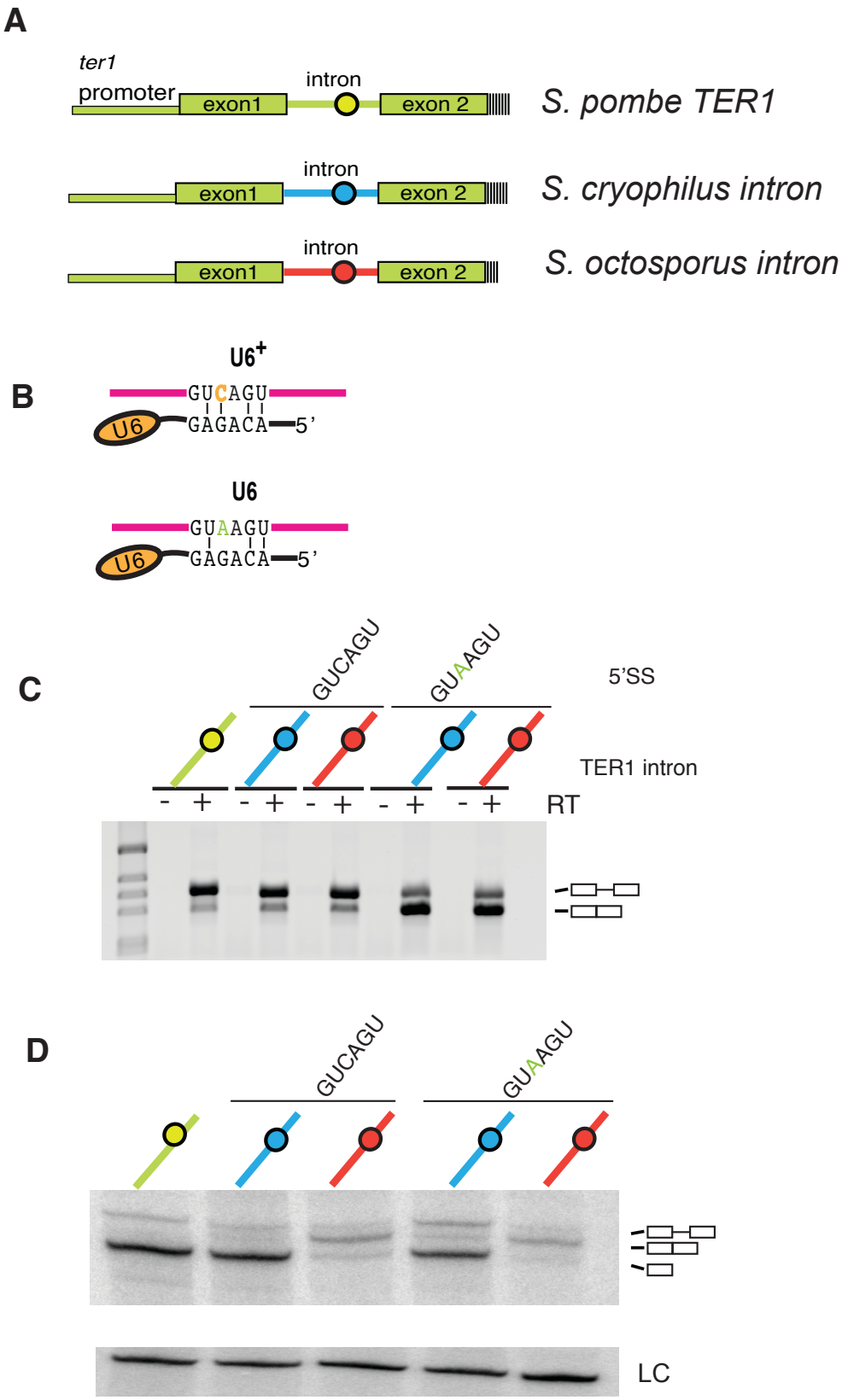

A
ter1
promoter
exon1
intron
exon 2
S. pombe TER1
intron
exon1
exon 2
S. cryophilus intron
intron
exon1
exon 2
S. octosporus intron
B
U6+
GUCAGU
U6
GAGACA
5'
U6
GUAAGU
U6
GAGACA
5'
C
GUCAGU
GUAAGU
5'SS
TER1 intron
- + - + - + - + - +
RT
D
GUCAGU
GUAAGU
LC

Figure 4.4: Hyper-stabilization of U6 with 5'ss as an alternative method for spliceosomal cleavage**.** A) Schematic of the three different fission yeast telomerase RNAs B) Illustration of U6 binding to 5'ss C) RT-PCR of 5'ss mutants D) Northern blot of the different *ter1* constructs carrying the fission yeast introns with or without 5'ss mutations.

These observations demonstrate that the spliceosome has a conserved function in generating the 3' ends of telomerase RNA in fission yeast. More importantly, it suggests there are multiple ways to promote cleavage over splicing, indicating uncoupling of the two steps of splicing could have a critical role in gene regulation and RNA processing.

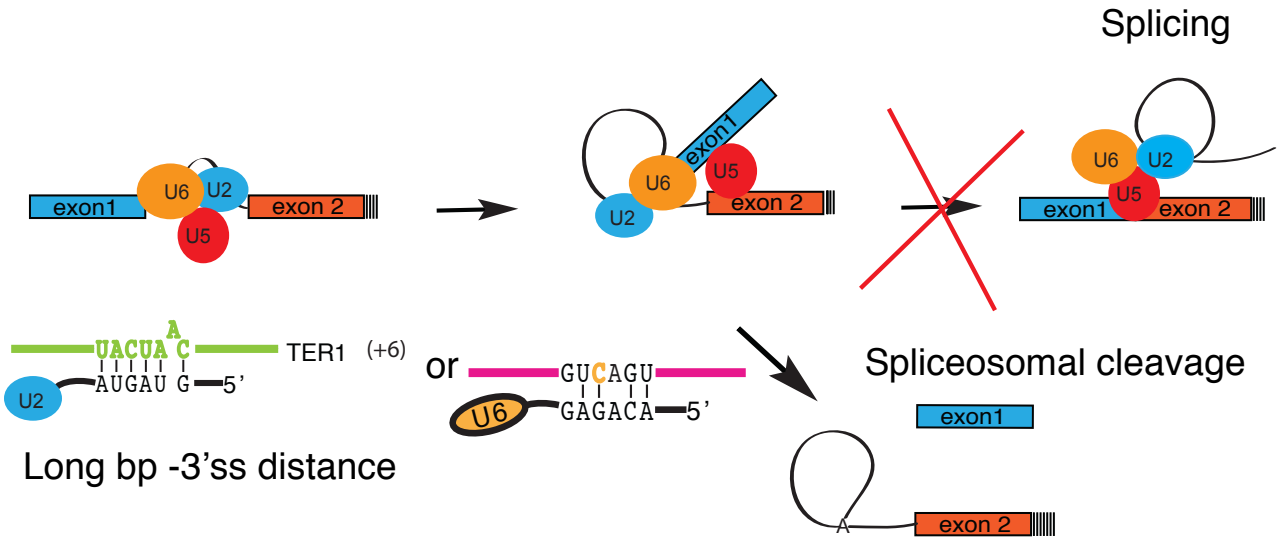


Figure 4.5: Model illustrating how sequences within the spliceosome can convert the function of spliceosome from splicing to 3' end processing.

IV.4: Discussion

The realization that a combination of sequence elements can switch the function of spliceosome from splicing to 3' end processing was surprising at first [165]. In the case of protein coding genes, uncoupling the two steps of splicing and releasing splicing intermediates will block gene expression. Here, we have shown that majority of protein coding genes in *S. pombe* alleviate this problem by maintaining a short distance between

bp and 3'ss, allowing completion of the two steps of splicing. Surprisingly, we demonstrate that weaker branch sites are selected in introns that contain a long distance between bp and 3'ss, presumably to prevent uncoupling the two steps of splicing. More importantly, we establish that spliceosomal cleavage is a general mechanism for telomerase RNA processing in fission yeast. In addition to providing insights into telomerase RNA biogenesis, these studies shed light on the mechanism of splicing in non-coding RNAs, which is still poorly understood.

IV.4.a: Selection against uncoupling elements in S. pombe

Pioneering studies established that UACUAAC is the preferred branch site for U2 snRNA binding from yeast to humans, and hence essential for splicing [167]. Elegant crosslinking studies had demonstrated that U2 snRNP is associated with the intron until completion of splicing [168]. These studies had led to the conclusion that presence of a strong branch site promotes the first step of splicing and does not influence splicing afterwards. The first clue that U2 – BS interaction might be disrupted prior to the 5' splice site cleavage came from the observation of trans-splicing products between the 5' end of U2 and the 3' exon of a reporter gene in budding yeast[164]. However, the significance of the disruption of U2-BS interaction has remained elusive. We have shown that a strong branch site is critical for inhibition of the 2nd step of splicing for TER1 processing [165]. Whether disruption of U2-BS interaction is a general phenomenon for completion of splicing has remained elusive. To solve this conundrum, we analyzed genome wide the strength of branch sites in *S. pombe.* From these studies it is clear that

the inhibitory role of branch site is context dependent. Though introns with a short distance between bp and 3'ss can tolerate strong branch site, there is a bias against strong branch sites in introns with long distance between bp and 3'ss (Fig. 4.1). In light of our results and the observation that U2snRNP components SF3a and SF3b are destabilized after the first step of splicing [158, 169], suggests strong U2-BS interaction is detrimental for the second step of splicing.

Furthermore, at least for one intron that contained the uncoupling elements, we detected a cleaved isoform in a protein-coding gene Atp23 indicating that spliceosome could function in reducing the levels of mRNA thereby playing a critical role in gene regulation. Nevertheless, there exists a fervent anti-correlation between the strength of branch site and distance between bp and 3'ss.

IV.4.b: U6 hyper-stabilization with 5'ss generates 3' end of telomerase in S. cryophilus and S. octosporus

The identification that the spliceosome generates the mature form of telomerase RNA in *S. pombe* led to the question how general is this mechanism for TER processing. With this intention, we identified the telomerase RNA in closely related species of *S. cryophilus* and *S. octosporus.* Despite displaying little sequence similarity, the telomerase RNA introns from *S. cryophilus* and *S. octosporus* are cleaved efficiently in the context of TER1, demonstrating a conserved role for the spliceosome in TER processing across fission yeast species.

For promoting cleavage over splicing, *S. pombe* TER1 utilizes a combination of a strong branch site and a long distance between bp and 3'ss [165]. In contrast, those uncoupling elements are absent in other species of fission yeast. Instead, both TER undergoes extensive base pair interactions with U6 snRNA via their 5'ss, thereby stabilizing the first step conformation and impeding the second step.

These observations add to the growing evidence that supports the 'two state' model of the spliceosome [170]. It is now clear that the spliceosome has to transition between the two conformations to catalyze the two transesterifications. Our finding is reminiscent of studies in budding yeast where stabilizing U6 snRNA base pair interaction with 5'ss in the context of a protein coding gene inhibits the transition to the second step [166]. On the other hand, weakening the 5'ss base pair interaction with U6 allows completion of splicing.

In essence, these studies suggest that extensive base pair interaction of either U2 or U6 with branch site and 5'ss respectively, can inhibit transition to the second step, thereby promoting release of the 5' exon via the discard pathway (Fig. 4.5). Furthermore, these results also indicate that spliceosomal cleavage rather than identity of sequence elements, is under strong selective pressure for telomerase RNA processing in fission yeast. The observation that introns are present beyond the 3' ends of telomerase RNAs in several species of *Candida* and *Aspergilli* also suggest that spliceosome has a general function in telomerase RNA processing in yeast and filamentous fungi [127, 171].

IV.4.c: Implications for evolution of splice sites

Considering the indispensable role for U6 and U2 snRNA in splicing, the splice site binding regions, ACAGAG in U6 and GUAGUA in U2 are essentially identical from yeast to humans [95]. In contrast, the 5'ss and branch sites are loosely conserved in mammals. Furthermore, recent studies that sequenced the lariat of housekeeping genes in humans identified the branch sites to be extremely degenerate [172]. Although, this paradox has been noted earlier the exact reasons has not been established [167]. Our observation demonstrating that extensive base pair interactions between 5'ss and U6, BS and U2 can uncouple the two steps of splicing, suggests introns in protein coding genes have evolved to allow coupling of the two steps at the expense of splice strength. In conclusion, our studies establish that introns in non-coding RNAs and mRNAs have evolved different sequences that enable them to utilize the spliceosome for 3' end processing or splicing, respectively.

Chapter Five:

Conclusion and Future directions

V.1: Significance of using spliceosomal cleavage for 3' end processing

Research examining how linear chromosome ends solve the 'end protection' and the 'end replication' problems has been one of the most fruitful areas of science. Within this field, vast knowledge has been gained about how telomeres and telomerase allow eukaryotes to safely maintain linear chromosome ends. In addition, critical roles for telomeres and telomerase in cancer and degenerative diseases have been uncovered. The rapid progress in our understanding of telomerase was possible due to the insights obtained from model organisms. Although, many key concepts and principles have emerged regarding the role of telomerase in telomere maintenance, we still know little about how telomerase biogenesis occurs in eukaryotes. Since small nuclear RNA are extremely conserved between *S. pombe* and humans [173], elucidating TER1 processing provided an excellent tool not only for uncovering telomerase biogenesis, but also for probing aspects of non-coding RNA processing that remain unclear.

The observation that only the first step of splicing is required for TER1 processing was surprising initially. From elucidating splicing mechanism in protein coding genes, it was clear that the two steps of splicing are tightly coupled. However, in this context, release of splicing intermediates will lead to degradation and thereby block gene expression. Nevertheless, due to the contrasting pieces of evidence in hand, we decided to solve a

major question plaguing the field: How is the spliceosome converted into a 3' end processing machinery?

The key for dissecting RNA elements that promote the first step of splicing but hinder the second step, was the development of a reporter system in which similar levels of spliced and cleaved form are produced and can be quantified in a direct assay. Using this unique approach, we have shown that a long distance between the branch point and 3'ss inhibits the second step of splicing. More importantly, we have demonstrated that strong branch site base pair interactions with U2 snRNA can uncouple the two steps of splicing and the 5' exon can be released via the discard pathway. These observations lead to critical questions which is the major contribution of the research in Chapter Three and Four to the fields of RNA processing, splicing and telomerase biogenesis.

One of the most pressing questions put forth in Chapter Three and Four is why the spliceosome is used to generate 3' ends of telomerase RNA when other energy efficient ways are available? In other words, are there conditions in which splicing can be completed allowing an elegant way to control telomerase levels. Studying conditions in which telomerase activity is no longer required would be an obvious starting point. *S. pombe* cells grow rapidly in log phase when plenty of nutrients are available. However, cell division slows down with decrease in nutrient availability and halts on reaching stationary phase. Since telomerase promotes cell proliferation, it is possible that it is actually detrimental for stationary phase progression and must be turned off. Preliminary experiments indicate that the spliced form of telomerase RNA accumulates as cells enter

stationary phase (Fig. 5.1B). In contrast, Prp43 mRNA levels decrease during the same time period consistent with the finding in Chapter Three that Prp43 hinders the second step of splicing (Fig. 5.1A).

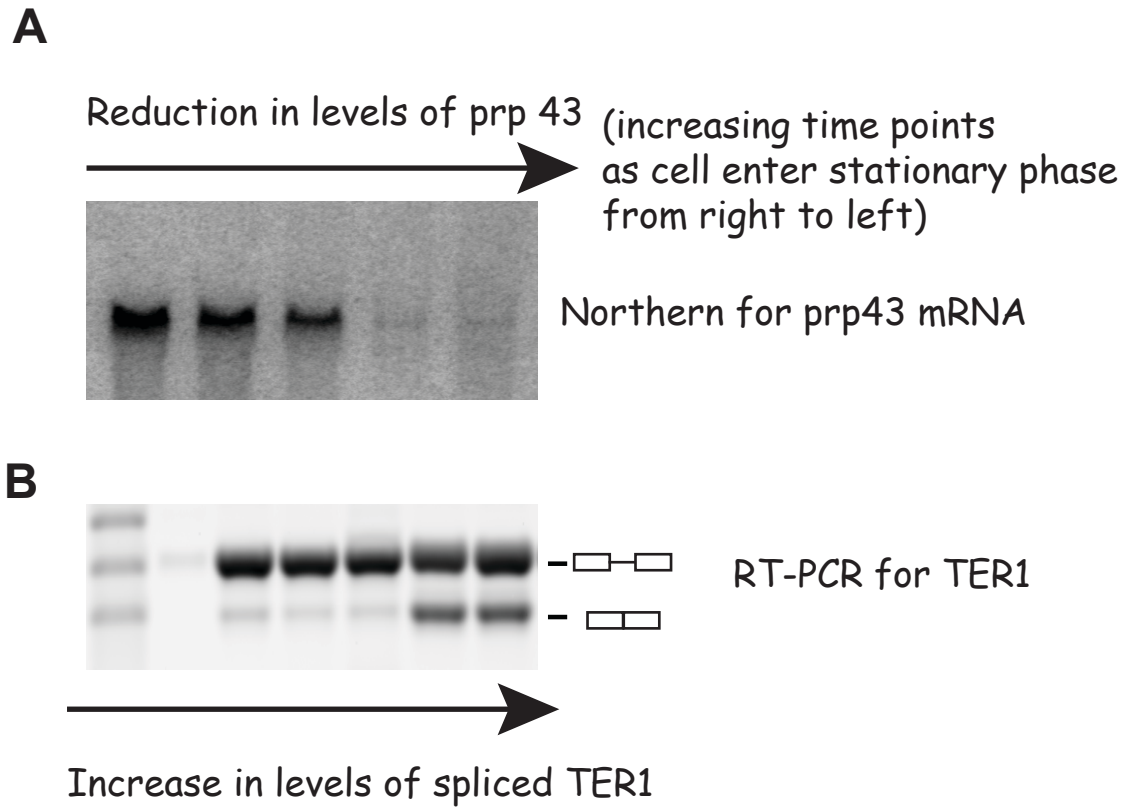


Figure 5.1: Spliced form increases in stationary phase. A) Northern blot of total RNA isolated from wild type cells at increasing time points as cells enter stationary phase and probed for Prp 43 mRNA. B) RT-PCR detecting the precursor and spliced forms of TER1 in the same time points as A.

With this observation as a starting point, one could perform a series of experiments to test if spliceosomal cleavage is used to regulate telomerase levels. An important limitation in studying the function of spliced form is that despite being polyadenylated, it is extremely unstable suggesting that additional mechanisms exist within the cell to ensure low levels

of spliced TER1. We have shown that the nuclear exosome component Rrp6 downregulates spliced TER1 (data not shown). Thus, spliced form can be stabilized, by expressing *ter1-55* mutant, which promotes completion of splicing, in a *rrp6* deletion strain. In this context, we can examine if telomerase is functional by performing a Southern blot to determine telomere length. In addition, telomerase activity assays can provide insights whether the spliced form can support catalysis. In summary, experiments suggested here can pinpoint if spliceosomal cleavage regulates telomerase RNA levels in different cellular conditions.

V.2: Role for spliceosomal cleavage in controlling mRNA levels

Generally speaking, transcriptional networks play an important role in controlling gene expression[174]. However, some genes must be regulated in response to sudden changes in environment, and achieving this via transcription is difficult. Equally important, RNA processing provides additional layers of regulation thereby fine tuning gene expression [79, 80]. Considering this point, the detection of a cleaved form for at least one protein coding gene, Atp23, suggests that spliceosomal cleavage may be required for controlling mRNA levels.

Under circumstances in which an excess amount of protein is detrimental to the cell, the ability to cleave and degrade the 5' exon, thereby lowering the amount of mRNA available for translation, would allow the cells to rapidly adapt to environmental changes. An obvious experiment to test this hypothesis and continue from the results obtained in Chapter Four is to replace the cleaved intron with a well spliced intron to examine

whether completion of splicing is detrimental to the corresponding cellular process. The change in splicing can be measured by quantitative RT-PCR and protein expression can be detected via western blotting. In essence, these experiments will not only suggest new insights on gene regulation but will change our approach regarding function of introns in protein coding genes.

In light of our findings in Chapter Three and Four, the notion that spliceosomal cleavage may have non-canonical functions in controlling gene expression is gaining momentum. For example, it has been discovered that introns that undergo a slow 2nd step of splicing, are the source of small silencing RNA in the opportunistic fungal pathogen *Cryptococcus neoformans[175]*. Due to suboptimal intronic sequences, the spliceosome stalls on the substrate mRNAs after the first step of splicing. Following the arrest, the spliceosome coupled and Nuclear RNAi complex (SCANR) uses the debranched lariat intermediate as a template for small-RNA biosynthesis which feeds into the silencing machinery and down regulates spliced mRNA. In summary, our studies are shedding light on the non-canonical functions of the spliceosome in regulating gene expression.

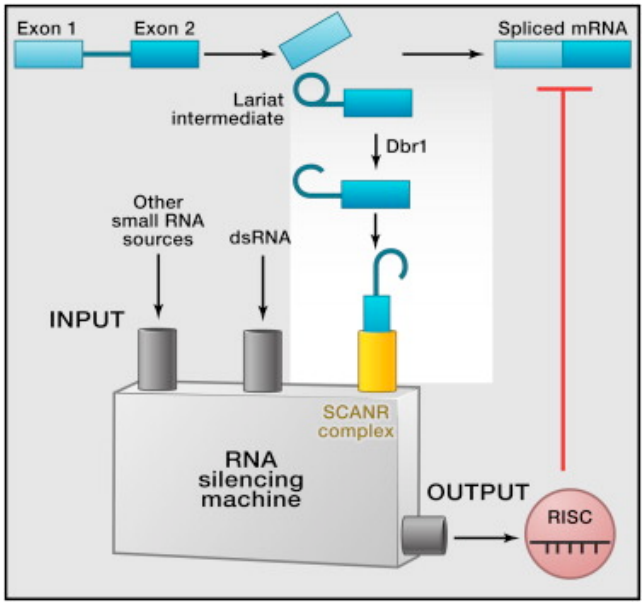


Figure 5.2: A spliceosome-based source of small RNAs programs the silencing machinery in *C. neoformans*, adapted from [176]. Suboptimal introns with slow second step of splicing stall the spliceosome and the lariat intron is used as a substrate for generating small RNAs for RNA silencing. The resulting small RNAs are incorporated into the RNA interference-silencing complex (RISC) that degrades spliced form.

V.3: Implications for Telomerase RNA processing in eukaryotes

Though spliceosomal cleavage could have additional functions in controlling telomerase expression and gene regulation, the results obtained in Chapter Three and Chapter Four point towards a general role for the spliceosome in telomerase RNA and possibly non-coding RNA processing. More importantly, the conservation of intronic sequences in recently identified TER from *Candida* and *Aspergilli* indicate that spliceosomal cleavage is a near obligatory mechanism for telomerase RNA processing at least in lower eukaryotes. Although it seems redundant, dissecting the sequence elements required for splicing uncoupling in *Candida* and *Aspergilli* TER will be crucial for elucidating aspects of the splicing mechanism that are only poorly understood [127, 171].

Similar to TER1, longer polyadenylated forms have been detected for human telomerase RNA (hTR) but whether they are precursors is unclear. As discussed earlier, the degenerate splice sites in higher eukaryotes makes it extremely difficult to identify introns in humans. The main evidence pointing towards a function for spliceosome in hTR processing stems from studies in which inhibition of the spliceosome using the drug isoginketin [177] yields accumulation of precursor (Gaspari M and Baumann P, unpublished). However, this may be due to an indirect effect and further research is required to dissect this problem. Considering that U2 snRNA in *S. pombe* and humans are remarkably similar [173] and Prp22, Prp43 are conserved in humans, manipulating these components could provide novel insights into hTR processing [178, 179].

One argument against extrapolating the data in *S. pombe* to humans is that the mode of spliceosome assembly is different between the two. Due to the large size of human introns, spliceosome assembly occurs by exon definition, whereas short introns in *S. pombe* can be recognized by intron definition. Although, the mode of intron recognition is different, the mechanism of catalysis is essentially identical from yeast to humans [95]. For example, it has been shown *in vitro* using human splicing extracts, that increasing the distance between branch point and 3'ss to 66 nucleotides in the context of pADml reporter system, leads to uncoupling of the two steps of splicing [92] (Kannan, R and Baumann, P; unpublished). However, to uncover the role of the branch site in this regard, a systematic analysis similar to those performed in budding yeast and *S. pombe* is essential. Considering this point, a recent study sequencing lariats of human introns has revealed a strict distance constraint for the branch point from the 3'ss, irrespective of the

total length of the intron [172]. In conclusion, these results indicate that insights obtained from *S. pombe* splicing can enhance our understanding of the human spliceosome.

V.4: To determine other RNAs that are processed by spliceosomal cleavage

The observation that either U6 hyper-stabilization with 5'ss or strong branch site base pair interactions with U2 snRNA can impede the transition to the second step of splicing thereby promoting 5' exon accumulation indicates different mechanisms are in place for promoting spliceosomal cleavage. For uncovering other RNAs that are cleaved, the obvious candidates would be non-coding RNAs that contain introns. In most conditions, the cleaved 5' exons would be degraded. In contrast, if the 3' ends are stabilized by proteins similar to TER1 or due to formation of higher order structures [180], spliceosomal cleavage could result in a functional non-coding RNA.

To identify non-coding RNAs that are cleaved, we can apply the knowledge gained from Chapters Three and Four. We know that a cleaved 5'exon would lack a polyA tail and hence be degraded by the nuclear exosome. However, a major problem in enriching noncoding RNA is the contamination of ribosomal RNA. Thus, in the context of *rrp6* deletion, depleting mRNA and ribosomal RNA from total RNA thereby enriching non-polyadenylated RNA within the cell, will be the first step in identification of cleaved RNAs. Furthermore, to determine the 3' end of cleaved form, we can add polyA tail *in vitro* using poly A polymerase and perform deep sequencing of the samples. Our goal would be to identify RNAs that end ahead of the 5'ss. Once identified, we can make mutations in the splice sites to test if the RNA is actually generated by spliceosomal cleavage. In summary, these results are expected to improve our understanding of

telomerase RNA biogenesis and possibly other non-coding RNA biogenesis in *S. pombe* and potentially humans.

References:

1. Nosek, J., P. Kosa, and L. Tomaska, *On the origin of telomeres: a glimpse at the pre-telomerase world.* Bioessays, 2006. **28**(2): p. 182-90.
2. Ishikawa, F. and T. Naito, *Why do we have linear chromosomes? A matter of Adam and Eve.* Mutat Res, 1999. **434**(2): p. 99-107.
3. Polo, S.E. and S.P. Jackson, *Dynamics of DNA damage response proteins at DNA breaks: a focus on protein modifications.* Genes & Development, 2011. **25**(5): p. 409-433.
4. Peterson, C.L. and J. Cote, *Cellular machineries for chromosomal DNA repair.* Genes & development, 2004. **18**(6): p. 602-16.
5. de Lange, T., *How Telomeres Solve the End-Protection Problem.* Science, 2009. **326**(5955): p. 948-952.
6. McClintock, B., *The Stability of Broken Ends of Chromosomes in Zea Mays.* Genetics, 1941. **26**(2): p. 234-82.
7. Muller, H.J. and E. Altenburg, *The Frequency of Translocations Produced by X-Rays in Drosophila.* Genetics, 1930. **15**(4): p. 283-311.
8. Orr-Weaver, T.L., J.W. Szostak, and R.J. Rothstein, *Yeast transformation: a model system for the study of recombination.* Proceedings of the National Academy of Sciences, 1981. **78**(10): p. 6354-6358.
9. Lieber, M.R., *The mechanism of double-strand DNA break repair by the nonhomologous DNA end-joining pathway.* Annu Rev Biochem, 2010. **79**: p. 181-211.
10. Helleday, T., et al., *DNA double-strand break repair: from mechanistic understanding to cancer treatment.* DNA Repair (Amst), 2007. **6**(7): p. 923-35.
11. Weinert, T. and L. Hartwell, *The RAD9 gene controls the cell cycle response to DNA damage in Saccharomyces cerevisiae.* Science, 1988. **241**(4863): p. 317-322.
12. Callegari, A.J. and T.J. Kelly, *Shedding light on the DNA damage checkpoint.* Cell Cycle, 2007. **6**(6): p. 660-6.
13. Szostak, J.W. and E.H. Blackburn, *Cloning yeast telomeres on linear plasmid vectors.* Cell, 1982. **29**(1): p. 245-55.
14. de Lange, T., et al., *Structure and variability of human chromosome ends.* Mol Cell Biol, 1990. **10**(2): p. 518-27.
15. Kipling, D. and H.J. Cooke, *Hypervariable ultra-long telomeres in mice.* Nature, 1990. **347**(6291): p. 400-2.
16. Griffith, J.D., et al., *Mammalian telomeres end in a large duplex loop.* Cell, 1999. **97**(4): p. 503-14.
17. Bilaud, T., et al., *Telomeric localization of TRF2, a novel human telobox protein.* Nat Genet, 1997. **17**(2): p. 236-9.
18. Broccoli, D., et al., *Human telomeres contain two distinct Myb-related proteins, TRF1 and TRF2.* Nat Genet, 1997. **17**(2): p. 231-5.
19. Kim, S.H., P. Kaminker, and J. Campisi, *TIN2, a new regulator of telomere length in human cells.* Nat Genet, 1999. **23**(4): p. 405-12.
20. Li, B., S. Oestreich, and T. de Lange, *Identification of human Rap1: implications for telomere evolution.* Cell, 2000. **101**(5): p. 471-83.
21. Baumann, P. and T.R. Cech, *Pot1, the putative telomere end-binding protein in fission yeast and humans.* Science, 2001. **292**(5519): p. 1171-5.

22. Hockemeyer, D., et al., *Recent expansion of the telomeric complex in rodents: Two distinct POT1 proteins protect mouse telomeres.* Cell, 2006. **126**(1): p. 63-77.
23. Kibe, T., et al., *Telomere protection by TPP1 is mediated by POT1a and POT1b.* Mol Cell Biol, 2010. **30**(4): p. 1059-66.
24. Wang, F., et al., *The POT1-TPP1 telomere complex is a telomerase processivity factor.* Nature, 2007. **445**(7127): p. 506-10.
25. Ye, J.Z., et al., *POT1-interacting protein PIP1: a telomere length regulator that recruits POT1 to the TIN2/TRF1 complex.* Genes Dev, 2004. **18**(14): p. 1649-54.
26. de Lange, T., *Shelterin: the protein complex that shapes and safeguards human telomeres.* Genes & Development, 2005. **19**(18): p. 2100-2110.
27. Karlseder, J., et al., *p53- and ATM-Dependent Apoptosis Induced by Telomeres Lacking TRF2.* Science, 1999. **283**(5406): p. 1321-1325.
28. d'Adda di Fagagna, F., et al., *A DNA damage checkpoint response in telomere-initiated senescence.* Nature, 2003. **426**(6963): p. 194-8.
29. Takai, H., A. Smogorzewska, and T. de Lange, *DNA damage foci at dysfunctional telomeres.* Curr Biol, 2003. **13**(17): p. 1549-56.
30. Denchi, E.L. and T. de Lange, *Protection of telomeres through independent control of ATM and ATR by TRF2 and POT1.* Nature, 2007. **448**(7157): p. 1068-71.
31. Guo, X., et al., *Dysfunctional telomeres activate an ATM-ATR-dependent DNA damage response to suppress tumorigenesis.* EMBO J, 2007. **26**(22): p. 4709-19.
32. Sfeir, A., et al., *Loss of Rap1 Induces Telomere Recombination in the Absence of NHEJ or a DNA Damage Signal.* Science, 2010. **327**(5973): p. 1657-1661.
33. Bae, N.S. and P. Baumann, *A RAP1/TRF2 complex inhibits nonhomologous end-joining at human telomeric DNA ends.* Mol Cell, 2007. **26**(3): p. 323-34.
34. Sarthy, J., et al., *Human RAP1 inhibits non-homologous end joining at telomeres.* EMBO J, 2009. **28**(21): p. 3390-9.
35. Sfeir, A., et al., *Mammalian telomeres resemble fragile sites and require TRF1 for efficient replication.* Cell, 2009. **138**(1): p. 90-103.
36. Zhong, F.L., et al., *TPP1 OB-fold domain controls telomere maintenance by recruiting telomerase to chromosome ends.* Cell, 2012. **150**(3): p. 481-94.
37. Nandakumar, J., et al., *The TEL patch of telomere protein TPP1 mediates telomerase recruitment and processivity.* Nature, 2012. **492**(7428): p. 285-9.
38. Watson, J.D., *Origin of concatemeric T7 DNA.* Nat New Biol, 1972. **239**(94): p. 197-201.
39. Olovnikov, A.M., *[Principle of marginotomy in template synthesis of polynucleotides].* Dokl Akad Nauk SSSR, 1971. **201**(6): p. 1496-9.
40. 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, 1973. **41**(1): p. 181-90.
41. Vega, L.R., M.K. Mateyak, and V.A. Zakian, *Getting to the end: telomerase access in yeast and humans.* Nat Rev Mol Cell Biol, 2003. **4**(12): p. 948-59.
42. Greider, C.W. and E.H. Blackburn, *Identification of a specific telomere terminal transferase activity in Tetrahymena extracts.* Cell, 1985. **43**(2 Pt 1): p. 405-13.
43. Lingner, J. and T.R. Cech, *Purification of telomerase from Euplotes aediculatus: requirement of a primer 3' overhang.* Proceedings of the National Academy of Sciences, 1996. **93**(20): p. 10712-10717.

44. Lingner, J., et al., *Reverse transcriptase motifs in the catalytic subunit of telomerase.* Science, 1997. **276**(5312): p. 561-7.
45. Hayflick, L. and P.S. Moorhead, *The serial cultivation of human diploid cell strains.* Exp Cell Res, 1961. **25**: p. 585-621.
46. Harley, C.B., A.B. Futcher, and C.W. Greider, *Telomeres shorten during ageing of human fibroblasts.* Nature, 1990. **345**(6274): p. 458-60.
47. Shay, J.W., O.M. Pereira-Smith, and W.E. Wright, *A role for both RB and p53 in the regulation of human cellular senescence.* Exp Cell Res, 1991. **196**(1): p. 33-9.
48. Hara, E., et al., *Cooperative effect of antisense-Rb and antisense-p53 oligomers on the extension of life span in human diploid fibroblasts, TIG-1.* Biochem Biophys Res Commun, 1991. **179**(1): p. 528-34.
49. Bodnar, A.G., et al., *Extension of life-span by introduction of telomerase into normal human cells.* Science, 1998. **279**(5349): p. 349-52.
50. Horn, S., et al., *TERT promoter mutations in familial and sporadic melanoma.* Science, 2013. **339**(6122): p. 959-61.
51. Huang, F.W., et al., *Highly recurrent TERT promoter mutations in human melanoma.* Science, 2013. **339**(6122): p. 957-9.
52. Ouellette, M.M., W.E. Wright, and J.W. Shay, *Targeting telomerase-expressing cancer cells.* J Cell Mol Med, 2011. **15**(7): p. 1433-42.
53. Sahin, E. and R.A. Depinho, *Linking functional decline of telomeres, mitochondria and stem cells during ageing.* Nature, 2010. **464**(7288): p. 520-8.
54. Calado, R.T. and N.S. Young, *Telomere diseases.* N Engl J Med, 2009. **361**(24): p. 2353-65.
55. Mitchell, J.R., E. Wood, and K. Collins, *A telomerase component is defective in the human disease dyskeratosis congenita.* Nature, 1999. **402**(6761): p. 551-5.
56. Vulliamy, T., et al., *The RNA component of telomerase is mutated in autosomal dominant dyskeratosis congenita.* Nature, 2001. **413**(6854): p. 432-5.
57. Yamaguchi, H., et al., *Mutations in TERT, the gene for telomerase reverse transcriptase, in aplastic anemia.* N Engl J Med, 2005. **352**(14): p. 1413-24.
58. Heiss, N.S., et al., *X-linked dyskeratosis congenita is caused by mutations in a highly conserved gene with putative nucleolar functions.* Nat Genet, 1998. **19**(1): p. 32-8.
59. Zhong, F., et al., *Disruption of telomerase trafficking by TCAB1 mutation causes dyskeratosis congenita.* Genes & development, 2011. **25**(1): p. 11-6.
60. Kiss, T., *Biogenesis of small nuclear RNPs.* J Cell Sci, 2004. **117**(Pt 25): p. 5949-51.
61. Batista, L.F., et al., *Telomere shortening and loss of self-renewal in dyskeratosis congenita induced pluripotent stem cells.* Nature, 2011. **474**(7351): p. 399-402.
62. Egan, E.D. and K. Collins, *Biogenesis of telomerase ribonucleoproteins.* RNA, 2012. **18**(10): p. 1747-59.
63. Weinrich, S.L., et al., *Reconstitution of human telomerase with the template RNA component hTR and the catalytic protein subunit hTRT.* Nat Genet, 1997. **17**(4): p. 498-502.
64. Aigner, S., et al., *Euplotes telomerase contains an La motif protein produced by apparent translational frameshifting.* EMBO J, 2000. **19**(22): p. 6230-9.
65. Mollenbeck, M., et al., *The telomerase-associated protein p43 is involved in anchoring telomerase in the nucleus.* J Cell Sci, 2003. **116**(Pt 9): p. 1757-61.

66. Aigner, S. and T.R. Cech, *The Euplotes telomerase subunit p43 stimulates enzymatic activity and processivity in vitro.* RNA, 2004. **10**(7): p. 1108-18.
67. O'Connor, C.M. and K. Collins, *A novel RNA binding domain in tetrahymena telomerase p65 initiates hierarchical assembly of telomerase holoenzyme.* Mol Cell Biol, 2006. **26**(6): p. 2029-36.
68. Greider, C.W. and E.H. Blackburn, *A telomeric sequence in the RNA of Tetrahymena telomerase required for telomere repeat synthesis.* Nature, 1989. **337**(6205): p. 331-7.
69. Singh, M., et al., *Structural basis for telomerase RNA recognition and RNP assembly by the holoenzyme La family protein p65.* Mol Cell, 2012. **47**(1): p. 16-26.
70. Chapon, C., T.R. Cech, and A.J. Zaug, *Polyadenylation of telomerase RNA in budding yeast.* RNA, 1997. **3**(11): p. 1337-51.
71. Noel, J.F., et al., *Budding yeast telomerase RNA transcription termination is dictated by the Nrd1/Nab3 non-coding RNA termination pathway.* Nucleic Acids Res, 2012. **40**(12): p. 5625-36.
72. Hsin, J.P. and J.L. Manley, *The RNA polymerase II CTD coordinates transcription and RNA processing.* Genes Dev, 2012. **26**(19): p. 2119-37.
73. Seto, A.G., et al., *Saccharomyces cerevisiae telomerase is an Sm small nuclear ribonucleoprotein particle.* Nature, 1999. **401**(6749): p. 177-80.
74. Mouaikel, J., et al., *Hypermethylation of the cap structure of both yeast snRNAs and snoRNAs requires a conserved methyltransferase that is localized to the nucleolus.* Mol Cell, 2002. **9**(4): p. 891-901.
75. Gallardo, F., et al., *Live cell imaging of telomerase RNA dynamics reveals cell cycle-dependent clustering of telomerase at elongating telomeres.* Molecular cell, 2011. **44**(5): p. 819-27.
76. Kiss, T., E. Fayet-Lebaron, and B.E. Jady, *Box H/ACA small ribonucleoproteins.* Mol Cell, 2010. **37**(5): p. 597-606.
77. Fu, D. and K. Collins, *Distinct biogenesis pathways for human telomerase RNA and H/ACA small nucleolar RNAs.* Molecular cell, 2003. **11**(5): p. 1361-72.
78. Venteicher, A.S., et al., *A human telomerase holoenzyme protein required for Cajal body localization and telomere synthesis.* Science, 2009. **323**(5914): p. 644-8.
79. Sharp, P.A., *The centrality of RNA.* Cell, 2009. **136**(4): p. 577-80.
80. Darnell, J.E., Jr., *Reflections on the history of pre-mRNA processing and highlights of current knowledge: a unified picture.* RNA, 2013. **19**(4): p. 443-60.
81. Chow, L.T., et al., *An amazing sequence arrangement at the 5' ends of adenovirus 2 messenger RNA.* Cell, 1977. **12**(1): p. 1-8.
82. Berget, S.M., C. Moore, and P.A. Sharp, *Spliced segments at the 5' terminus of adenovirus 2 late mRNA.* Proceedings of the National Academy of Sciences, 1977. **74**(8): p. 3171-3175.
83. Keren, H., G. Lev-Maor, and G. Ast, *Alternative splicing and evolution: diversification, exon definition and function.* Nat Rev Genet, 2010. **11**(5): p. 345-55.
84. Ward, A.J. and T.A. Cooper, *The pathobiology of splicing.* J Pathol, 2010. **220**(2): p. 152-63.
85. Patel, A.A. and J.A. Steitz, *Splicing double: insights from the second spliceosome.* Nat Rev Mol Cell Biol, 2003. **4**(12): p. 960-70.

86. Staley, J.P. and C. Guthrie, *Mechanical devices of the spliceosome: motors, clocks, springs, and things.* Cell, 1998. **92**(3): p. 315-26.
87. Jacquier, A., *Self-splicing group II and nuclear pre-mRNA introns: how similar are they?* Trends Biochem Sci, 1990. **15**(9): p. 351-4.
88. Kuhn, A.N. and N.F. Kaufer, *Pre-mRNA splicing in Schizosaccharomyces pombe: regulatory role of a kinase conserved from fission yeast to mammals.* Curr Genet, 2003. **42**(5): p. 241-51.
89. Zhang, M.Q. and T.G. Marr, *Fission yeast gene structure and recognition.* Nucleic Acids Res, 1994. **22**(9): p. 1750-9.
90. Will, C.L. and R. LVºhrmann, *Spliceosome Structure and Function.* Cold Spring Harbor Perspectives in Biology, 2011. **3**(7).
91. Romfo, C.M. and J.A. Wise, *Both the polypyrimidine tract and the 3' splice site function prior to the first step of splicing in fission yeast.* Nucleic Acids Res, 1997. **25**(22): p. 4658-65.
92. Reed, R., *The organization of 3' splice-site sequences in mammalian introns.* Genes Dev, 1989. **3**(12B): p. 2113-23.
93. Reich, C.I., et al., *Mutations at the 3' splice site can be suppressed by compensatory base changes in U1 snRNA in fission yeast.* Cell, 1992. **69**(7): p. 1159-69.
94. Horowitz, D.S., *The mechanism of the second step of pre-mRNA splicing.* Wiley Interdiscip Rev RNA, 2012. **3**(3): p. 331-50.
95. Will, C.L. and R. Luhrmann, *Spliceosome structure and function.* Cold Spring Harb Perspect Biol, 2011. **3**(7).
96. Yean, S.L. and R.J. Lin, *U4 small nuclear RNA dissociates from a yeast spliceosome and does not participate in the subsequent splicing reaction.* Mol Cell Biol, 1991. **11**(11): p. 5571-7.
97. Crispino, J.D., B.J. Blencowe, and P.A. Sharp, *Complementation by SR proteins of pre-mRNA splicing reactions depleted of U1 snRNP.* Science, 1994. **265**(5180): p. 1866-9.
98. Tarn, W.Y. and J.A. Steitz, *SR proteins can compensate for the loss of U1 snRNP functions in vitro.* Genes Dev, 1994. **8**(22): p. 2704-17.
99. Konarska, M.M., J. Vilardell, and C.C. Query, *Repositioning of the reaction intermediate within the catalytic center of the spliceosome.* Mol Cell, 2006. **21**(4): p. 543-53.
100. Smith, D.J., C.C. Query, and M.M. Konarska, *trans-Splicing to Spliceosomal U2 snRNA Suggests Disruption of Branch Site-U2 Pairing during Pre-mRNA Splicing.* Molecular cell, 2007. **26**(6): p. 883-890.
101. Crotti, L.B., D. Bacikova, and D.S. Horowitz, *The Prp18 protein stabilizes the interaction of both exons with the U5 snRNA during the second step of pre-mRNA splicing.* Genes Dev, 2007. **21**(10): p. 1204-16.
102. Semlow, D.R. and J.P. Staley, *Staying on message: ensuring fidelity in pre-mRNA splicing.* Trends Biochem Sci, 2012. **37**(7): p. 263-73.
103. Pickrell, J.K., et al., *Understanding mechanisms underlying human gene expression variation with RNA sequencing.* Nature, 2010. **464**(7289): p. 768-72.
104. Mitrovich, Q.M., et al., *Evolution of yeast noncoding RNAs reveals an alternative mechanism for widespread intron loss.* Science, 2010. **330**(6005): p. 838-41.
105. Hopfield, J.J., *Kinetic proofreading: a new mechanism for reducing errors in biosynthetic processes requiring high specificity.* Proc Natl Acad Sci U S A, 1974. **71**(10): p. 4135-9.

106. Ninio, J., *Kinetic amplification of enzyme discrimination.* Biochimie, 1975. **57**(5): p. 587-95.
107. Zaher, H.S. and R. Green, *Fidelity at the molecular level: lessons from protein synthesis.* Cell, 2009. **136**(4): p. 746-62.
108. Libby, R.T. and J.A. Gallant, *The role of RNA polymerase in transcriptional fidelity.* Mol Microbiol, 1991. **5**(5): p. 999-1004.
109. Burgess, S.M. and C. Guthrie, *A mechanism to enhance mRNA splicing fidelity: the RNA-dependent ATPase Prp16 governs usage of a discard pathway for aberrant lariat intermediates.* Cell, 1993. **73**(7): p. 1377-91.
110. Xu, Y.Z. and C.C. Query, *Competition between the ATPase Prp5 and branch region-U2 snRNA pairing modulates the fidelity of spliceosome assembly.* Mol Cell, 2007. **28**(5): p. 838-49.
111. Query, C.C. and M.M. Konarska, *Splicing fidelity revisited.* Nat Struct Mol Biol, 2006. **13**(6): p. 472-4.
112. Mayas, R.M., et al., *Spliceosome discards intermediates via the DEAH box ATPase Prp43p.* Proc Natl Acad Sci U S A, 2010. **107**(22): p. 10020-5.
113. Mayas, R.M., H. Maita, and J.P. Staley, *Exon ligation is proofread by the DExD/H-box ATPase Prp22p.* Nat Struct Mol Biol, 2006. **13**(6): p. 482-90.
114. Wahl, M.C., C.L. Will, and R. Luhrmann, *The spliceosome: design principles of a dynamic RNP machine.* Cell, 2009. **136**(4): p. 701-18.
115. Leonardi, J., et al., *TER1, the RNA subunit of fission yeast telomerase.* Nat Struct Mol Biol, 2008. **15**(1): p. 26-33.
116. Box, J.A., et al., *Spliceosomal cleavage generates the 3' end of telomerase RNA.* Nature, 2008. **456**(7224): p. 910-4.
117. Fu, D. and K. Collins, *Distinct biogenesis pathways for human telomerase RNA and H/ACA small nucleolar RNAs.* Mol Cell, 2003. **11**(5): p. 1361-72.
118. Tang, W., et al., *Telomerase RNA biogenesis involves sequential binding by Sm and Lsm complexes.* Nature, 2012. **484**(7393): p. 260-4.
119. Brun, C., D.D. Dubey, and J.A. Huberman, *pDblet, a stable autonomously replicating shuttle vector for Schizosaccharomyces pombe.* Gene, 1995. **164**(1): p. 173-7.
120. Baumann, P. and T.R. Cech, *Protection of telomeres by the Ku protein in fission yeast.* Mol Biol Cell, 2000. **11**(10): p. 3265-75.
121. Burke, J.D. and K.L. Gould, *Molecular cloning and characterization of the Schizosaccharomyces pombe his3 gene for use as a selectable marker.* Molecular & general genetics : MGG, 1994. **242**(2): p. 169-76.
122. Schneider, T.D., *A brief review of molecular information theory.* Nano Commun Netw., 2010. **1**(3): p. 173-180.
123. Blackburn, E.H. and K. Collins, *Telomerase: an RNP enzyme synthesizes DNA.* Cold Spring Harbor perspectives in biology, 2011. **3**(5).
124. Webb, C.J. and V.A. Zakian, *Identification and characterization of the Schizosaccharomyces pombe TER1 telomerase RNA.* Nat Struct Mol Biol, 2008. **15**(1): p. 34-42.
125. Wilusz, C.J. and J. Wilusz, *Eukaryotic Lsm proteins: lessons from bacteria.* Nat Struct Mol Biol, 2005. **12**(12): p. 1031-6.

126. Hilleren, P.J. and R. Parker, *Cytoplasmic degradation of splice-defective pre-mRNAs and intermediates.* Molecular cell, 2003. **12**(6): p. 1453-65.
127. Gunisova, S., et al., *Identification and comparative analysis of telomerase RNAs from Candida species reveal conservation of functional elements.* RNA, 2009. **15**(4): p. 546-59.
128. Gooding, C., et al., *A class of human exons with predicted distant branch points revealed by analysis of AG dinucleotide exclusion zones.* Genome biology, 2006. **7**(1): p. R1.
129. Harigaya, Y. and R. Parker, *Global analysis of mRNA decay intermediates in Saccharomyces cerevisiae.* Proceedings of the National Academy of Sciences of the United States of America, 2012. **109**(29): p. 11764-9.
130. Baillat, D., et al., *Integrator, a multiprotein mediator of small nuclear RNA processing, associates with the C-terminal repeat of RNA polymerase II.* Cell, 2005. **123**(2): p. 265-76.
131. Hernandez, N. and A.M. Weiner, *Formation of the 3' end of U1 snRNA requires compatible snRNA promoter elements.* Cell, 1986. **47**(2): p. 249-58.
132. Moldon, A., et al., *Promoter-driven splicing regulation in fission yeast.* Nature, 2008. **455**(7215): p. 997-1000.
133. Bentley, D.L., *Rules of engagement: co-transcriptional recruitment of pre-mRNA processing factors.* Current opinion in cell biology, 2005. **17**(3): p. 251-6.
134. Huang, T., J. Vilardell, and C.C. Query, *Pre-spliceosome formation in S.pombe requires a stable complex of SF1-U2AF(59)-U2AF(23).* Embo J, 2002. **21**(20): p. 5516-26.
135. Coolidge, C.J., R.J. Seely, and J.G. Patton, *Functional analysis of the polypyrimidine tract in pre-mRNA splicing.* Nucleic Acids Res., 1997. **25**(4): p. 888-896.
136. Fu, X.D., et al., *The role of branchpoint and 3'-exon sequences in the control of balanced splicing of avian retrovirus RNA.* Genes & Dev., 1991. **5**(2): p. 211-220.
137. Umen, J.G. and C. Guthrie, *The second catalytic step of pre-mRNA splicing.* RNA 1995. **1**(9): p. 869-885.
138. Potashkin, J., K. Naik, and K. Wentz-Hunter, *U2AF homolog required for splicing in vivo.* Science, 1993. **262**(5133): p. 573-575.
139. Zamore, P.D. and M.R. Green, *Identification, purification, and biochemical characterization of U2 small nuclear ribonucleoprotein auxiliary factor.* Proc. Natl. Acad. Sci. USA, 1989. **86**(23): p. 9243-9247.
140. Zamore, P.D. and M.R. Green, *Biochemical characterization of U2 snRNP auxiliary factor: an essential pre-mRNA splicing factor with a novel intranuclear distribution.* EMBO (European Molecular Biology Organization) Journal, 1991. **10**(1): p. 207-214.
141. Singh, R., J. Valcarcel, and M.R. Green, *Distinct binding specificities and functions of higher eukaryotic polypyrimidine tract-binding proteins.* Science, 1995. **268**(5214): p. 1173-1176.
142. Haraguchi, N., et al., *Mutations in the SF1-U2AF59-U2AF23 complex cause exon skipping in Schizosaccharomyces pombe.* J Biol Chem, 2007. **282**(4): p. 2221-8.
143. Meyer, M., et al., *Deciphering 3'ss selection in the yeast genome reveals an RNA thermosensor that mediates alternative splicing.* Molecular cell, 2011. **43**(6): p. 1033-9.
144. Wassarman, D.A. and J.A. Steitz, *Interactions of small nuclear RNA's with precursor messenger RNA during in vitro splicing.* Science, 1992. **257**(5078): p. 1918-1925.
145. Xu, Y.Z. and C.C. Query, *Competition between the ATPase Prp5 and branch region-U2 snRNA pairing modulates the fidelity of spliceosome assembly.* Mol Cell., 2007. **28**(5): p. 838-849.

146. Spingola, M., et al., *Genome-wide bioinformatic adn molecular analysis of introns in Saccharomyces cerevisiae.* RNA, 1999. **5**(2): p. 212-234.
147. Cellini, A., E. Felder, and J.J. Rossi, *Yeast pre-messenger RNA splicing efficiency depends on critical spacing requirements between the branch point and 3'splice site.* EMBO (European Molecular Biology Organization) Journal, 1986. **5**(5): p. 1023-1030.
148. Luukkonen, B.G. and B. Seraphin, *The role of branchpoint-3' splice site spacing and interaction between intron terminal nucleotides in 3' splice site selection in Saccharomyces cerevisiae.* Embo J, 1997. **16**(4): p. 779-92.
149. Corvelo, A., et al., *Genome-wide association between branch point properties and alternative splicing.* PLoS Computational Biology, 2010. **6**(11): p. e1001016.
150. Gao, K., et al., *Human branch point consensus sequence is yUnAy.* Nucleic Acids and Res., 2008. **36**(7): p. 2257-2267.
151. Chua, K. and R. Reed, *An upstream AG determines whether a downstream AG is selected during catalytic step II of splicing.* Mol. Cell. Biol., 2001. **21**(5): p. 1509-1514.
152. Roscigno, R.F., M. Weiner, and M.A. Garcia-Blanco, *A mutational analysis of the polypyrimidine tract of introns. Effects of sequence differences in pyrimidine tracts on splicing.* J Biol Chem, 1993. **268**(15): p. 11222-9.
153. Patterson, B. and C. Guthrie, *A U-rich tract enhances usage of an alternative 3' splice site in yeast.* Cell 1991. **64**(1): p. 181-187.
154. Chiara, M.D., et al., *Evidence that U5 snRNP recognizes the 3' splice site for catalytic step II in mammals.* EMBO (European Molecular Biology Organization) Journal, 1997. **16**(15): p. 4746-4759.
155. Agafonov, D.E., et al., *Semiquantitative proteomic analysis of the human spliceosome via a novel two-dimensional gel electrophoresis method.* Mol Cell Biol, 2011. **31**(13): p. 2667-82.
156. Hoskins, A.A., et al., *Ordered and dynamic assembly of single spliceosomes.* Science, 2011. **331**(6022): p. 1289-95.
157. Zhuang, Y., A.M. Goldstein, and A.M. Weiner, *UACUAAC is the preferred branch site for mammalian mRNA splicing.* Proc. Natl. Acad. Sci. USA, 1989. **86**: p. 2752-2756.
158. Bessonov, S., et al., *Isolation of an active step I spliceosome and composition of its RNP core.* Nature, 2008. **452**(7189): p. 846-50.
159. Fouser, L.A. and J.D. Friesen, *Mutations in a yeast intron demonstrate the importance of specific conserved nucleotides for the two stages of nuclear mRNA splicing.* Cell, 1986. **45**(1): p. 81-93.
160. Horning, H., M. Aebi, and C. Weissmann, *Effect of mutations at the lariat branch acceptor site on b-globin pre-mRNA splicing in vitro.* Nature (London), 1986. **324**: p. 589-591.
161. Jacquier, A. and M. Rosbash, *RNA splicing and intron turnover are greatly diminished by a mutant yeast branch point.* Proc. Natl. Acad. Sci. USA, 1986. **83**(16): p. 5835-5839.
162. Tseng, C.K., H.L. Liu, and S.C. Cheng, *DEAH-box ATPase Prp16 has dual roles in remodeling of the spliceosome in catalytic steps.* RNA, 2011. **17**(1): p. 145-154.
163. Smith, D.J., M.M. Konarska, and C.C. Query, *Insights into branch nucleophile positioning and activation from an orthogonal pre-mRNA splicing system in yeast.* Mol Cell., 2009. **34**(3): p. 333-343.

164. Smith, D.J., C.C. Query, and M.M. Konarska, *trans-splicing to spliceosomal U2 snRNA suggests disruption of branch site-U2 pairing during pre-mRNA splicing.* Mol Cell, 2007. **26**(6): p. 883-90.
165. 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, 2013. **27**(6): p. 627-38.
166. Konarska, M.M., J. Vilardell, and C.C. Query, *Repositioning of the Reaction Intermediate within the Catalytic Center of the Spliceosome.* Molecular cell, 2006. **21**(4): p. 543-553.
167. Zhuang, Y.A., A.M. Goldstein, and A.M. Weiner, *UACUAAC is the preferred branch site for mammalian mRNA splicing.* Proc Natl Acad Sci U S A, 1989. **86**(8): p. 2752-6.
168. Wassarman, D.A. and J.A. Steitz, *Interactions of small nuclear RNA's with precursor messenger RNA during in vitro splicing.* Science, 1992. **257**(5078): p. 1918-25.
169. Fabrizio, P., et al., *The evolutionarily conserved core design of the catalytic activation step of the yeast spliceosome.* Mol Cell, 2009. **36**(4): p. 593-608.
170. Konarska, M.M. and C.C. Query, *Insights into the mechanisms of splicing: more lessons from the ribosome.* Genes & Development, 2005. **19**(19): p. 2255-2260.
171. Kuprys, P.V., et al., *Identification of Telomerase RNAs from Filamentous Fungi Reveals Conservation with Vertebrates and Yeasts.* PLoS One, 2013. **8**(3): p. e58661.
172. Gao, K., et al., *Human branch point consensus sequence is yUnAy.* Nucleic Acids Research, 2008. **36**(7): p. 2257-2267.
173. Brennwald, P., G. Porter, and J.A. Wise, *U2 small nuclear RNA is remarkably conserved between Schizosaccharomyces pombe and mammals.* Mol Cell Biol, 1988. **8**(12): p. 5575-80.
174. Blais, A. and B.D. Dynlacht, *Constructing transcriptional regulatory networks.* Genes & Development, 2005. **19**(13): p. 1499-1511.
175. Dumesic, P.A., et al., *Stalled spliceosomes are a signal for RNAi-mediated genome defense.* Cell, 2013. **152**(5): p. 957-68.
176. Kim Guisbert, K.S. and E.J. Sontheimer, *Quit stalling or you'll be silenced.* Cell, 2013. **152**(5): p. 938-9.
177. O'Brien, K., et al., *The biflavonoid isoginkgetin is a general inhibitor of Pre-mRNA splicing.* J Biol Chem, 2008. **283**(48): p. 33147-54.
178. Ohno, M. and Y. Shimura, *A human RNA helicase-like protein, HRH1, facilitates nuclear export of spliced mRNA by releasing the RNA from the spliceosome.* Genes Dev, 1996. **10**(8): p. 997-1007.
179. Gee, S., et al., *Cloning of mDEAH9, a putative RNA helicase and mammalian homologue of Saccharomyces cerevisiae splicing factor Prp43.* Proc Natl Acad Sci U S A, 1997. **94**(22): p. 11803-7.
180. Wilusz, J.E., et al., *A triple helix stabilizes the 3' ends of long noncoding RNAs that lack poly(A) tails.* Genes Dev, 2012. **26**(21): p. 2392-407.

Appendix I: Publications

1. Tang, W., **Kannan, R**., Blanchette, M., and Baumann, P. Telomerase RNA biogenesis involves sequential binding by Sm and Lsm complexes. *Nature, 484: 260-264, 2012.*

2. **Kannan, R.**, Hartnett, S., Voelker, R.B., Berglund, J.A., Staley, J.P., and Baumann, P. Intronic sequence elements impede exon ligation and trigger a discard pathway that yields functional telomerase RNA in fission yeast. *Genes & Development, 27(6): 627-38, 2013.*

3. **Kannan, R** and Baumann, P. U6 hyper-stabilization with 5'splice site generates the 3' end of telomerase RNA in *S. cryophilus* and *S. octosporus*. In Preparation.

Appendix II : Conferences Attended and Presentations

1. **Kannan, R.**, Hartnett, S., Voelker, R.B., Berglund, J.A., Staley, J.P., and Baumann, P. Intronic sequence elements impede exon ligation and trigger a discard pathway that yields functional telomerase RNA in fission yeast, HHMI meeting, Janelia Farm, September 2012. Poster presentation.

2. **Kannan R** and Baumann P. Telomerase RNA Biogenesis: Spliceosomal cleavage versus Splicing, Keystone Symposia, Protein-RNA Interactions in Biology and Disease, Santa Fe, New Mexico, March 2012. Oral presentation.

3. **Kannan R** and Baumann P. Telomerase RNA Biogenesis: Spliceosomal cleavage versus Splicing, RNA Conference, Kyoto, Japan, June 2011. Poster presentation.

4. **Kannan R** and Baumann P. Identifying the factors involved in spliceosomal cleavage of telomerase RNA in *S. pombe*, EMBO Conference, Telomeres and the DNA Damage Response, Marseille, France, Sep 2010. Poster presentation.

Appendix III: Collaborators' Contributions

Sean Hartnett contributed to the *prp22* and *prp43* mutant characterization and performed spotting assays in Figure 3.8; Roger and Andy performed the computational analysis shown in Figure 4.1 and Figure 4.2; Jon Staley motivated and advised on discard pathway mutants and mechanism of splicing.