January 2020

Termination-Independent Role Of Rat1 In Cotranscriptional Splicing In Budding Yeast

Zuzer Hakimuddin Dhoondia
Wayne State University

TERMINATION INDEPENDENT ROLE OF RAT1 IN COTRANSCRIPTIONAL SPLICING IN BUDDING YEAST

by

ZUZER HAKIMUDDIN DHOONDIA

DISSERTATION

Submitted to the Graduate School

of Wayne State University

Detroit, Michigan

in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

2020

MAJOR: BIOLOGICAL SCIENCES

Approved By:

_____

Advisor **Date**

_____

_____

_____

_____

DEDICATION

To my

Teachers, friends & family

ACKNOWLEDGEMENTS

I would like to thank my supervisor Dr. Athar Ansari for giving me an opportunity to work in his group. He gave me freedom to explore the field and was always present for discussion and troubleshooting. I thank him for all the positive support and guidance. His knowledge in the different field of science is motivational. His advice on both research and soft skills including writing and presentations have been priceless and will forever influence me in my further career. In the end, we managed to build up pieces of present work and connected with a past research in the field to write a really nice story.

I am very thankful to have worked with Dr. Roger Pique-Regi, who has been an amazing mentor in learning skills of Bioinformatics. He is brilliant with skills in R and I am trying to use the best of my time to interact and learn from him. He was very calm and understanding in explaining and always ready to help when needed. His guidance has been very important for my research and has giving me skill set that I will be definitely using it future. I thank Dr. Victoria Meller and Dr. Lori Pile for being an amazing committee member who always kept me and checked on for the progress through my project in lab meetings, committee meeting and molecular biology division meetings. I also thank all of them for the constructive comments on the work as well as suggesting direction when needed. Dr. Meller and Dr. Pile have always inspired me to see a topic under the bigger context in the cell/system, and through them I learned the importance of seeing any work with lens of constructive criticism.

I also thank lab co-workers Neha, Nadra, Scott, Bianca, Katie, Ricci, and Mike for a wonderful friendly lab atmosphere. They all were a second family to me who were very helpful, understanding and co-operative. I also thank Neha, Scott and Nadra for sharing

their expertise in several techniques at the beginning of my Ph.D program. Katie, Mike, Ricci, and Alisa have been very helpful in proofreading my writing and they have all contributed me to improve my writing skills. Thank you for my all-time friends Ojas Natarajan, Tejeshwar Rao, Mohammed-Hussain Bharmal, Feng Tao, Katie Dwyer, Mike O'Brien, Neha Agarwal, Progga Sen, and Mitra Asgari. I had a wonderful time with you all and enjoyed discussion over scientific and nonscientific topics. All of you have contributed in a meaningful way in my journey of Ph.D.

My work is incomplete without mentioning the contribution from the seven undergraduate who I had a pleasure to work with and learned important skills to manage a team. Special thanks to Zahid, Julie, Marva, Hesham and Zehn who have worked very hard towards the advancement of this project. Also, Thanks to Alisa, Nandita, Katsumi, and Akshay. I enjoyed working with you all and big thank you for a cheerful environment in the lab.

Lastly, my very special thanks to my parents, Hakimuddin, Munira, and Zainab. It is with your constant support and blessings I am fortunate to achieve a dream career with endless opportunity. A special mention of my brother, Dr. Saifuddin Dhoondia, who has been a biggest motivation in what I am today. Without his support I wouldn't be in the field science. My sweet sister, Sakina, has been a wonderful companion to cheerful conversation when research is on a down slide. I thank both my siblings for being part of my family and sharing so many wonderful things in this life. I love laughing with them. Samina & Zahra for that even being far away they were always close and supportive. I thank them all my previous teachers and family members for shaping me the person who

I am today. Many thanks to Tasneem, my wife and a special friend for all the great moments together. I have been very blessed in my life to have your love and friendship.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1. INTRODUCTION

According to central dogma of molecular biology proposed by Francis Crick, expression of genetic information takes place in two steps. The first step is transcription, which is the synthesis of mRNA from the DNA template, and the second step is translation that involves synthesis of protein from the mRNA (Crick, 1958). With the exception of a few viruses, the repository of genetic information in all life forms is DNA. Genetic information is stored in units called genes. During transcription, a multi-subunit molecular machine called DNA-dependent RNA polymerase (RNAP) makes mRNA based on the gene. In prokaryotes and archaea, there is a single RNA polymerase that transcribes different types of RNAs including mRNA (Zhang et al., 1999a). In eukaryotes, however, there are at least three distinct RNA polymerases (Weinmann and Roeder, 1974). RNAPI transcribes ribosomal RNA (rRNA), whereas RNAPIII transcribes 5S rRNA, transfer RNA (tRNA), and small nuclear U6 snRNA (Weinmann and Roeder, 1974). RNAPII transcribes messenger RNA (mRNA) as well as a number of different types of non-coding RNAs including small interfering RNA (siRNA), micro RNA (miRNA), small nuclear RNA (snRNA) and small nucleolar RNA (snoRNA) (Weinmann et al., 1974).

RNAPII-mediated transcription is a tightly regulated process that involves participation of a number of accessory protein factors like activators, repressors, chromatin modifiers, and the general transcription factors. A defect in any of these factors may adversely affect the process of transcription and can be detrimental to the cell. Since the level of proteins in a cell is determined to a great extent by transcription, understanding the molecular mechanism underlying transcription by RNAPII is critical for understanding gene expression. The molecular basis underlying transcription by RNAPII

is remarkably conserved from yeast to humans. Elucidation of transcription in genetically tractable budding yeast Saccharomyces cerevisiae therefore has contributed immensely to our understanding of the process of transcription by RNAPII.

1.1 RNAPII transcription cycle

The RNAPII plays a central role in transcription in eukaryotic cells. Transcription cycle of RNAPII consists of four consecutive steps; initiation, elongation, termination and reinitiation. Transcription is coupled to three RNA processing events; capping, splicing and cleavage-polyadenylation (Figure 1.1). The molecular structure of RNAPII allow steps of transcription and cotranscriptional RNA processing events to be accomplished in a highly coordinated manner. RNAPII is composed of 12 subunits (Bushnell and Kornberg, 2003). The catalytic core of RNAPII is composed of ten subunits, to which the heterodimer Rpb4/Rpb7 joins to form the 12-subunit complex. Importantly, the largest subunit of RNAPII, Rpb1, consists of a carboxy-terminal-domain (CTD) with a variable number of $Y_1S_2P_3T_4S_5P_6S_7$ heptapeptide repeats. The number of heptapeptide repeats varies from 26 in yeast to 52 in humans (Corden, 1990; Corden et al., 1985; Hsin and Manley, 2012). Posttranslational modification of the CTD by phosphorylation of tyrosine at position 1, threonine at position 4, and serine at positions 2, 5 and 7 of the heptapeptide repeat is regulated by various kinases and phosphatases (Bataille et al., 2012; Corden, 2013). Phosphorylation/dephosphorylation of the CTD facilitates coordinated recruitment of transcription factors, RNA processing factors and the histone modifying enzyme. Alteration in phosphorylation state of the CTD results in the recruitment of various accessory factors during different stages of transcription and cotranscriptional RNA processing. The proteins that recognize the specific phosphorylation state of CTD contain

a CTD interacting domain (CID) that helps in binding of these proteins to RNAPII. This phosphorylation pattern of the CTD is dynamic and is tightly regulated during different steps of transcription and co-transcriptional RNA processing events (Bataille et al., 2012; Hsin and Manley, 2012; Jeronimo et al., 2013).

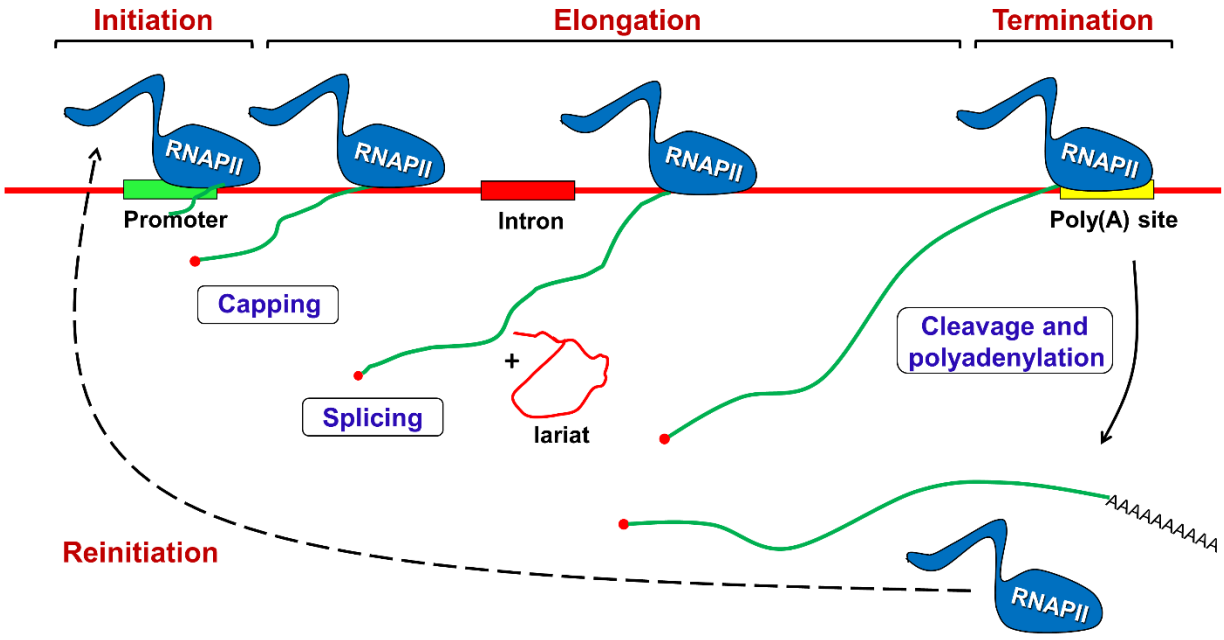

Figure1. 1 Schematic representation of transcription and transcription-coupled RNA processing. The four stages of the transcription cycle are labelled in red and cotransciptional RNA processing steps are labelled in purple.

1.2 Initiation of transcription

The promoter region positioned upstream of the gene is often nucleosome-free but may contain a complex of histones and non-histone proteins around which DNA is wrapped. Before transcription can be initiated, the promoter region must be depleted of nucleosomes to make it accessible for binding to activators, general transcription factors, and RNAPII. SWI/SNF and RSC, which are ATP-dependent chromatin modifying complexes, remodel or displace nucleosomes in and around the promoter making the

region accessible to transcription machinery (Buratowski and Kim, 2010; Tosi et al., 2013; Wilson et al., 1996).

In *S. cerevisiae*, the signal for the transcription of the gene is through the recruitment of the gene-specific activator protein at the Upstream Activating Sequence (UAS) (Hahn and Young, 2011; Zawel and Reinberg, 1992). UAS elements are located within a few hundred base pairs upstream of the transcription start site (TSS) of the gene. The binding of the activator protein often triggers the recruitment of the co-activator complex called Mediator, which in turn facilitates recruitment of the general transcription factors (Kubik et al., 2017; Warfield et al., 2017). Mediator is a 25-subunit complex that acts as a bridge linking the activator with the general transcription machinery (Flanagan et al., 1991). Although Mediator was discovered as a co-activator, its role in the transcription cycle was found to be more comprehensive. Evidence suggests that Mediator could be a general transcription factor. A recently emerging view is the role of Mediator in transcription is not restricted to the initiation step but may extend to the elongation and termination steps as well (Donner et al., 2010; Mukundan and Ansari, 2011; Takahashi et al., 2011).

Mediator binding is followed by the recruitment of six general transcription factors; TFIID or SAGA, TFIIB, TFIIA, TFIIF, TFIIE and TFIIH. These six general transcription factors (GTFs) are recruited in a sequential manner and together constitute the preinitiation complex (PIC). The first general transcription factor to bind to the promoter is TFIID or SAGA complex. TFIID is a multi-subunit factor consisting of TBP (TATA-binding protein) and 14 TBP-associated factors (TAFs) (Green, 2005). SAGA (Spt-Ada-Gcn5 acetyltransferase) is also a megacomplex composed of 19 subunits. Previous

studies that investigated SAGA and TFIID recruitment had classified promoters based on the presence or absence of a TATA-box. Specifically, TFIID was found recruited to TATA-less promoters, which constitutes a major class of promoters present in yeast (Li et al., 2002). Only 10% of the genes in yeast contain a TATA-box, whereas a majority of promoter's lack a canonical TATA-box. The TATA-containing promoters were believed to recruit the SAGA complex (Huisinga and Pugh, 2004; Li et al., 2002). Contrary to these findings, it was recently shown that TFIID and SAGA are required for the transcription of all protein-coding genes in yeast irrespective of the presence or absence of the TATA-box (Baptista et al., 2017; Warfield et al., 2017).

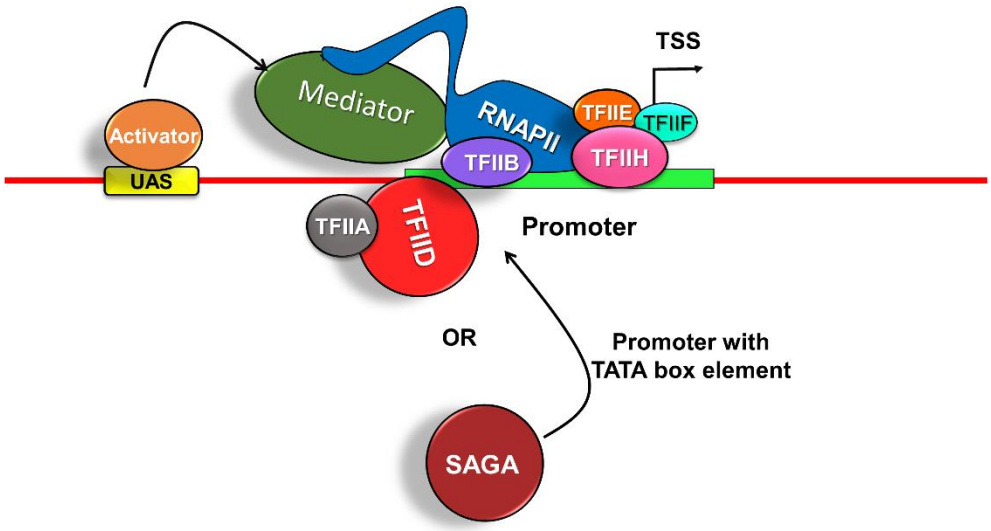

Figure1. 2 The pre-initiation complex of RNAPII assembled on the promoter.

The next GTF that binds to the promoter region through its interaction with a DNA element called TFIIB recognition element (BRE), is TFIIB. It is a single subunit factor which interacts with the TATA binding protein (TBP), and RNAPII (Deng and Roberts, 2005; Lagrange et al., 1998). This association plays a vital role in the transcription start site (TSS) selection (Deng and Roberts, 2007; Liu et al., 2010; Sun and Hampsey, 1996).

Following this, TFIIA is recruited, which helps to stabilize the association of TFIIB with DNA (Meisterernst and Roeder, 1991). This signals TFIIF and RNAPII recruitment at the promoter and finally, TFIIE and TFIIH join at the promoter to form the pre-initiation complex (PIC). Thus, the PIC contains an activator, a co-activator, and the general transcription factors TFIID, TFIIB, TFIIA, TFIIF, TFIIH, and TFIIE (Figure 1.2). *In S. cerevisiae*, RNAPII scans for the initiator element or transcription start site (TSS) situated 40-120 bp downstream of the TATA box, while in higher eukaryotes the TSS is located approximately 30 bp downstream of the TATA element (Giardina and Lis, 1993; Murakami et al., 2015).

Once the PIC is assembled, the two strands of DNA separate to form a 'transcription bubble' that helps to change the conformation of the PIC from a closed to an open complex structure (Giardina and Lis, 1993; Grünberg and Hahn, 2013). The initiation of transcription and synthesis of the first 15-20 nucleotides of nascent mRNA takes place while RNAPII is still sitting on the promoter and is in complex with the GTFs. The recruitment of the last general transcription factor, TFIIH, to the pre-initiation complex (PIC), has two important functions. First, the Ssl2 subunit of TFIIH has helicase activity, which acts as a wrench and unwinds the DNA downstream of the TSS to facilitate collapse of the transcription bubble. Second, the Kin28 subunit of TFIIH kinase module phosphorylates the Ser5 residue of Rpb1-CTD. This phosphorylation, together with the Ssl2-mediated collapse of the transcription bubble, marks the 'promoter clearance' step during which RNAPII dissociates from the promoter-associated initiation complex and transitions to elongation step (Kim et al., 2009). Several rounds of abortive initiation by RNAPII result in transcripts of around 10 nucleotides length before the promoter

clearance (Murakami et al., 2015). Leaving a part of the assembled initiation complex behind at the promoter, RNAPII begins to elongate the transcript. TFIID, TFIIA, TFIIE, TFIIH and SAGA are part of the scaffold left at the promoter, which can be utilized for the re-initiation of transcription (Murakami et al., 2013).

1.2.1 Capping

The Rbp1-CTD phosphorylation at the serine-5 residue is not merely the signal for promoter clearance, but also facilitates the recruitment of capping enzymes. The addition of the 7-methylguanosyl cap to the 5' end of the nascent transcript stabilizes the nascent RNA and prevents its degradation by the nuclear RNases. The process to add the 7-methylguanosyl cap is the first cotranscriptional RNA processing event that occurs in three steps (Ghosh and Lima, 2010). In the first step, Cet1 phosphatase removes the γ-phosphate group from the exposed 5' triphosphate group. In the next step, an addition of an inverted guanylyl group by Ceg1 guanylyltrasferase takes place. In the final step, methylation of the terminal guanine residue occurs at the N7 position by Abd1 methyl transferase (Cho et al., 1997; Schroeder et al., 2000) (Figure 1.3). The cap-binding complex (CBC) associates with the 7-methyl-guanosine (m7G) cap and helps in subsequent pre-mRNA splicing and export of mRNA events (Schwer and Shuman, 1996). At this step, incorrectly capped transcripts are subjected to cotranscriptional degradation by Rat1-Rai1 complex (Xiang et al., 2009).

1.3 Elongation of transcription by RNAPII on DNA full of roadblocks

RNA emerging from the RNAPII exit channel is a signal for the recruitment of elongation factors Spt4-Spt5 and is one of the earliest events that marks the transition of transcription from initiation to elongation. The mechanism by which the Spt4-Spt5

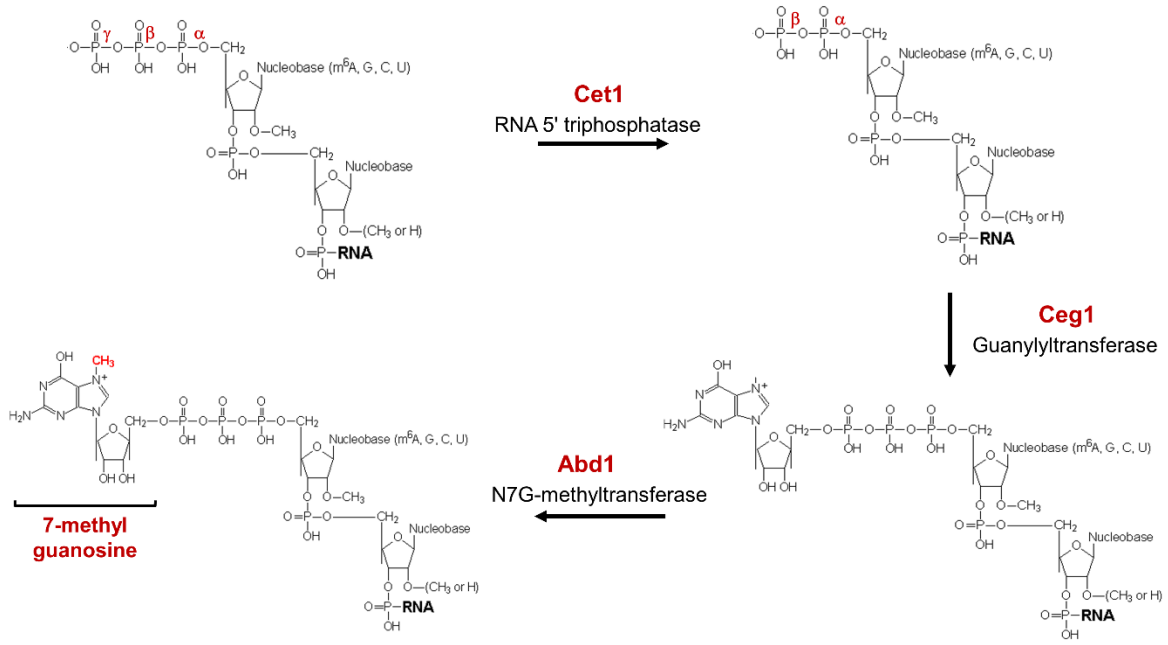

Figure1. 3 Biochemical reactions in the enzymatic 5′-cap formation of mRNA.

heterodimer regulates elongation is not entirely clear. Structural studies suggest Spt4-Spt5 may bind directly to RNAPII, bridging its central cleft. This arrangement may prevent dissociation of the DNA template from the transcribing polymerase and enhance the processivity of the elongating RNAPII (Martinez-Rucobo et al., 2011). Spt5 is unique as it contains a short repeat sequence at its C-terminus known as the C-terminal region (CTR) (Swanson et al., 1991). The CTR of Spt5 is phosphorylated by the cyclin dependent kinase Ctk1/Bur1 and its cyclin pair Ctk2/Bur2 (Liu et al., 2009). Likewise, the Ctk1/Bur1 kinase also phosphorylates Rpb1-CTD at the tyrosine-1 and serine-2 residues (Mayer et al., 2012; Wood and Shilatifard, 2006). Both phosphorylation events signal the recruitment of accessory factors that help to improve RNAPII processivity along the DNA template, and to overcome the nucleosomal barrier for completing the productive round of transcription. The elongation factors that help passage of the polymerase through chromatin at this stage include the histone modifying enzyme (Set1), histone chaperones

(Spt6 and Spt16) and the chromatin remodeling complexes (RSC, Chd1 and ISWI). The CTR of Spt5 recruits the Pol II-associated factor 1 (PAF1) complex (Liu et al., 2009; Zhou et al., 2009) which, along with Spt4-Spt5, recruits histone H3K4 methyltransferase Set1, a subunit of COMPASS (Complex of Proteins Associated with Set1) that ensures H3K4 dimethylation as well as trimethylation of chromatin templates (Miller et al., 2001). Both FACT (Facilitates Chromatin Transcription) and Spt6 are histone chaperones that facilitate assembly and disassembly of nucleosomes, thus allowing the polymerase to move through the chromatin template smoothly (Figure 1.4) (Belotserkovskaya et al., 2003; Kim et al., 2004a).

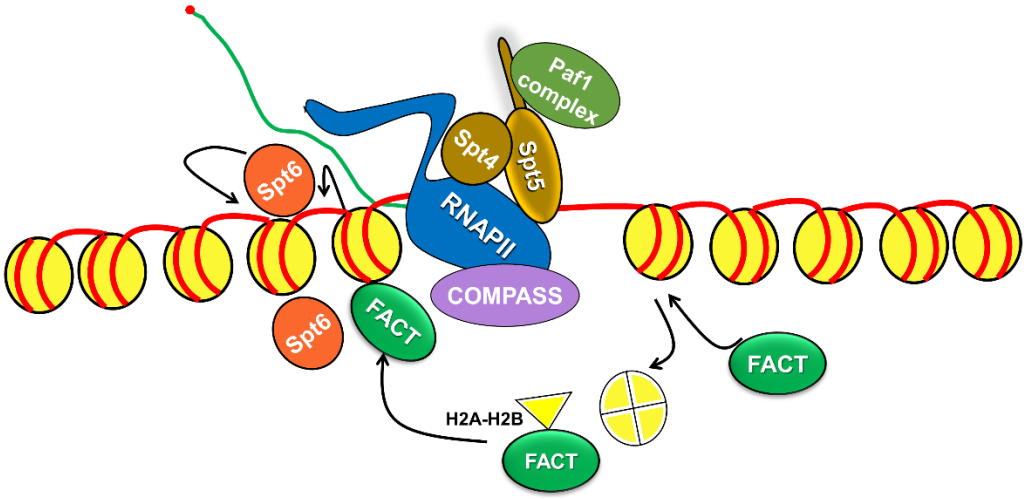

Figure1. 4 Schematic representation of elongation complex associated with RNAPII

The presence of A/T-rich sequences in the DNA template serves as an additional barrier for the elongating RNAPII. The polymerase, while maneuvering through A/T-rich sequences, often backtracks leading to the misalignment of the active site of the enzyme with the 3' end of elongating mRNA. TFIIS is the elongation factor that helps the backtracked RNAPII to resume elongation. TFIIS stimulates 3'→ 5' exoribonuclease

activity of polymerase, thereby realigning the RNAPII active site with the 3' end of RNA (Hahn, 2004; Kulish and Struhl, 2001).

1.3.1 Splicing removes non-coding regions from mRNA

S. cerevisiae differs from most other eukaryotic organisms as its genome consists of fewer intron containing genes (Neuvéglise et al., 2011). Nearly 95% of the genes in budding yeast are without an intron. Introns represent non-coding sequences (Spingola et al., 1999). Most of the intron containing genes in yeast contain only a single intron, which is typically located near the 5' end of the gene (Woolford and Peebles, 1992). Intron containing genes represent a mere 3.8% of the total genes in yeast, yet they contribute to nearly one third of the total mRNA produced in yeast cells. Out of ~300 intron containing genes in yeast, nearly 101 represent genes that code for ribosomal proteins. These 101 genes that code for ribosomal proteins are highly expressed and they represent approximately 90% of the splicing substrate in actively growing yeast cells (Ares et al., 1999; Warner, 1999).

The occurrence of cotranscriptional splicing was first observed for chorion pre-mRNA in fruit flies, *Drosophila melanogaster* (Osheim et al., 1985). Further research revealed that splicing of most eukaryotic introns also occurs cotranscriptionally. In *S. cerevisiae*, splicing of most introns was believed to occur post-transcriptionally when the polymerase reaches 500 nucleotides downstream of the 3' splice sites (Lacadie et al., 2006). It was argued that the spliceosome assembles during transcription, but the short length of yeast introns made it difficult for the spliceosome to complete the splicing process before termination of transcription. Recent advances in the global analysis of nascent RNA, however, has explicitly shown that most yeast introns are also spliced

cotranscriptionally (Barrass et al., 2015; Carrillo Oesterreich et al., 2010). Specifically, splicing was detected when the polymerase is just 40 nucleotides downstream of the 3' splice sites (Carrillo Oesterreich et al., 2010).

The splicing of pre-mRNA takes place by two transesterification reactions that remove the intron and ligate the two exons to form a mature mRNA (Grabowski et al., 1985; Padgett et al., 1984; Ruskin et al., 1984). Splicing is facilitated by a large multi-subunit complex called spliceosome (Brody and Abelson, 1985; Will and Lührmann, 2011). The ordered assembly of five small nuclear ribonucleoprotein particles (snRNPs), and a number of non-snRNP proteins over the intron form a functional spliceosomal complex (Wahl et al., 2009; Will and Lührmann, 2011) (Figure 1.5). The snRNP component contains U-rich snRNAs, U1, U2, U4, U5, and U6 as well as associated interacting proteins that have a specific function to recognize conserved RNA sequences within the intron. Briefly, in the first step U1 snRNP and U2 snRNP bind to the 5' splice site and the branch point, respectively (Parker et al., 1987; Ruby and Abelson, 1988). The base pairing of U2 snRNA occurs with the conserved branchpoint sequence on the intron except at an adenosine residue (Parker et al., 1987). In the subsequent step, the preformed tri-snRNP of U4/U6 plus U5 joins to form the complete pre-catalytic spliceosome (Staley and Guthrie, 1998). In the tri-snRNP, the U4 and U6 snRNAs are extensively base paired forming a three-way junction. U4 and U6 snRNA base pairing is then unwound which assists in the formation of a short duplex between the 5′ and 3′ ends of the U2 and U6 snRNA, respectively (Hardin et al., 2015; Lescoute and Westhof, 2006). U6 further base pairs with the 5' splice site that results in destabilization of the U1 snRNP. This is followed by a series of conformational changes including the release of U1 and

U4 snRNPs, bringing the two splice sites in close proximity to form an active spliceosome (Brow, 2002). As a result, the branchpoint adenosine bulges out of an intramolecular helix and becomes available for the first nucleophilic attack. This results in a nucleophilic attack by the 2' hydroxyl group of the adenosine to the exon-5' splice site junction. The first transesterification step leads to the formation of a lariat structure with a phosphodiester bond between the 5' end of intron sequence GU and the adenosine present at the branchpoint. In the second transesterification step, the 3' hydroxyl group present at the end of the first exon breaks the phosphodiester bond between the 3' SS and the second exon. This results in joining of the two adjacent exons and release of the intron lariat along with U2, U5, and U6 snRNPs, thus completing the process of splicing (Wahl et al., 2009).

In addition to snRNPs, non-snRNPs play a critical role in regulation of assembly, activation and disassembly of spliceosomes during the process of splicing. A special mention of protein complex associated with Prp19 known as NineTeen Complex (NTC) is necessary as it is docked onto the spliceosome during the association of the tri-snRNP (U4/U6.U5) and it stays until splicing is complete. So far eight core proteins; Cef1, Snt309, Prp46, Syf1, Syf2, Syf3/Clf1, Isy1 and Ntc20, along with Prp19 have been recognized in NineTeen Complex (Fabrizio et al., 2009). In addition to this, eighteen other accessory proteins were found to be associated with Prp19 (Chanarat and Sträßer, 2013). The central coiled-coil domain of Prp19, which is crucial for its tetramerization, facilitates the assembly of NTC and the spliceosomal complex. The protein composition of the NTC changes during different phases of splicing in response to remodeling of proteins. Specifically, the ubiquitination activity of Prp19 facilitates rearrangements within the spliceosome (Ohi et al., 2003). Furthermore, it has been demonstrated that NTC is

essential for the stable interaction of the U5 and U6 snRNPs with the spliceosome in budding yeast (Chan and Cheng, 2005).

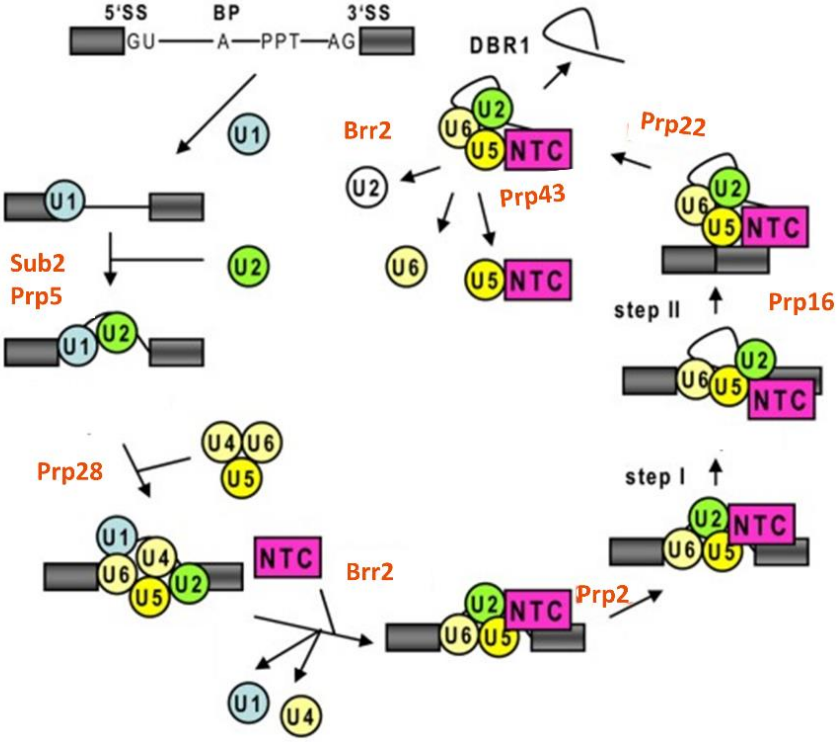

Figure1. 5 Dynamic association of snRNPs and non snRNPs during the splicing reaction. The assembly of U1, U2, U4, U5 and U6 snRNPs in coordination with eight RNA helicases and NineTeen complex components are represented in spliceosome assembly and disassembly (adapted from (Koncz et al., 2012)).

A group of helicases are involved in regulating conformational rearrangements of protein–RNA and RNA–RNA during the splicing reaction. Research to date has identified eight conserved RNA helicases. Of these, four helicases (Prp2, Prp16, Prp22 and Prp43) belong to the DEAH-box family, three RNA helicases (Prp5, Sub2 and Prp28) belong to the DEAD-box family, and one (Brr2/Ski2) to the Ski-2 like family (Cordin and Beggs, 2013). RNA helicases from the DEAD-box family (Sub2, Prp5 and Prp28) are known to function before the association of the NTC with the spliceosome. Except for Prp5, Sub2 and Prp28, the other five helicases are modulated through NTC and the accessory

proteins that associate with NTC during splicing (de Almeida and O'Keefe, 2015). Among them, Prp2 is required for the first catalytic step, whereas Prp16 and Prp22 participate in the second catalytic step of the pre-mRNA splicing (Kim and Lin, 1993, 1996; Schwer and Gross, 1998; Schwer and Guthrie, 1991). Prp22 does not require ATP in the second step of splicing. After completion of the second step of splicing, Prp22 functions to release spliced mRNA from the spliceosomal complex in an ATP-dependent manner (Schwer and Gross, 1998; Wagner et al., 1998). Lastly, Prp43 plays a role in pre-mRNA splicing as well as in rRNA processing (Arenas and Abelson, 1997; Combs et al., 2006; Leeds et al., 2006).

1.4 Termination of transcription

When the elongating RNAPII complex approaches the 3′ end of genes, two processes facilitate the termination of transcription cycle. First, the nascent transcript is cleaved, polyadenylated and released. Second, the ternary complex of DNA-RNAPII-RNA is disrupted following which the polymerase dissociates from the DNA template

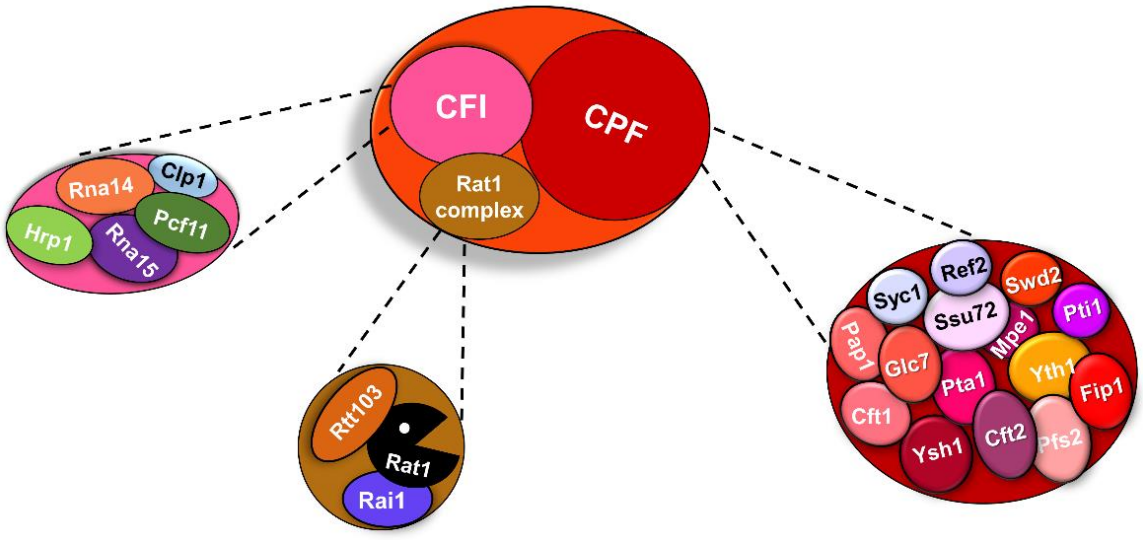

Figure1. 6 Termination complexes in *S.cerevisiae.* CF1 and CPF complexes are involved in cleavage and polyadenylation of mRNA, while Rat1 functions after this step to disengage RNAPII from DNA template.

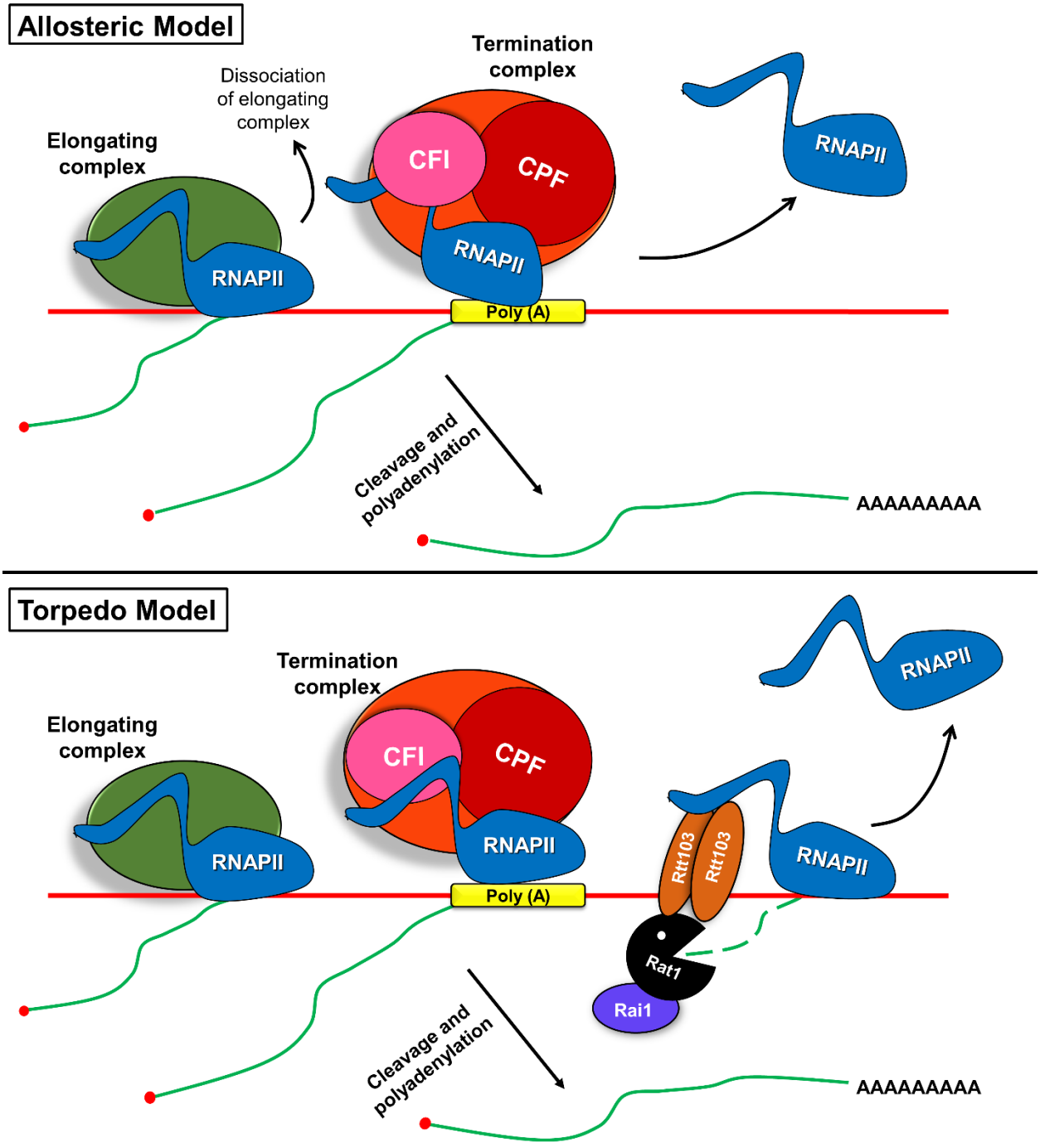

Figure1. 7 Schematic representation of the two proposed models explaining termination of RNAPII transcription (concept adapted from Luo and Bentley 2006). Allosteric model (top panel) shows change in conformation during the transition from elongation to termination when transcribing poly(A) site (indicated with change in shape). Torpedo model (bottom panel) shows degradation of uncapped RNA transcribed by RNAPII downstream of the poly(A) cleavage site by Rat1 exoribonuclease activity.

(Mischo and Proudfoot, 2013). The successful execution of these two steps results in

termination of transcription. Termination requires *cis*-acting elements and *trans*-acting factors. Inhibition of termination by mutation of either *cis*-acting sequence elements or the *trans*-acting protein factors can cause the polymerase to continue the transcription of downstream genes. This phenomenon is known as transcription interference (Candelli et al., 2018; Proudfoot, 1986). Furthermore, termination is a critical step as it allows recycling of RNAPII back to the promoter for reinitiation of transcription (Dieci and Sentenac, 1996), stabilizes mRNA by polyadenylation, and helps in recruitment of the ribonucleoprotein particles (RNPs) that facilitate nuclear to cytoplasmic transport of the mRNA (Lykke-Andersen and Jensen, 2007; Richard and Manley, 2009).

Over the years, in-depth analyses of the molecular mechanism underlying termination of transcription has led to the emergence of two models; the allosteric model and torpedo model (Figure 1.7). The ‘allosteric’ model suggests that termination ensues due to a change in the conformation of the elongating complex leading to the recruitment of the 3' end processing factors and displacement of the anti-termination factors (Logan et al., 1987). In budding yeast, Cleavage Factor 1 (CF1) and Cleavage Polyadenylation Factor (CPF) complexes bring about cleavage of nascent mRNA followed by addition of a poly(A) tail to the 3’ end of mRNA (Figure 1.6). The ‘torpedo’ model proposes that an exoribonuclease uses the 5' uncapped nascent RNA, generated following cleavage and polyadenylation of precursor mRNA, as a substrate. The degradation of the uncapped nascent RNA leads to dissociation of RNAPII from the DNA template (Connelly and Manley, 1988; Proudfoot, 1989). There is evidence in support of both these models. Each model on its own, however, is insufficient to explain the process of termination.

Termination possibly occurs by a hybrid model that involves a combination of allosteric and torpedo mechanisms (Luo et al., 2006; Rosonina et al., 2006).

In addition to CF1 and CPF complexes, termination of transcription by RNAPII also requires the *cis-acting* element called polyadenylation signal (PAS). Transcription of the polyadenylation signal coincides with phosphorylation of the CTD at the serine-2 residue, which in turn facilitates the recruitment of CF1 complex through its Pcf11 subunit, which contains the CTD interaction domain (CID) (Gross and Moore, 2001; Kessler et al., 1996). RNAPII undergoes a brief pause upon transcription of the PAS, which helps in the recruitment of CPF complex (Bienroth et al., 1991; Keller et al., 1991). The CPF subunit, Ysh1, cleaves the nascent transcript using its endonuclease activity. Following the cleavage, poly(A) polymerase (Pap1) adds a poly (A) tail of about 70 nucleotides in length (Zhao et al., 1999). The phosphorylation of the CTD at the serine-2 and threonine-4 residues recruits another termination factor called Rtt103 (Jasnovidova et al., 2017; Kim et al., 2004b). . Rtt103 forms homodimers through coiled-coil domain interaction (43). Rtt103, which has a CTD interaction domain is in a complex with Rat1 (Xrn2 in mammals) and Rai1. Rat1 degrades uncapped mRNA by its 5'→3' exoribonuclease activity. This degradation of RNA disrupts association of the RNAPII from the DNA template leading to termination of the transcription cycle (Fong et al., 2015; Kim et al., 2004b). The exoribonuclease activity of Rat1 is essential but not sufficient for disengaging the polymerase from the DNA template (Pearson and Moore, 2013).

1.5 Discovery of Rat1 in budding yeast

Rat1 (Ribonucleic Acid Trafficking 1) was first described as a protein involved in transport of poly(A)-containing mRNA from the nucleus to the cytoplasm in budding yeast. In a temperature-sensitive mutant of Rat1 called *rat1-1*, mRNA was cleaved and polyadenylated but failed to move out of the nucleus (Amberg et al., 1992). Simultaneously, Rat1 was also identified in two other independent UV-mutagenesis studies in budding yeast with different functional phenotypes. In the first study, Rat1 was accidently identified in a screen for genes involved in the translocation of proteins to mitochondria. The temperature-sensitive trait of Rat1 in this mutant, however, was later identified as an independent phenotype not related to the mitochondrial translocation defect. A detailed sequence examination of the *RAT1* gene helped to establish its homology to the cytoplasmic exoribonuclease, Xrn1 (also referred to as Kem1) (Kenna et al., 1993). Similar to Xrn1, Rat1 was found to be a 5'→3' exoribonuclease that digests 5' end monophosphorylated RNA substrate (Stevens and Poole, 1995). Rat1 was cited previously in a few studies as *HKE1* (Homolog of KEM1). In another independent study, Rat1 was identified as a suppressor of a mutation in the TFIIIC-binding site in the promoter of tRNA genes. Here, the gene for Rat1 was referred to as *TAP1* (Transcription Activation Protein 1), as the temperature-sensitive mutant *tap1-1* was able to overcome the transcriptional activation defect of tRNA gene with a mutated TFIIIC-binding site (Di Segni et al., 1993).

The above mentioned phenotype of nuclear retention of poly(A) transcripts resulted due to the substitution of tyrosine to cytosine at position 657 of Rat1, whereas

the tRNA internal promoter activation independent of TFIIIC was observed in the mutant of Rat1 exhibiting tyrosine to histidine substitution at 683 position (Aldrich et al., 1993; Luo et al., 2006). Although both these mutations associated with Rat1 were situated outside the exoribonuclease domain of Rat1, the functional effect of the mutant protein varied significantly. It was proposed that both these phenotypes of Rat1 protein were apparently due to an underlying change in the chromosomal structure caused by the mutant Rat1 protein.

1.6 Rat1 is an evolutionary conserved protein

Rat1 is a 5′→3′ exoribonuclease belonging to the Xrn-family of nucleases. It is a highly conserved protein with homologs present in fission yeast, flies, worms, zebra fish, plants, mice and humans (Johnson, 1997; Kastenmayer and Green, 2000; Richter et al., 2016; Shobuike et al., 1995; Sugano et al., 1994; Zhang et al., 1999b). Except for flies, Rat1 is referred by different names in these organisms. Rat1 is referred to as Dhp1 in fission yeast, DhmI in mice, and Xrn2 in worms, zebra fish, plants, and humans. Unlike other organisms, there exists three Rat1 homologs in *Arabidopsis thaliana.* They are referred to as AtXrn2, AtXrn3 and AtXrn4 (Kastenmayer and Green, 2000). They are believed to have evolved by gene duplication. Rat1/Xrn2 are essential for viability only in yeast and worms (Amberg et al., 1992; Miki et al., 2014a). Except for AtXrn4, Rat1 and its homologs are predominantly localized in the nucleus. Loss of nuclear localization of AtXrn4 was attributed to the partial loss of its bipartite nuclear localization sequence (Kastenmayer and Green, 2000).

The three-dimensional architecture of Rat1 homologs Dhp1 and Xrn2 from fission yeast and worms respectively was elucidate by X-ray diffraction approach (Richter et al., 2016; Xiang et al., 2009). The crystal structure of fission yeast Dhp1 in complex with its interacting protein Rai1 revealed the presence of two conserved regions in the protein called conserved region 1 (CR1) and conserved region 2 (CR2). CR1 spans residues 1-387, whereas CR2 spans residues 602-766. These regions are separated by a lesser conserved linker region (Figure 1.8). The CR1 region is comprised of acidic residues and is a part of the active site of the enzyme. The CR2 region forms a pocket like structure that surrounds the active site. This arrangement is significantly different from other related nucleases that are in open conformation. The pocket conformation of the active site

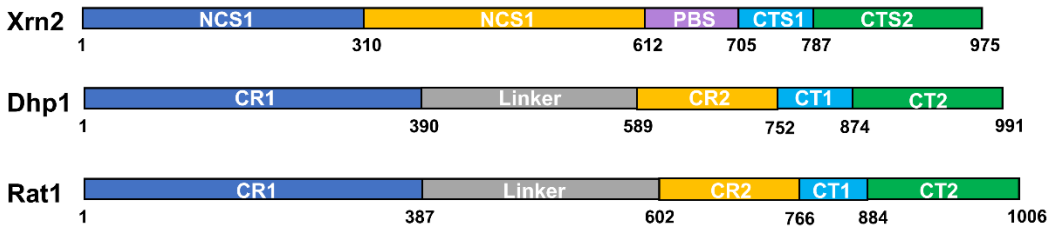

Figure1. 8 Representation of domain organization of Dhp1 from *S. pombe,* Rat1 from *S. cerevisiae, and* Xrn2 from *C.elegans*. CR1 and CR2 represents the two conserved regions. NCS1 and NCS2 are the two nuclease core segments. PBS (purple) is the Paxt-1 binding domain. CTS and CT (in light blue and green) represent poorly conserved C-terminal regions (adapted from Xiang et.al., 2009 and Richter et.al., 2016).

contains seven acidic amino acids and requires Mg^{+2} or Mn^{+2} divalent ions for the exoribonuclease activity. The region following CR2 is poorly conserved and the interaction of this region (residues 766-884) with the CR1 is known to form the Rai1 binding region (Xiang et al., 2009). A different scheme was adopted to define the conserved and less conserved regions of Xrn2 from worms. The rationale behind this characterization was to dissect the regions of Xrn2 with respect to its functional interaction with Paxt-1. The homolog of Paxt-1, however, has not been found in yeast (Miki et al.,

2014b). Xrn2 binding domain (XTBD) of Paxt-1 is located at the N-terminal of the protein that interacts with the region spanning residues 613-706 of Xrn2. This region overlapped with the CR2 region of fission yeast Dhp1. The full length protein can be divided into five functional domains, which are nuclease core segment 1 (NCS1, residues 1-310), nuclease core segment 2 (NCS2, residues 311-612), Paxt-1 binding segment (PBS; 613-705), C-terminal segment 1 (CTS1; 706–787) and C-terminal segment 2 (CTS2; 788–975) (Richter et al., 2016). The evidence of the sequence similarity across species strongly supports the notion that Rat1 and its homologs are evolutionarily conserved proteins. Functional conservation of the protein across species was established using complementation analysis in a temperature-sensitive mutant of yeast. Fission yeast Dhp1 as well as two plant homologs of Xrn2 (AtXrn2 and AtXrn3) rescued the temperature-sensitive phenotype of budding yeast *rat1-1* strain (Kastenmayer and Green, 2000; Sugano et al., 1994). Similarly, murine homolog of Rat1 (Dhm1) rescued the temperature-sensitive phenotype of fission yeast *dhp1-1* mutant (Shobuike et al., 1995). The similarity of the sequence, structure, and functional complementation across species confirms that Rat/Xrn2 is an evolutionary conserved protein.

1.7 Rat1/Xrn2 functions in the termination of transcription

Rat1 has been found to associate with the terminator region of genes. Its role in termination of transcription has been proven through both single gene as well as genomewide analyses (Baejen et al., 2017; Fong et al., 2015; Kim et al., 2004b; Larochelle et al., 2018; Miki et al., 2017; West et al., 2004). Following cleavage and polyadenylation, RNAPII continues to transcribe past the polyadenylation site, resulting in a nascent transcript that is not capped at its 5' end. This makes the newly synthesized

nascent transcript a substrate for Rat1 exoribonuclease activity. The degradation of the nascent transcript destabilizes the association of RNAPII with the DNA template, leading to its dissociation from the template and ultimately resulting in termination of transcription. The role of Rat1 in termination is in alignment with the torpedo model. A genomewide analysis, however, found only 30% of the genes in budding yeast are dependent on Rat1 for termination of their transcription (Baejen et al., 2017). Thus, unlike CF1 and CPF complexes, Rat1 is not an essential termination factor in budding yeast. The Rat1 mammalian homolog, Xrn2, exhibits a similar role in termination of transcription (Kim et al., 2004b; West et al., 2004).

The mechanistic details of the role of Rat in termination of RNAPII-mediated transcription is not clearly understood. Due to the lack of a uniformly defined *in vitro* biochemical assay, the role of Rat1 and its interacting partner Rai1 in termination of transcription cannot be validated to its role inside the cell. Two out of three studies that attempted to develop an *in vitro* assay for testing the role of Rat1 in termination pointed out that Rat1 and Rai1 on their own are capable for terminating transcription (Dengl and Cramer, 2009; Park et al., 2015; Pearson and Moore, 2013). Furthermore, it was also demonstrated that NTP misincorporation and the length of RNA transcript undergoing degradation by Rat1 are key factors that dissociate RNAPII during transcription (Park et al., 2015). Another independent study, however, came to a different conclusion. This study showed that Rat1/Rai1 cannot facilitate termination independently in an *in vitro* transcription system (Dengl and Cramer, 2009). This failure of Rat1 to bring about termination in this study was cited due to the use of a scaffold elongation complex instead of a promoter-driven elongation complex (Pearson and Moore, 2013). Thus, factors

required by Rat1 for facilitating termination and the mechanism of RNAPII dissociation by the protein remain obscure. Rat1-mediated termination, however, is dependent on promoter-driven transcription, possibly due to promoter-terminator crosstalk or gene looping. This view is supported by the observation that Rat1 is localized to both the 5′ and 3′ ends of a gene (Baejen et al., 2017; Kim et al., 2004b). This *in vitro* study further demonstrated that Rat1 exoribonuclease activity is crucial but not sufficient for its termination function as CTD interacting factor Rtt103 was able to rescue the termination defective phenotype of exoribonuclease inactive Rat1 mutant (Pearson and Moore, 2013). This finding suggests that factors capable of binding the CTD can disengage RNAPII independent of Rat1 activity. In another study, a similar observation was recorded with CTD binding protein Pcf11. Pcf11 is a subunit of CF1 termination complex that functions to facilitate the cleavage step during transcription termination (Zhang et al., 2005). These studies imply that RNAPII disassociation from the DNA template does not require cleavage or Rat1 exonuclease activity but instead was shown to be mediated by poly(A)-site dependent conformational change of the elongation complex**.** (Pearson and Moore, 2013; Zhang et al., 2015, 2005).

In line with the inconclusive nature of Rat1 function in termination in budding yeast, there are conflicting reports about the role of Xrn2 in termination of transcription in mammals as well. According to one report, there is no significant change in RNAPII profile downstream of the gene upon inactivation of Xrn2 in mammals, thereby supporting the view that Xrn2 is not a general termination factor just like in budding yeast, and there exists an unexplored mechanism that can disengage RNAPII from the DNA template during transcription termination (Nojima et al., 2015). There are, however, two reports that

strongly support an active role of Xrn2 in termination of transcription by RNAPII. In the first study, complementation of cells inactivated of native Xrn2 with an exoribonuclease defective mutant of Xrn2 resulted in the polymerase reading through the termination signal across most protein coding genes. It was suggested that there exists a kinetic competition between the elongation complex and Xrn2, which is altered in an exoribonuclease defective mutant of Xrn2 leading to RNAPII signal downstream of the poly(A) sites (Fong et al., 2015). In the second study, auxin-induced degradation of Xrn2 by degron approach resulted in the stabilization of RNA downstream of the poly(A) site on a genomewide scale. The logical conclusion of this result is that Xrn2 plays a general role in termination of transcription of protein coding genes in mammalian systems (Eaton et al., 2018).

Xrn2 has also been found in *C. elegans*, where it functions in the termination step of transcription just like in yeast and higher eukaryotes (Richter et al., 2016). Xrn2 of *C. elegans* forms a complex with Paxt-1. The homolog of Paxt-1, however, has not been found in yeast. Xrn2 binding domain (XTBD) of Paxt-1 is located at the N-terminal of the protein (Miki et al., 2014b). Similar to Rat1 in yeast, Xrn2 in *C. elegans* was shown to be important in transcription termination of only a subset of genes termed Xrn2-dependent genes (XDT) (Miki et al., 2017). Thus, Rat1/Xrn2 may not be considered a general termination factor in *C. elegans*. This study not only demonstrated that termination in *C. elegans* differs from that in yeast and humans, but also suggested that the promoter of the genes play a critical role in regulating termination in *C. elegans.* These conclusions were based both on the basis of single gene analysis as well as genomewide analysis. The study went on to show that both Xrn2-dependent and Xrn2-independent genes

exhibited recruitment of Xrn2 at the 3’ end of genes. All the genes carrying Xrn2 at the terminator did not exhibit dependence on the protein for their termination. These results strongly suggest that the mere presence of a factor does not guarantee its role in the process. Metagene analysis showing crosslinking of a known termination factor towards the 3’ end of genes in yeast and mammals should be corroborated with a functional assay to confirm its role in the process.

In addition to termination at the 3’ end of genes, Rat1/Xrn2 have also been reported to terminate transcription within the body of the gene if transcripts are aberrantly capped or unspliced (Bousquet-Antonelli et al., 2000; Brannan et al., 2012; Jimeno-González et al., 2010).

The role of Rat1 is not restricted to RNAPII-mediated termination of transcription. The factor has been implicated in termination of transcription by RNA polymerase I (RNAPI) as well. Briefly, Rat1’s exoribonuclease activity is involved in degrading RNAPI associated nascent transcripts after endonucleolytic cleavage of the rRNA by Rnt1. Following the cleavage, Rat1 and Rai1 function to degrade the nascent transcript generated by RNAPI downstream of the cleavage site (El Hage et al., 2008). This process is analogous to the torpedo mechanism proposed for RNAPII termination. Furthermore, Sen1 was identified as the helicase that works in coordination with Rat1(Kawauchi et al., 2008). In *rat1-1 sen1-1* double mutant, the termination read-through phenotype of RNAPI was found to be more pronounced as compared to in either single mutant (Kawauchi et al., 2008). Mass spectrometry analysis of affinity purified Xrn2 preparations revealed the presence of TTF-I and Rsf1, which are termination factors that function in RNAPI

transcription. The possibility of Xrn2 playing a similar role in RNAPI transcription in mammals cannot be ruled out (Brannan et al., 2012).

1.8 Role of Rat1/Xrn2 in the elongation of transcription

The role of Rat1/Xrn2 is not restricted to the termination step of transcription. There is a substantial amount of evidence implicating Rat1 and its human homolog Xrn2 in the elongation step of transcription. First, *rpb1-N488D,* which is a mutant of RNAPII that exhibits slow elongation rate phenotype in yeast, rescued the growth-defective phenotype of *rat1-1* at elevated temperature. RNAPII density-ChIP within the body of genes implied fast elongation rate in *rat1-1* mutant. Furthermore, RNAPII ChIP signal in the *rpb1-N488D* mutant was lower as compared to its level in the *rpb1-N488D–rat1-1* double mutant (Jimeno-González et al., 2010). An interpretation of these results is that elongation rate of RNAPII is influenced by Rat1. Second, RNAPII CTD serine-2 phosphorylation is a critical mark that signals recruitment of elongation factors during the transcription cycle. In *rat1-1* mutant, an increase in CTD serine-2 phosphorylation signal was observed apparently due to the increased loading of Ctk1 on the transcription track. The faster elongation rate observed in *rat1-1* mutant was apparently due to increased CTD serine-2 phosphorylation associated with the mutant (Jimeno-González et al., 2014). Consequently, combining the *rat1-1* allele with the *rpb1-E1103G* mutant, which causes a faster polymerase elongation rate, resulted in an even stronger growth defect (Jimeno-González et al., 2010; Malagon et al., 2006). Third, in mammalian cells, an independent chemical genetic screen performed to identify putative substrates of the Cdk9, which is a subunit of P-TEFb complex, identified Xrn2 as a substrate. This study further found that the phosphorylation of threonine-439 residue of Xrn2 is critical for its recruitment onto the

chromatin as well as for its termination function (Sansó et al., 2016). This study along with the results shown in yeast described above imply that the role of Rat1/Xrn2 in termination is influenced by kinases that play a key role in the elongation phase of transcription. Fourth, in the presence of catalytically inactive Xrn2 in the mammalian system, RNAPII occupancy was observed further downstream of the termination region. An identical phenotype was observed in *rpb1-E1126G*, which is the yeast mutant of RNAPII that exhibits fast elongation rate. The similarity in phenotype of RNAPII occupancy downstream of the termination region in *rpb1-E1126G* and Xrn2 knockdown mutant indicate that RNAPII elongation rate is critical for termination of transcription (Fong et al., 2015). Lastly, analysis of factors associated with the early productive and late stage RNAPII elongation complexes in yeast revealed enrichment of Rat1 and Rai1 in the early RNAPII elongation complex (Harlen and Churchman, 2017). These results corroborated the previous finding that Rat1 and Rai1 occupy the 5' end of the gene (Kim et al., 2004b; Nojima et al., 2015). These observations argue that Rat1 plays a broader role in the transcription cycle than previously thought, and there is a need to revisit the role of Rat1 in transcription and cotranscriptional processing.

1.9 Dhp1, fission yeast homolog of Rat1, in heterochromatin formation

Dhp1, the fission yeast homolog of Rat1, has a novel role in heterochromatin formation (Chalamcharla et al., 2015; Tucker et al., 2016). There are, however, two contradicting hypotheses explaining the involvement of Dhp1 in heterochromatin dynamics. According to one view, Dhp1-mediated termination is an obligate requirement for heterochromatin assembly, while the alternative view postulates that Dhp1-mediated heterochromatin formation is independent of its termination function. According to the first

model, Dhp1 facilitates non-canonical termination at the site that is targeted for heterochromatin formation. During this termination process, Dhp1 associates with a diverse set of binding partners like Clr4 methyltransferase complex and a host of RNA binding and processing factors. These factors in turn bring about heterochromatin assembly (Chalamcharla et al., 2015). An independent study published at the same time, however, contradicted the termination-heterochromatin link, and instead advocated the view that they are independent events regulated by Dhp1 (Tucker et al., 2016).

1.10 Rat1 is involved in the nuclear quality control mechanism

The nuclear localization and presence of 5'→3' exoribonuclease activity allow Rat1/Xrn2 to function in nuclear RNA surveillance by degrading aberrant transcripts that are improperly spliced or capped. In fact, both Rat1 and Xrn2 have been implicated in the degradation of aberrant transcripts in yeast as well as in mammals. In budding yeast, Rat1 was linked to degradation of unspliced transcripts in the mutant of splicing factor Prp2, while Xrn2 was reported to co-transcriptionally degrade unspliced transcripts in human cells treated with splicing inhibitor spliceostatin (Bousquet-Antonelli et al., 2000; Davidson et al., 2012). Prp2 is a helicase that functions in the first step of splicing in budding yeast. In *prp2-1,* which is the temperature-sensitive mutant of Prp2, splicing was compromised, and unspliced transcripts accumulated at non-permissive temperature. In the double mutant *prp2-1/rat1-1*, a higher level of unspliced transcripts was observed compared to *prp2-1* single mutant (Bousquet-Antonelli et al., 2000). The authors concluded that the increased level of unspliced transcripts in the double mutant was due to stabilization of unspliced transcripts in the absence of Rat1 activity. In mammalian cells, depletion of Xrn2 by RNAi mechanism resulted in accumulation of unspliced transcripts

under conditions when splicing was compromised either by 3' splice site mutation or by inhibition of splicing by spliceostatin (Davidson et al., 2012). Both these studies indicated that Rat1/Xrn2 can function in degrading unspliced transcripts. There was, however, no explanation regarding how Rat1/Xrn2 recognizes and gains access to its substrate for degradation of the unspliced transcript. This is significant as Rat1/Xrn2 prefers degrading 5' monophosphorylated RNA due to the closed confirmation of its active site pocket (Richter et al., 2016; Xiang et al., 2009).

Nuclear decapping of RNA can lead to the formation of transcripts that are monophosphorylated at the 5' end, and therefore are a substrate for the exoribonuclease activity of Rat1/Xrn2. Studies in yeast have demonstrated that Rat1 interacts with Rai1 (Din1 in *S. pombe)* and this interaction can facilitate processing of an RNA substrate at the 5′ end in a manner that makes it a substrate for Rat1 exoribonuclease activity. A homolog of Rai1 is present in higher eukaryotes including flies, worms, mouse, and human. In these organisms, Rai1 is referred to as Dom3Z (Xue et al., 2000). The interaction between homologs of Rai1 and Rat1, however, is not conserved in these organisms (Xiang et al., 2009). Furthermore, in yeast, there exists a cytoplasmic homolog of nuclear Rai1, known as Decapping exonuclease 1 Dxo1 (Chang et al., 2012). Specifically, using *in vitro* studies it was demonstrated that Rai1 possess a wide range of enzymatic activities including 5' pyrophosphohydralase activity and decapping activity towards unmethylated, capped RNA substrates in *S. pombe* and *S. cerevisiae,* respectively (Jiao et al., 2010; Xiang et al., 2009). *S. pombe* Rai1 pyrophosphohydralase activity hydrolyzes the 5'-end triphosphate of an uncapped RNA (pppRNA) to release diphosphate and a 5' monophosphorylated RNA, which could be degraded by the 5'→3'

nuclear exoribonuclease of Rat1, suggesting the possible role of the protein in the quality control mechanism. Supporting this idea, Rat1 and Rai1 have been reported to target RNA that lacks a proper cap structure in yeast (Jiao et al., 2010; Jimeno-González et al., 2010). Unlike decapping enzyme Dcp2, which acts on methylated capped transcripts, Rai1 has been shown to possess an unusual decapping activity towards unmethylated capped transcripts. This activity of Rai1 was shown to release GpppN from an unmethylated capped RNA in *S. cerevisiae* under *in vivo* conditions. In *rai1Δ*, incompletely capped mRNAs accumulate under nutritional stress conditions like glucose or amino acid starvation (Jiao et al., 2010). Unlike Rai1, Dxo1 contains decapping activity towards methylated RNA substrates and intrinsic 5′→3′ exoribonuclease activity, but it lacks pyrophosphohydralase activity. This was demonstrated by *in vitro* studies using Dxo1 from the yeast, *Kluyveromyces lactis* (Chang et al., 2012).

Mammalian cells contain a homolog of Rai1p and Dxo1p termed Dom3Z, and it possesses the collective enzymatic actives of both proteins (Jiao et al., 2013). Since Dom3Z doesn't interact with mammalian Xrn2, an alternate quality control mechanism of Xrn2 was proposed (Brannan et al., 2012). In this study, Xrn2 was found to associate with decapping proteins Edc3, Dcp1a, and Dcp2. This association was proposed to function first in cotranscriptional decapping, generating a 5' monophosphate end on the nascent RNA for Xrn2-mediated degradation. This degradation further leads to premature termination of transcription. Promoters in humans are mostly bidirectional in nature that commonly generate upstream anti-sense transcripts initiating from the promoter of protein-coding genes (Core et al., 2008; Seila et al., 2008). Xrn2 was also shown to limit

upstream anti-sense transcription from bidirectional promoters by a similar mechanism (Brannan et al., 2012).

Lastly, aberrant pre-rRNA transcripts were found to be eliminated by coordinated action of both 3'→5' exoribonuclease of nuclear exosome and 5'→3' exoribonuclease activity of Rat1 in budding yeast (Fang et al., 2005). Hypo-modified tRNAs, such as those lacking proper methylation required for tRNA function, can also be degraded by Xrn2 (Chernyakov et al., 2008).

1.11 Rat1 is also involved in RNA processing

Rat1/Xrn2 perform another important function in eukaryotic cells. They are essential for processing of rRNA. Genes coding for rRNA in eukaryotes are localized in the rDNA locus localized in nucleolar compartments (Figure 1.9). It consists of multiple,

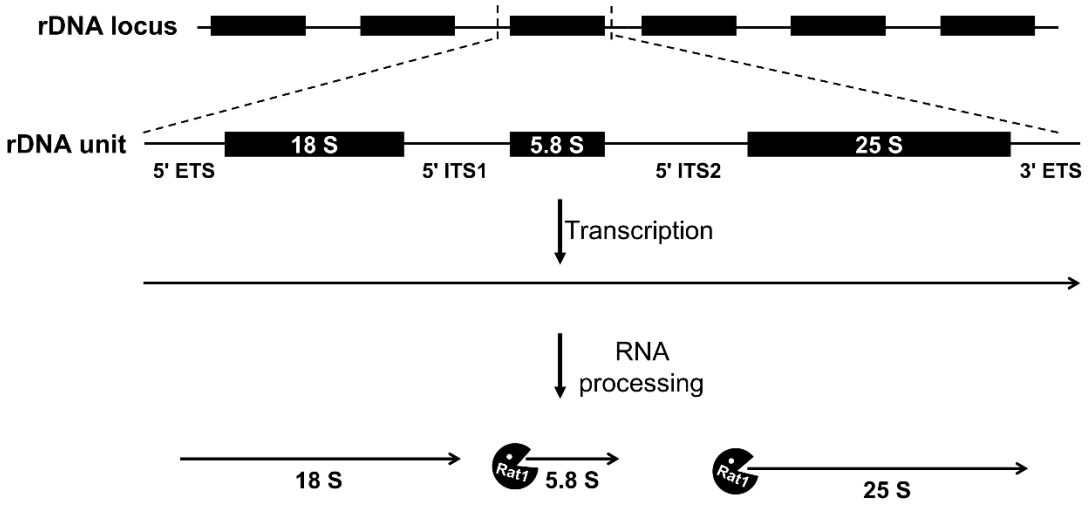

Figure1. 9 Transcription and RNA processing at rDNA locus. Ribosomal RNA (rRNA) gene units organized in tandem repeats at rDNA locus. A single unit is shown in detail with 18S, 5.8S and 25S genes External transcribed spacer (ETS) and internal transcribed spacer (ITS) are the regions processed from the pre-rRNA transcript (adapted from Henry et.al., 1994).

tandem repeats of genes coding for different species of rRNA. Each repeat consists of

genes coding for 18S rRNA, 5.8S rRNA, and 25S rRNA flanked at the 5' and 3' ends by external transcribed spacer elements 5' ETS and 3' ETS. Two internal transcribed spacer elements (ITS), ITS1 and ITS2, separate the 18S rRNA, 5.8S rRNA and 25 rRNA gene. Each rDNA repeat is transcribed as a polycistronic pre-35S rRNA. Rat1 plays a crucial role in processing of pre-35S rRNA into mature rRNA molecules (Amberg et al., 1992). Processing of 5.8S pre-rRNA was affected in the absence of Rat1 and accumulation of longer precursor 5.8S rRNA was observed in mutants of Rat1 (Amberg et al., 1992; Henry et al., 1994). It was further identified that deletion of Rai1, the interacting partner of Rat1, resulted in a similar increase in level of the longer precursor of 5.8S rRNA in cells. These results imply that both Rat1 and Rai1 function together in the processing of the 5.8S rRNA (Fang et al., 2005; Xue et al., 2000). Rat1 also functions in the processing of pre-27S rRNA in addition to its role in the processing of the 5.8S rRNA (Geerlings et al., 2000). Xrn2 was similarly shown to be involved in processing of the 5' ends of 5.8S and 28S rRNAs in mouse cells and plants (Wang and Pestov, 2011; Zakrzewska-Placzek et al., 2010). This evidence suggests that there exists a conserved mechanism by which exonuclease Rat1/Xrn2 function in processing of precursor rRNA from yeast to mammal (Wang and Pestov, 2011).

Small nucleolar RNA (snoRNA) is a class of non-coding RNA that is mostly derived from intronic sequences. A few snoRNA, however, are of non-intronic origin and are transcribed directly from the genomic location outside introns. snoRNA are involved in the site-specific processing and modification of the pre-rRNAs (Vincenti et al., 2007; Weinstein and Steitz, 1999). There are two models explaining the synthesis of snoRNA. According to the first model, snoRNA synthesis from an intron requires the activity of

debranching enzyme, Dbr1. Dbr1-mediated debranching of the lariat produces the substrate that is processed by the exoribonuclease activity of Rat1 into mature snoRNA (Ooi et al., 1998; Petfalski et al., 1998). Alternatively, two snoRNA can be transcribed as a polycistronic unit, which is then cleaved by the endonuclease activity of Rnt1. Rat1 functions in the subsequent step by trimming the available 5' end of the precursor snoRNA (Chanfreau et al., 1998; Petfalski et al., 1998; Qu et al., 1999).

In *C. elegans*, Xrn2 was found to degrade a subset of mature microRNAs (miRNA). Using the let-7 family of miRNA, it was discovered that Xrn2-mediated degradation occurs after disengaging miRNA from the argonaute protein. Furthermore, it was determined that only mature miRNA and not the pre-miRNA are the substrate of Xrn2. It was found that Xrn2-mediated destabilization of let-7 does not occur if the miRNA-argonaute complex is bound to the target sequence (Chatterjee and Grosshans, 2009). When the miRNA-argonaute complex is not interacting with its target, Xrn2 can disassemble the miRNA from the complex and subsequently degrade it. This role of Xrn2 has been exhibited in an *in vitro* system, which has yet to be further validated using studies *in vivo*. These results suggest that Xrn2 plays a role in homeostasis of miRNA in worms.

Studies over the years have established that Rat1/Xrn2 play role in diverse cellular processes; including transcription, rRNA processing, heterochromatin formation, and RNA quality control mechanism. Specifically, Rat1 role in the initiation and elongation steps of transcription suggests that it has much broader role in transcription cycle than previously anticipated. Thus, there is a need to revisit the currently defined role of Rat1 in transcription and identify the physiological role of Rat1 during transcription cycle. In the

next chapter, I have discussed how I identified a novel role of Rat1 in the process of co-transcriptional splicing in budding yeast. To address how Rat1 is involved in multiple processes of RNA metabolism, we performed a mass spectrometry analysis of the native complex from budding yeast. In the final chapter, I will support the basis of diverse roles of Rat1 with data from the mass spectrometry analysis of the Rat1-complex.

CHAPTER 2. A NOVEL ROLE OF RAT1 IN COTRANSCRIPTIONAL SPLICING OF PRE-MRNA IN BUDDING YEAST

2.1 Abstract

Rat1 and its human homolog Xrn2 have been implicated in multiple nuclear processes. Here we report a novel role of Rat1 in splicing of mRNA. We observed an increase in the level of unspliced transcripts in mutants of Rat1. Accumulation of unspliced transcripts was not due to the surveillance role of Rat1 in degrading unspliced mRNA, or an indirect effect of Rat1 function in termination of transcription, or due to an increased elongation rate in Rat1 mutants. ChIP-Seq analysis revealed Rat1 crosslinking to the introns of a subset of yeast genes. Mass spectrometry analysis revealed presence of Clf1, Isy1, Yju2, Prp43, and Sub2 splicing factors in purified Rat1 preparation. Furthermore, recruitment of Prp2 splicing factor on the intron was compromised in the Rat1 mutant. On the basis of these results, we propose that Rat1 has a novel role in splicing of a subset of mRNA in budding yeast.

2.2 Introduction

Rat1/Xrn2 are 5′→3′ exoribonucleases belonging to the Xrn-family of nucleases (Nagarajan et al., 2013). They are highly conserved proteins with homologs present in budding yeast, fission yeast, flies, plants, worms, mice and humans (Johnson, 1997; Kastenmayer and Green, 2000; Richter et al., 2016; Shobuike et al., 1995; Sugano et al., 1994; Zhang et al., 1999b). Rat1/Xrn2 have been implicated in multiple aspects of RNA metabolism in eukaryotes. A multitude of studies carried out in different eukaryotic systems have demonstrated involvement of Rat1/Xrn2 in RNA trafficking, RNA quality control, RNA processing, promoter-associated transcription, elongation, and termination

steps of transcription (Amberg et al., 1992; Bousquet-Antonelli et al., 2000; Brannan et al., 2012; Jiao et al., 2010; Jimeno-González et al., 2010; Kim et al., 2004b; Miki et al., 2017; West et al., 2004).

Rat1 was discovered in budding yeast as a factor required for nucleo-cytoplasmic transport of mRNA. In a mutant of Rat1, mRNA was cleaved and polyadenylated, but failed to move out of the nucleus and hence it was named RNA trafficking factor or Rat1 (Amberg et al., 1992). Further investigation revealed that Rat1 and its higher eukaryotic homolog Xrn2 are involved in termination of RNAPII-mediated transcription in a manner dependent on their 5′→3′ exoribonucleases activity (Kim et al., 2004b; West et al., 2004). The poly(A)-dependent termination of transcription involves pausing of the polymerase beyond the poly(A)-site, which facilitates recruitment of CF1 and CPF 3' end processing/termination factors in yeast (Mischo and Proudfoot, 2013). The transcribing mRNA is cleaved downstream of the poly(A)-site by the Ysh1 subunit of CPF complex and is polyadenylated by poly(A)-polymerase (Butler and Platt, 1988; Jenny et al., 1996). The monophosphorylated mRNA downstream of the cleavage site is still attached to the elongated polymerase. Rat1/Xrn2 bind to the uncapped elongating mRNA, digests RNA by the 5′→3′ exoribonucleases activity, catches up with the elongating polymerase, and helps disengage the polymerase from the DNA template (Kim et al., 2004b). The enzymatic activity of Rat1 is essential but not sufficient for termination of transcription (Pearson and Moore, 2013). Genomewide studies have revealed that Rat1 is not a general termination factor like CF1 subunit Pcf11 and CPF subunit Ysh1, but is required for termination of transcription of nearly 35% of RNAPII-transcribed genes in budding yeast (Baejen et al., 2017). The Rat1 homolog Xrn2 facilitates termination of transcription

in higher eukaryotes by a similar mechanism (Eaton et al., 2018; Fong et al., 2015; West et al., 2004). Rat1/Xrn2-mediated termination is reminiscent of the ‘torpedo’ mechanism of termination of transcription in prokaryotes (Jin et al., 1992).

Rat1 may also be involved in termination at the pre-cleavage/polyadenylation step in accordance with the ‘allosteric’ mechanism. The pausing of the polymerase near the poly(A)-site, which is critical for both cleavage/polyadenylation and termination, is dependent on Rat1. In the *rat1-1* mutant, which exhibits a termination defect, there is hyperphosphorylation of CTD-serine-2 resulting in an increased elongation rate. The *rat1-1* termination defect, however, could be rescued by introduction of the *rpb1-N488D* allele, which causes a slower elongation rate (Jimeno-González et al., 2010). These results strongly suggest that, at least in budding yeast, Rat1 has dual roles in termination; a pre-cleavage role in pausing of the polymerase downstream of the poly(A)-site, and a post-cleavage role in degrading 5' monophosphorylated RNA followed by dissociation of the polymerase from the DNA template. Another important conclusion of these results is implication of Rat1 in the elongation step of transcription through its influence on hyperphosphorylation of the CTD at serine-2 (Jimeno-González et al., 2014). Indeed, combining the *rat1-1* allele with the *rpb1-E1103G* mutation, which causes a faster polymerase elongation rate, resulted in an enhanced growth defect (Jimeno-González et al., 2010; Malagon et al., 2006). Depletion of human Xrn2 in human cell lines resulted in promoter-proximal pausing, which also could be the consequence of the role of the protein in early elongation steps (Nojima et al., 2015).

In *C. elegans*, not all the genes that exhibit 3' end occupancy of Rat1 are dependent on the protein for termination. Whether Rat1 is required for termination is

determined by the promoter (Miki et al., 2017). In budding yeast, Rat1 could dismantle only promoter-driven transcription complexes in an *in vitro* assay (Pearson and Moore, 2013). Thus, Rat1 function in transcription could be linked to the promoter-terminator crosstalk. In fact, using 3C assay we demonstrated that promoter-terminator interaction or gene loping is compromised in the *rat1-1* mutant (Ansari, unpublished data). In keeping with the role in gene looping, Rat1/Xrn2 have been found crosslinked to both the 5′ and 3′ ends of genes (Baejen et al., 2017; Fong et al., 2015; Kim et al., 2004b; Nojima et al., 2015). Localization of Rat1/Xrn2 in the promoter-proximal region gave rise to the speculation that the protein is playing a role in promoter-associated transcription. Experimental evidence has implicated Rat1/Xrn2 in the initiation-elongation transition (Harlen and Churchman, 2017). In the absence of Xrn2, a peak of promoter-proximally paused polymerase was observed, thereby suggesting that Xrn2 is involved in early stages of transcription, possibly in facilitating the transition from initiation to elongation phase. Xrn2 has also been found to function in promoter directionality as an increase in polymerase signal was observed in the region upstream of the promoter of genes in the absence of Xrn2 activity (Brannan et al., 2012; Fong et al., 2015). A similar involvement of Rat1 in promoter directionality in budding yeast cannot be ruled out. Both Rat1 and Xrn2 have also been shown to play a role in quality control of mRNA by degrading aberrant uncapped transcripts by 5′→3′ exoribonuclease activity (Brannan et al., 2012; Jiao et al., 2010; Jimeno-González et al., 2010). In line with these results, human Xrn2 has been reported to coimmunoprecipitate with Edc3, Dcp1a and Dcp2 capping proteins (Brannan et al., 2012). No such interaction of Rat1 with capping enzymes has been

reported, but it has been found that Rat1 can similarly degrade uncapped transcripts in budding yeast as well (Jiao et al., 2010).

Rat1/Xrn2 have been implicated in degradation of splicing defective transcripts. Xrn2 was reported to degrade unspliced transcripts in humans, while Rat1 was linked to degradation of unspliced transcripts in the mutant of splicing factor Prp2 (Bousquet-Antonelli et al., 2000). We found accumulation of unspliced transcripts in mutants of Rat1 with a wild type Prp2 allele. An increase in unspliced transcript level was not due to stabilization of transcripts, an indirect effect of defective termination, or loss of polymerase pausing on intronic regions in the Rat1 mutant. We present evidence that Rat1 plays a direct role in splicing. We discovered physical interaction of Rat1 with the intronic sequences as well as with the splicing factors of NineTeen complex (NTC). Furthermore, recruitment of the Prp2 splicing protein on the intron was compromised in the absence of a functional Rat1 in the cell. Our results strongly suggest a novel role of Rat1 in splicing of primary transcripts in budding yeast.

2.3 Results

2.3.1 Unspliced mRNA accumulates in the absence of functional Rat1

Rat1 has been implicated in a variety of nuclear processes in yeast and higher eukaryotes. In order to have a comprehensive understanding of the role of Rat1 in RNA biogenesis in budding yeast, we monitored mRNA levels of selected yeast genes in *rat1-1*, which is the temperature-sensitive mutant of Rat1. RNA analysis was performed in the mutant cells grown at the permissive (25°C) and non-permissive (37°C) temperatures by RT-PCR approach. Briefly, the protocol involved isolation of total RNA from exponentially growing cells, reverse transcription of RNA using an oligo-dT primer, PCR amplification

of resultant cDNA using gene-specific primers and quantification of RT-PCR products as described in El Kaderi et al., (2009). Since Rat1 affects transcription of a number of mRNA species as well as 5.8S and 28S rRNAs, we used 5S rRNA as normalization control in all experiments used in this study. As expected, a decrease in mRNA level of most genes was observed upon shifting of cells to the non-permissive temperature, in accordance with the termination function of Rat1 (Figure 2.1 B). However, we observed an unusual

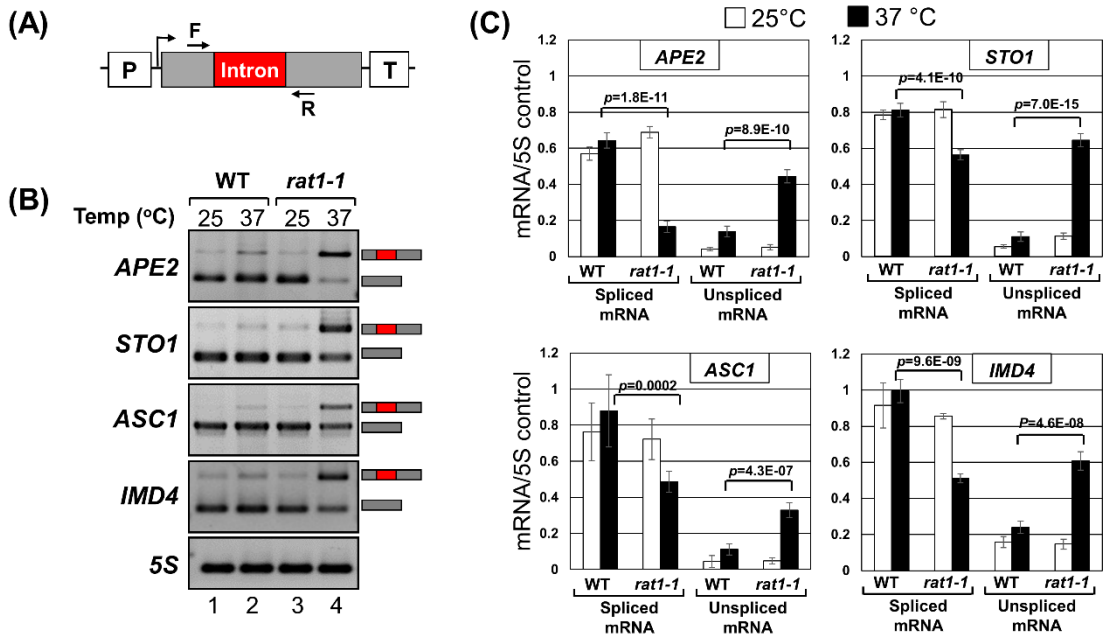

Figure 2. 1 Accumulation of unspliced transcripts in *rat1-1* mutant at non-permissive temperature. (A) Schematic depiction of a gene showing the position of primers F and R used in RT-PCR analysis. (B) Gel pictures showing RT-PCR products for the indicated genes in wild type (WT) and *rat1-1* mutant cells at the indicated temperatures. (C) Quantification of data shown in (B). 5S rRNA was used as normalization control.

result with the *APE2* gene. A longer transcript appeared upon shifting of mutant cells to the non-permissive temperature (Figure 2.1 B, lane 4). *APE2* is one of the few intron containing genes in budding yeast, and sequencing of PCR product revealed that the longer transcript was unspliced mRNA. To determine if the accumulation of unspliced transcripts in the Rat1 mutant is unique to *APE2* or if other intron containing genes exhibit a similar phenotype, we repeated the experiment with other intronic genes. Of all the

genes that we examined, *STO1*, *ASC1* and *IMD4* exhibited a similar trend of increased unspliced mRNA content in the *rat1-1* mutant at the elevated temperature (Figure 2.1 B, lane 4 and Figure 2.1 C, black bars). No such accumulation of unspliced transcripts was observed in the isogenic wild type cells at 37°C (Figure 2.1 B, lane 2 and Figure 2.1 C). Not all intron containing genes, however, exhibit increased unspliced transcript level in the mutant (data not shown).

To examine if accumulation of unspliced transcripts of a subset of yeast genes in the *rat1-1* strain was due to the specific inactivation of Rat1 and not due to a secondary mutation in another factor, we used the 'anchor away' approach. This technique involves selective depletion of a protein from the nucleus by anchoring it to ribosomes in the cytoplasm (Haruki et al., 2008). The approach takes advantage of the rapamycin-dependent heterodimerization of FK506 binding protein (FKBP12) with FKBP12-rapamycin-binding (FRB) domain (Belshaw et al., 1996; Chen et al., 1995). We inserted the FRB domain at the carboxy-terminus of Rat1 in a strain, which has FKBP12 fused to the ribosomal protein RPL13A, *fpr1* deletion and *tor1-1* mutation. The resultant strain was named *Rat1-AA*. In the presence of rapamycin, Rat1-FRB dimerized with ribosomal RPL13A-FKBP12 in the anchor away strain, leading to almost complete depletion of Rat1 from the nucleus within 60 minutes of addition of the antibiotic to the medium (Figure 2.2 A). We monitored levels of spliced and unspliced transcripts in the *Rat1-AA* strain in the presence of rapamycin (no Rat1 in the nucleus) and in the absence of rapamycin (Rat1 present in the nucleus) by RT-PCR as described above. We observed a 2-4-fold increase in the level of unspliced *APE2, STO1, MRK1, NMD2* and *LSB4* transcripts in the presence of rapamycin (Figure 2.2 C, lane 4 and Figure 2.2 D, red bars). There was no such

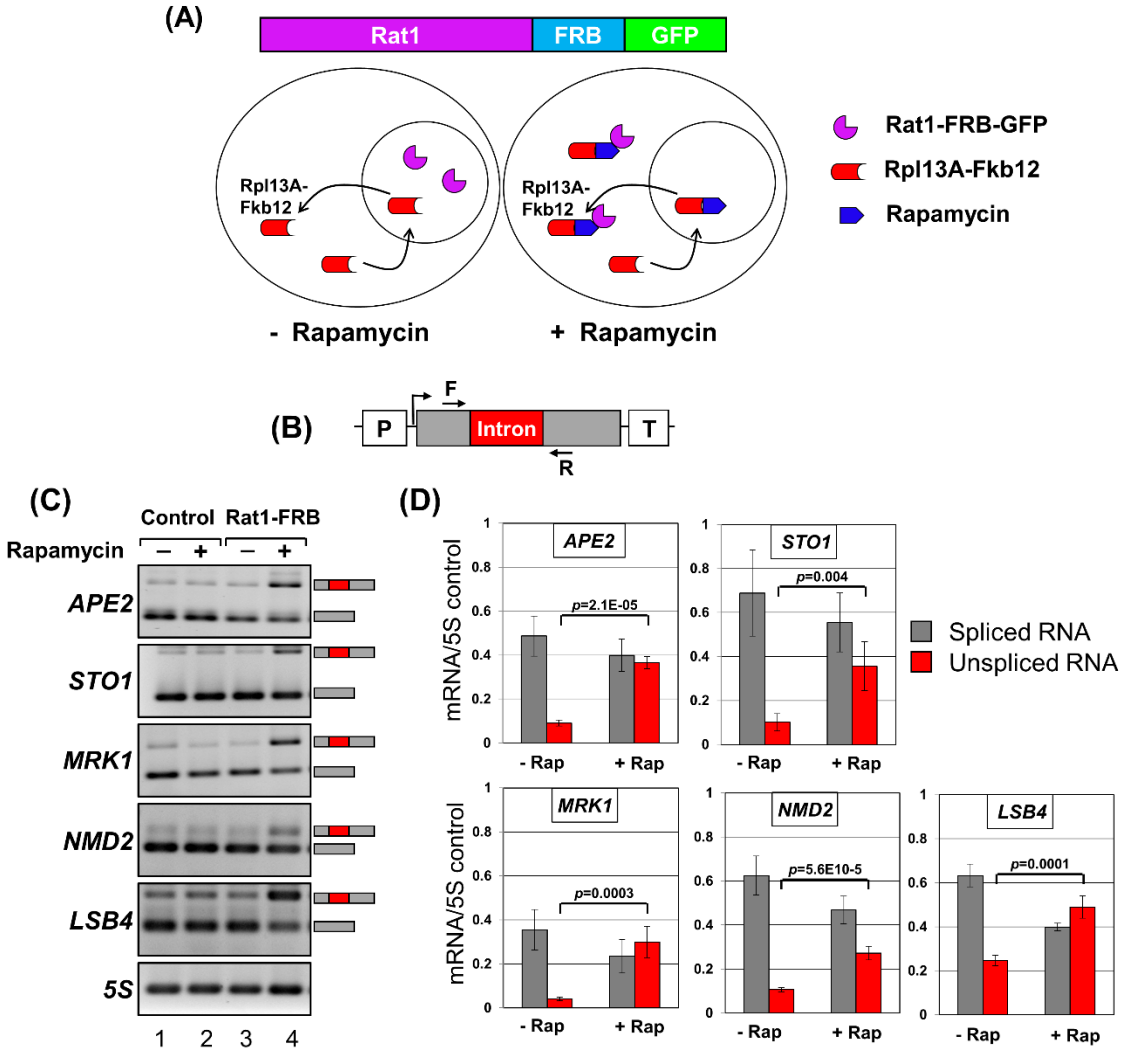

Figure 2. 2 Accumulation of unspliced transcripts upon nuclear depletion of FRB-tagged Rat1 in the presence of rapamycin. (A) Diagrammatic representation of anchor-away technique to selectively regulate compartmentalization of the nuclear protein Rat1 (concept adapted from Haruki et.al., 2008). (B) Schematic depiction of a gene showing the position of primers F and R used in RT-PCR analysis. (C) Gel pictures showing RT-PCR products for the indicated genes in FRB-tagged Rat1 strain and untagged strain (control) in the presence and absence of rapamycin. (D) Quantification of data in lanes 3 and 4 of (C). 5S rRNA was used as normalization control.

rapamycin-dependent buildup of unspliced mRNA in the isogenic wild type strain (Figure 2.2 C, lane 2 and Figure 2.2 D, red bars). The overall conclusion of these results is that there is an accumulation of unspliced transcripts in both the *rat1-1* strain as well as the *Rat1-AA* strain upon inactivation or nuclear depletion of Rat1. All intron-containing genes of budding yeast, however, do not exhibit accrual of unspliced RNA in the absence of

Rat1. Out of ten genes that we tested, only five exhibited accumulation of unspliced transcripts in the absence of Rat1.

To explain the presence of unspliced mRNA in the absence of Rat1 activity, we formulated four hypotheses:

Hypothesis #1: Rat1 plays a surveillance role and degrades unspliced transcripts by its exoribonuclease activity.

Hypothesis #2: Accumulation of unspliced transcripts is the consequence of an indirect effect of Rat1-mediated termination on splicing.

Hypothesis #3: Rat1 effect on elongation facilitates pausing of the polymerase during transcription of intronic regions to enable completion of splicing reactions.

Hypothesis #4: Rat1 has a novel role in splicing of primary transcripts.

We tested each of these hypotheses and the results are described below.

2.3.2 Rat1 is not the surveillance factor that degrades unspliced transcripts

The experiments described above monitored steady-state RNA levels, which is the net product of two opposing processes, transcription and degradation. To determine if Rat1 has a surveillance role in degrading unspliced mRNA by its exoribonuclease activity, we measured levels of unspliced and spliced transcripts by the transcription run-on (TRO) approach. The TRO assay detects nascent transcripts and therefore reflects transcription and rules out RNA stability making any contribution to the measured RNA content. The strand-specific TRO analysis was carried out using Br-UTP by the modification of the method described in Core et al., (2008). Briefly, exponentially growing yeast cells were permeabilized with sarkosyl, and allowed to resume transcription in the presence of Br-

labelled UTP for 2-5 minutes. Br-UTP-labelled nascent RNA was immunopurified, reverse transcribed using oligo-dT and amplified by PCR as described in (Dhoondia et al., 2017). TRO analysis revealed a 2-7 fold increase in the levels of unspliced transcripts of *APE2*, *STO1*, *MRK1* and *LSB4* in the presence of rapamycin compared to the absence of rapamycin in *Rat1-AA* strain (Figure 2.3 B and C).

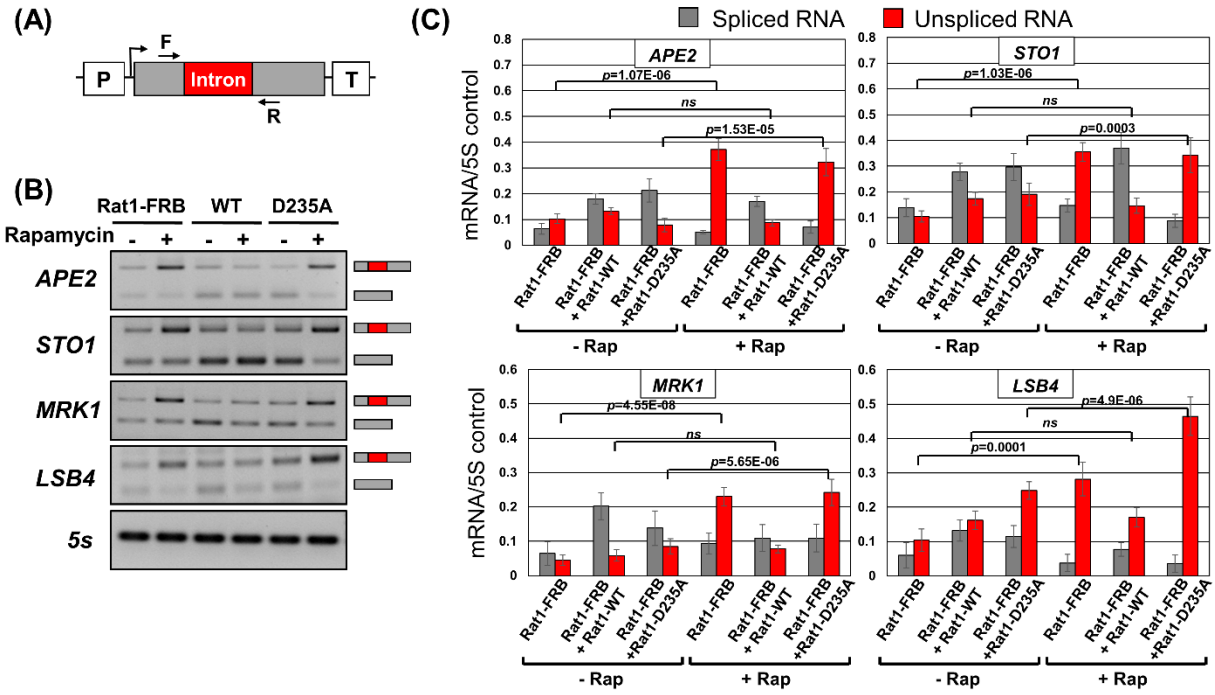

Figure 2. 3 Transcription Run-On (TRO) analysis showing that accumulation of nascent unspliced transcripts upon nuclear depletion of FRB-tagged Rat1 in the presence of rapamycin is rescued by wild type Rat1 but not the catalytically inactive mutant. (A) Schematic depiction of a gene showing the position of primers F and R used in TRO analysis. (B) Quantification of TRO signals for the indicated genes in FRB-tagged Rat1 strain in the presence and absence of rapamycin and upon complementation with WT Rat1 or catalytically inactive D235A Rat1 mutant. 5S rRNA was used as normalization control.

Since splicing occurs cotranscriptionally in both yeast and higher eukaryotes, these results suggest that Rat1 may be affecting cotranscriptional splicing of mRNA. The possibility of Rat1 playing a role in cotranscriptional degradation of unspliced transcripts, however, could not be ruled out. To find the answer, we transformed the *Rat1-AA* strain with a plasmid expressing either wild type Rat1 or a mutant form of Rat1 containing a

point mutation at an evolutionarily conserved catalytic residue (D235A). Denoted as exo-Rat1, the mutant is deficient in exoribonuclease activity (Kim et al., 2004b). TRO was repeated in the presence and absence of rapamycin. Expression of the wild type Rat1 rescued the accumulation of unspliced transcript in the presence of rapamycin, but catalytically inactive exo-Rat1 mutant was unable to do so (Figure 2.3 B and C). One of the interpretations of this result is that Rat1 catalytic activity is required for degradation of unspliced transcripts, and therefore there is a buildup of unspliced mRNA in the exo-Rat1 mutant.

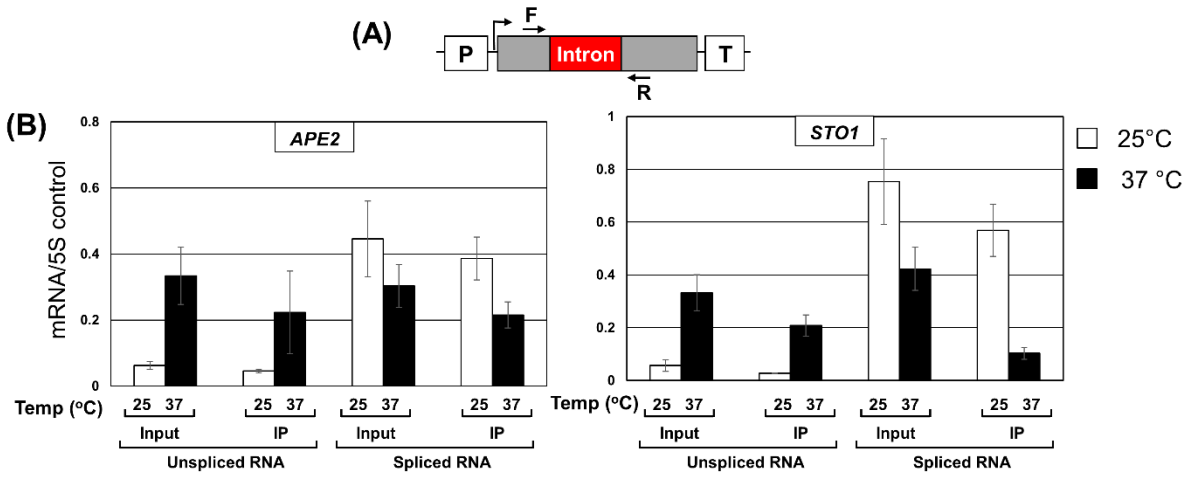

Figure 2. 4 RNA immunoprecipitation analysis shows that unspliced transcripts in *rat1-1* mutant are capped. (A) Schematic depiction of a gene showing the position of primers F and R used in RT-PCR analysis. (B) Affinity purified RNA using anti-m7G were reverse transcribed. Quantification of data shown in for the indicated genes in wild type (WT) and *rat1-1* mutant at the indicated temperatures. 5S rRNA was used as normalization control.

Rat1 is highly specific in choosing its RNA substrate. It can degrade only 5′ monophosphorylated RNA by its 5′→3′ exoribonuclease activity. 7' methylguanosine capped mRNA is not a suitable substrate of the Rat1 nuclease. Thus, Rat1 will be able to degrade unspliced mRNA only if it is somehow uncapped. We therefore examined the capping pattern of unspliced transcripts. RNA was isolated from mutant *rat1-1* cells grown at 25°C and 37°C and subjected to immunoprecipitation (RNA-IP) using antibodies

against 7’methylguanosine cap. Immunoprecipitated RNA was subjected to RT-PCR for *APE2* and *STO1* genes as described previously. RNA level was normalized by 5S control. As expected, there was an approximate 7-fold increase in the amount of unspliced mRNA of *APE2* and *STO1* at 37°C compared to 25°C (Figure 2.4 B). About 70 to 80% of both spliced and unspliced transcripts of *APE2* and *STO1* could be immunoprecipitated by anti-7’methylguanosine antibodies (Figure 2.4 B). These results clearly demonstrate that the degree of capping of unspliced mRNA is nearly to the same extent as that of spliced mRNA, and therefore unspliced transcripts could not be the substrate of Rat1 enzymatic activity. The overall conclusion of these results is that the accumulation of unspliced mRNA in the absence of Rat1 activity in yeast cells is not due to the surveillance role of Rat1 in degrading aberrant intron containing transcripts.

2.3.3 Rat1-mediated increase in the levels of unspliced transcripts is independent of Xrn1 and the exosome

The degradation of RNA relies on two well-conserved pathways in yeast, which involves both 5′→3′ and 3'→5' exoribonuclease activities by Xrn1 and Rrp6, respectively (Parker, 2012). Like Rat1, Xrn1 is a 5′→3′ exoribonuclease in budding yeast, but it is located in the cytoplasm where it functions to degrade the aberrant transcripts. Strikingly, both Rat1 and Xrn1 have the tendency to use RNA substrates with a 5'-monophosphate, and they are inactive towards RNA with a 5'-methylguanosine cap or a 5'-triphosphate, thus preventing the degradation of cellular RNA. Xrn1 is capable of rescuing the restrictive growth phenotype of *rat1-1* cells, but cannot complement the termination defect in the mutant (Luo et al., 2006). Therefore, to test whether Xrn1 shares the same substrate as Rat1 in influencing the levels of unspliced transcripts, we constructed *xrn1Δ* Rat1-AA

strain and determined the steady state mRNA levels for *APE2, STO1, NMD2, MRK1* and *LSB4* genes by RT-PCR analysis. The results indicate that deletion of Xrn1 does not influence the unspliced transcript levels (Figure 2.5), thereby suggesting that the increased level of unspliced transcripts is influenced only by Rat1.

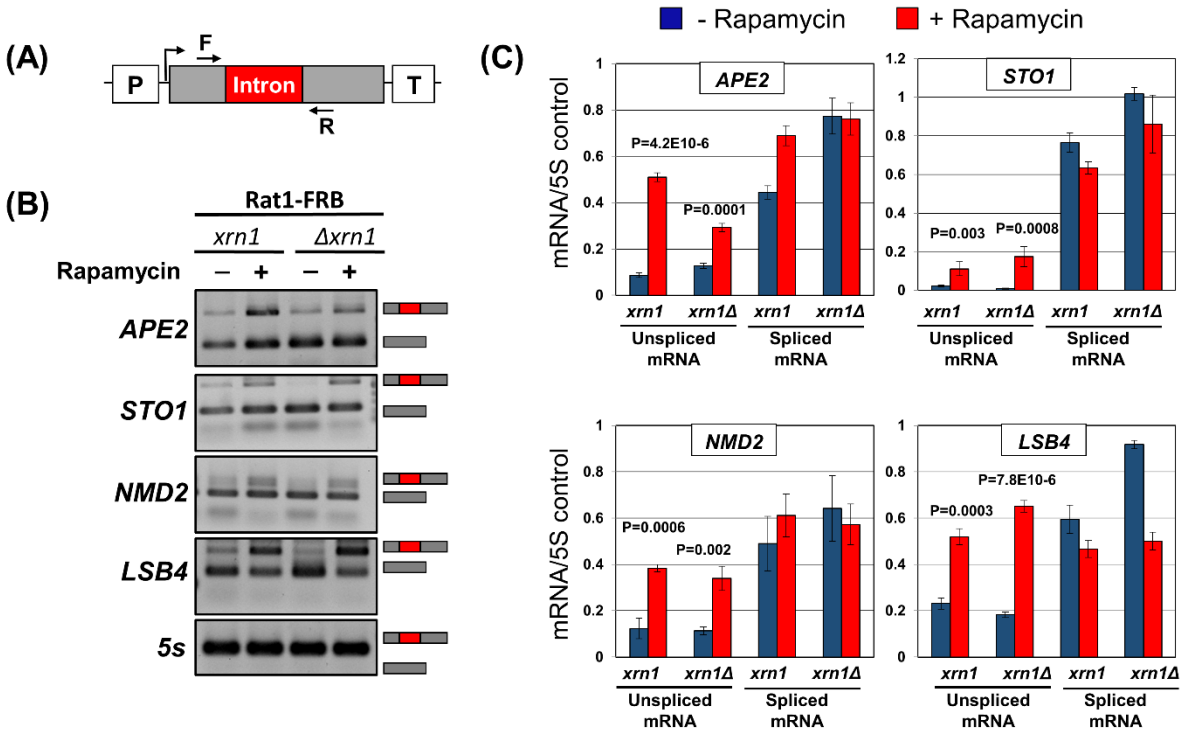

Figure 2. 5 Xrn1 deletion does not affect the level of unspliced transcript under nuclear depletion of Rat1. (A) Schematic depiction of a gene showing the position of primers F and R used in RT-PCR analysis. (C) Gel pictures showing RT-PCR products for the indicated genes in FRB-tagged Rat1 strain with and without Xrn1 in the presence and absence of rapamycin. (D) Quantification of data shown in (B). 5S rRNA was used as normalization control.

Many aberrant transcripts are degraded by nuclear exosomes in which the catalytic subunit Rrp6 rapidly eliminates them by its 3′→5′ exoribonuclease activity. We therefore examined if Rat1 has an overlapping function with Rrp6 in degrading unspliced transcripts. RT-PCR analysis shows that upon nuclear depletion of Rat1, *rrp6Δ* did not show any significant change in the levels of the unspliced transcripts of *APE2, STO1, NMD2, MRK1* and *LSB4* genes (Figure 2.6). These results clearly demonstrate that Xrn1

and rrp6 exoribonuclease activity is not involved in any way in degrading unspliced transcripts in budding yeast.

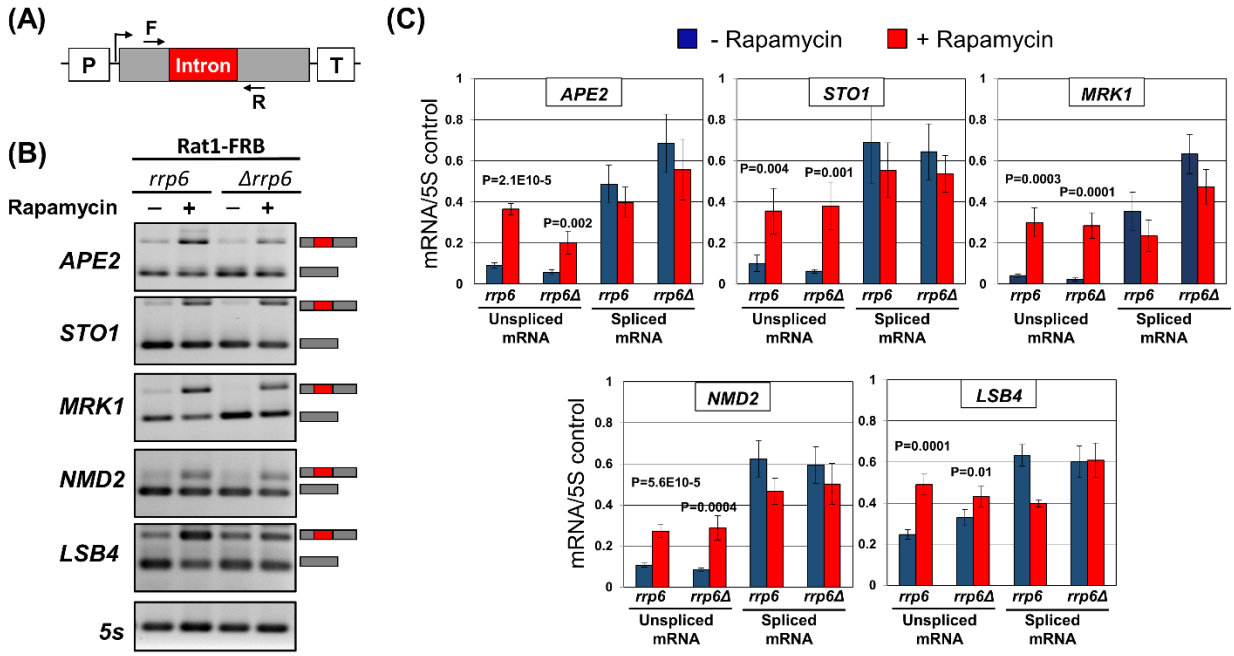

Figure 2. 6 Rrp6 deletion does not affect the level of unspliced transcript under nuclear depletion of Rat1. (A) Schematic depiction of a gene showing the position of primers F and R used in RT-PCR analysis. (C) Gel pictures showing RT-PCR products for the indicated genes in FRB-tagged Rat1 strain with and without Xrn1 in the presence and absence of rapamycin. (D) Quantification of data shown in (B). 5S rRNA was used as normalization control.

2.3.4 Accumulation of unspliced mRNA in the absence of Rat1 is not the consequence of defective termination

Rat1 and its higher eukaryotic homolog Xrn2 are transcription termination factors. We reasoned that if appearance of unspliced transcripts in the absence of Rat1 activity is the consequence of an indirect effect of defective termination on splicing, other termination-defective mutants may also display splicing defect.

Apart from Rat1, termination of transcription by RNAPII in yeast requires CF1 and CPF complexes. Both CF1 and CPF complexes are required for cleavage and polyadenylation of the messenger. The cleavage-polyadenylation of primary transcripts

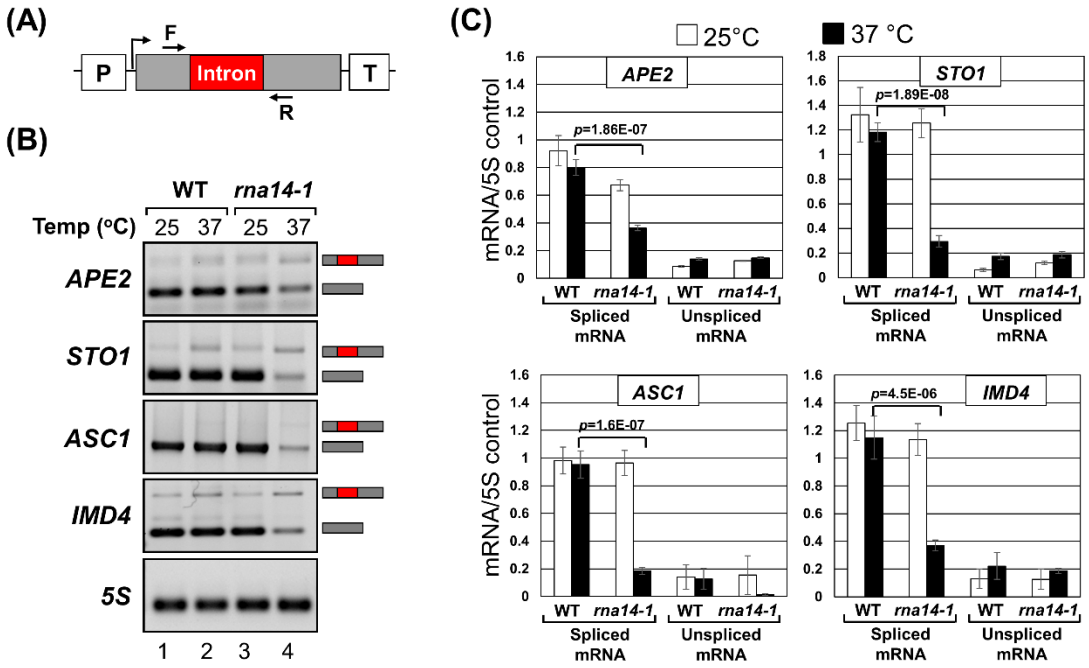

Figure 2. 7 Defective termination in *rna14-1* mutant does not result in accumulation of unspliced transcripts. (A) Schematic depiction of a gene showing the position of primers F and R used in RT-PCR analysis. (B) Gel pictures showing RT-PCR products for the indicated genes in wild type (WT) and *rna14-1* mutant at the indicated temperatures. (C) Quantification of data shown in (B). 5S rRNA was used as normalization control.

is coupled to poly(A)-dependent termination of transcription. A recent genomewide analysis revealed Rat1 being required for termination of transcription of nearly 35% of genes in budding yeast. In contrast, CF1 and CPF subunits have a more robust role in termination as they affect poly(A)-dependent termination of nearly 90% of yeast genes (Baejen et al., 2017). We therefore monitored levels of spliced and unspliced mRNA of *APE2*, *STO1*, *ASC1* and *IMD4* in the termination-defective *rna14-1* mutant. Rna14 is a subunit of CF1 complex and is essential for poly(A)-dependent termination of transcription (Minvielle-Sebastia et al., 1997). There was a decrease in the RNA level of *APE2*, *STO1*, *ASC1* and *IMD4* upon shifting of mutant cells to the elevated temperature (Figure 2.7 B, lane 4 and Figure 2.7 C, black bars), in agreement with the role of termination for optimal transcription. There was, however, no increase in the level of unspliced transcripts at the

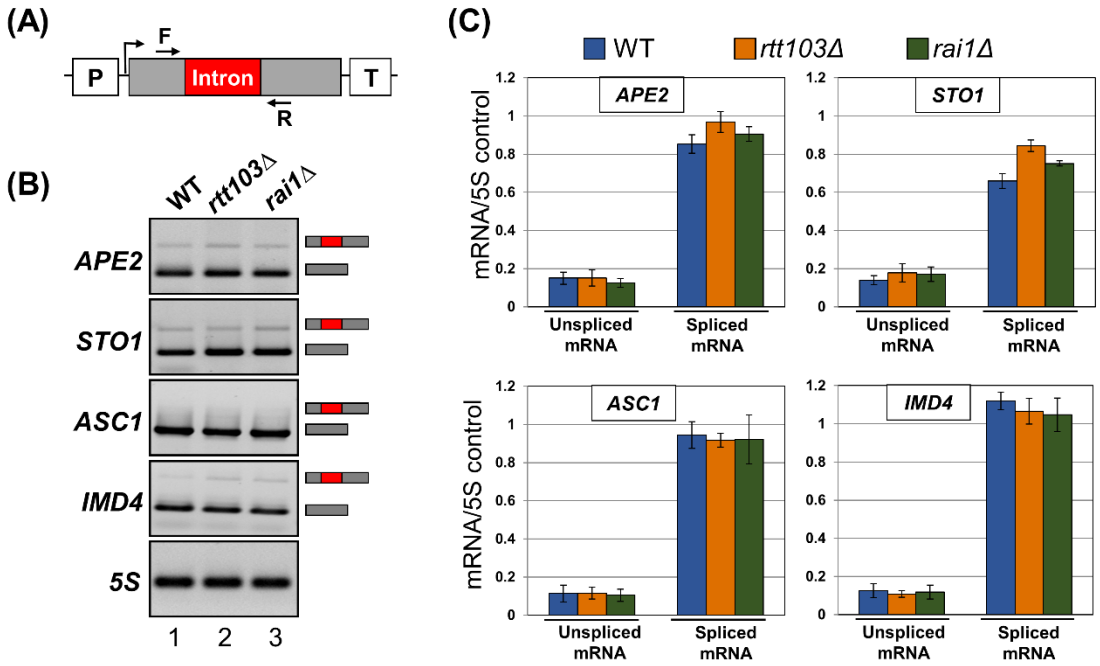

Figure 2. 8 Rat1 termination complex is not required for accumulation of unspliced transcripts. (A) Schematic depiction of a gene showing the position of primers F and R used in RT-PCR analysis. (B) Gel pictures showing RT-PCR products for the indicated genes in wild type (WT), *rtt103Δ* and *rai1Δ* strains. (C) Quantification of data shown in (B). 5S rRNA was used as normalization control.

non-permissive temperature (Figure 2.7 B, lane 4 and Figure 2.7 C, black bars). The amount of unspliced mRNA in the mutant at 37°C was similar to that in the isogenic wild type strain (Figure 2.7 B, lanes 2 and 4 and Figure 2.7 C, black bars). The defective termination in the Rna14 mutant therefore did not have an adverse effect on splicing. A corollary of these results is that the accumulation of unspliced mRNA in the mutants of Rat1 may not be the consequence of defective termination on splicing. In budding yeast, Rat1 is in a trimeric complex with Rai1 and Rtt103 (Kim et al., 2004b). Rai1 has pyrophosphohydralase activity and stimulates exonucleolytic activity of Rat1 (Xiang et al., 2009). Since Rat1 enzymatic activity is essential for its termination function, Rai1 has been found to affect Rat1-dependent termination in yeast to some extent (Kim et al., 2004b). Rai1, however, is not a termination factor per se. Although Rtt103 is not required for termination *in vivo*, it physically interacts with Rat1, and has been implicated in

termination of transcription *in vitro* under certain condition (Kim et al., 2004b; Pearson and Moore, 2013). We reasoned that if Rat1 exonuclease-dependent termination is affecting splicing, Rtt103 and Rai1 may have an influence on splicing. We therefore examined levels of spliced and unspliced transcripts in deletion mutants of Rai1 (*rai1Δ*) and Rtt103 (*rtt103Δ*) cells by RT-PCR as described above. There was no increase in the amount of unspliced transcripts of *APE2, STO1, ASC1* and *IMD4* in *rai1Δ* strain (Figure 2.8 B, lane 3, Figure 2.8 C, green bars) over isogenic wild type control (Figure 2.8 B, lane 1, Figure 2.8 C, blue bars). Similar results were obtained with *rtt103Δ* cells (Figure 2.8 B, lane 2, Figure 2.8 C orange bars).

These results strongly suggest the accumulation of unspliced transcripts in the mutants of Rat1 may not be an indirect consequence of Rat1-dependent termination defect. Failure of the nuclease-deficient exo-Rat1 mutant to rescue the splicing defective phenotype of Rat1 mutant, however, does not completely rule out the possibility that Rat1-dependent termination is somehow indirectly effecting splicing.

2.3.5 Is Rat1 facilitated pausing of polymerase during elongation of intronic regions to enable the recruitment of splicing factors?

Splicing of most introns occurs cotranscriptionally. During cotranscriptional splicing, the polymerase slows down or pauses while transcribing intronic regions. This may enable recruitment of splicing factors on splice sites leading to efficient execution of the splicing reaction. Consequently, the increased elongation rate, associated with some mutants of RNAPII, contributes to splicing defects (Braberg et al., 2013). The *rat1-1* mutant also exhibits enhanced RNAPII elongation rate by stimulating hyper-phosphorylation of CTD-serine-2 (Jimeno-González et al., 2014). The possibility of the

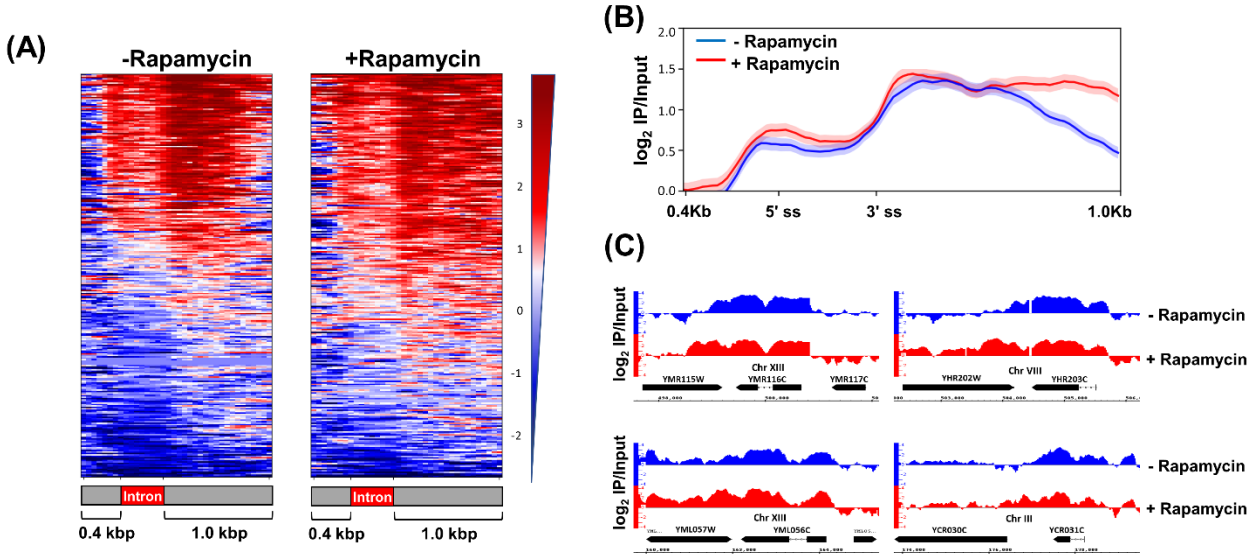

Figure 2. 9 RNAPII levels are similar across the intronic genes after nuclear depletion of Rat1. (A) Heatmaps of Rpb1 occupancy levels calculated as the $\log_2$ IP/Input ratio in Rat1-FRB strain before (-rapamycin) and after (+rapamycin) nuclear depletion of Rat1 from ChIP-seq experiments performed in Baejen et.al., 2017. Genes are sorted in the descending order of Rpb1 occupancy. The y- axis indicates individual genes from the 280 intron containing genes and the x-axis shows relative position of 0.4 Kbp upstream and 1.0 Kbp downstream around 5' ss and 3' ss intronic sites. (B) Averaged normalized occupancy profiles from the ChIP-seq data of RNAPII. Metagene plot was designed with intron size averaged to 300 bp and plotting 0.4 Kbp upstream and 1Kbp downstream. The profiles are aligned at both the 5' ss and 3' ss of 280 intron containing genes. The shaded areas around the ChIP-seq profile represent standard deviation ±1. (C) IGB browser view of $\log_2$ IP/Input Rpb1 (RNAPII) read counts from ChIP-seq experiment performed in Rat1-FRB tagged strain before (blue) and after (red) nuclear depletion of Rat1. Four representative ORFs shown *YMR116C, YHR203C, YML056C, YCR031C*.

increased elongation rate contributing to the splicing defect observed in *rat1-1* mutant cannot be ruled out.

We therefore examined if there is decreased pausing of polymerase on the intron in the absence of Rat1 activity. We used Rpb1-ChIP-Seq data from Baejen et al., (2017) to test our hypothesis. This study analyzed genomewide transcription using both 4tU-Seq and Rpb1-ChIP-Seq approach in *Rat1-AA* strain in the presence (Rat1 depleted from the nucleus) and absence of rapamycin (Rat1 present in the nucleus). We extracted Rpb1-ChIP-Seq data for intron containing genes from this study. There are nearly 301 intron containing genes in budding yeast. We extracted data for 280 of these genes. Rpb1-

ChIP-Seq reads were aligned with respect to the 5' and 3' splice sites of introns, encompassing a 400 bp window upstream of the 5' splice sites and a 1 kbp window downstream of the 3' splice sites as shown in Figure 2.9. In the presence of a functional Rat1 in the nucleus (-rapamycin), there was increased polymerase signal over the introns compared to the 400 bp upstream region. There was, however, no decrease in polymerase density over the intron upon nuclear depletion of Rat1 (+rapamycin) (Figure 2.9 A and Figure 2.9 B). Rpb1-ChIP profile of individual genes confirms that Rpb1 signal over the intron and splice sites is similar in the presence and absence of Rat1 in the nucleus (Figure 2.9 C). As expected, there was an increased RNAPII read-through signal in the downstream 1 kbp region under nuclear depletion of Rat1 due to the termination defect. These results indicate that accumulation of unspliced transcripts in Rat1 mutants is not an indirect consequence of loss of pausing of the polymerase over intronic regions due to faster elongation rate in the Rat1 mutant.

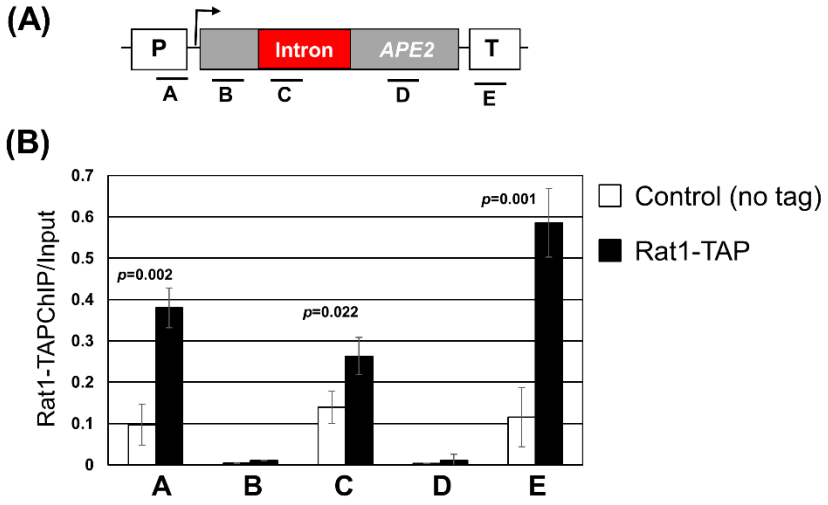

Figure 2. 10 Crosslinking of Rat1 to intron of *APE2* gene. (A) Schematic depiction of *APE2* indicating the position of ChIP primer pairs A, B, C, D, and E. (B) Quantitative analysis showing crosslinking of Rat1 to different regions of *APE2* in Rat1-TAP and WT (no tag) strains represented using black and white bars, respectively. The input signal represents DNA prior to immunoprecipitation.

2.3.6 Rat1 has a novel role in splicing of mRNA

The results described above could not provide a convincing explanation for accumulation of unspliced transcripts in the absence of functional Rat1 in the cell. An alternative possibility is that Rat1 has a novel role in splicing of primary transcripts. Therefore, we examined the direct involvement of Rat1 in splicing by employing three parallel approaches. We reasoned that if Rat1 is involved in splicing, (1) it will physically interact with introns; (2) it will exhibit a transient or stable interaction with the splicing factors; and (3) it will facilitate the recruitment of splicing factors on the intron.

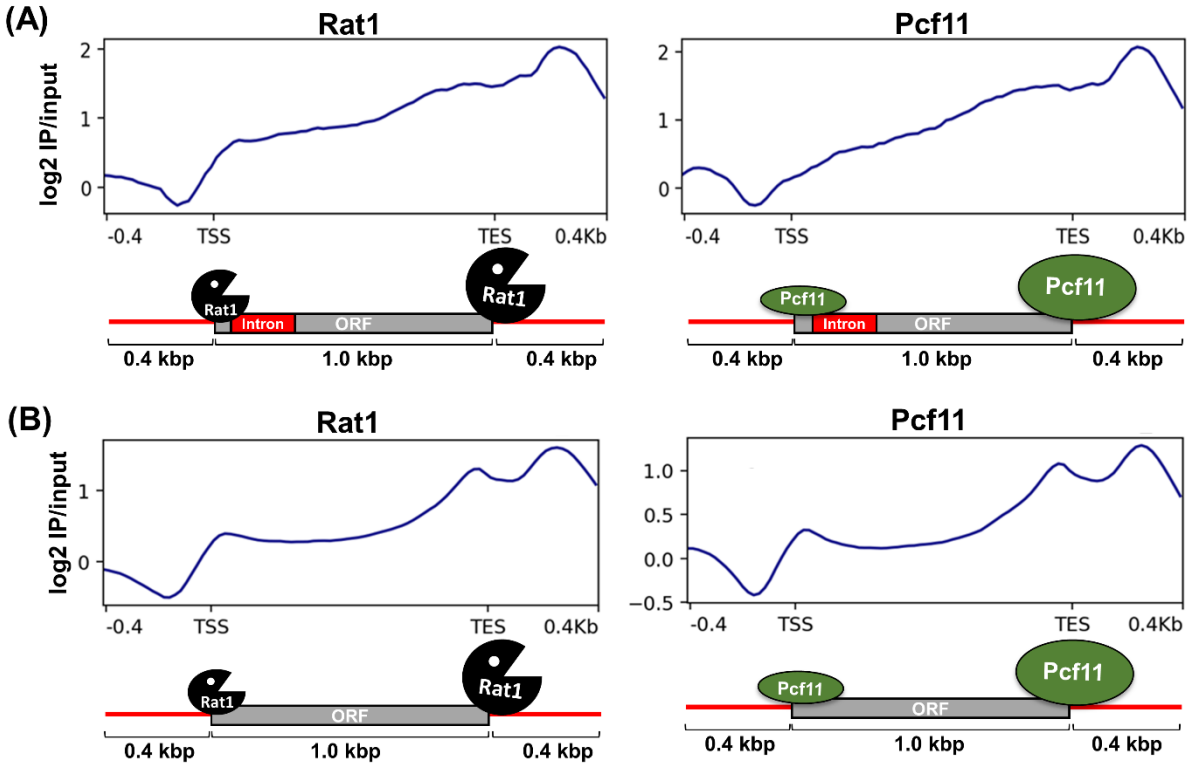

Figure 2. 11 Rat1 occupancy profile differs from Pcf11 at the intronic genes and not at the non-intronic genes. Average normalized plot of Rat1 and Pcf11 occupancy in WT strain, mapped to 2392 genes that are non-intronic (A) and 280 genes that are intronic. The metagene plots are are constructed with the coordinates of transcription start site (TSS) and transcription end site (TES) and distance of the open reading frame (ORF) averaged to 1 Kbp.

2.3.7 Rat1 crosslinks to introns

To determine if Rat1 contact the splice sites on an intron, we inserted a TAP-tag at the carboxy-terminus of Rat1. TAP-tag does not interfere with the termination function of Rat1 (data not presented). High resolution ChIP that employs extensive sonication was performed as described in El Kaderi et al., (2009). Splicing factors do not directly crosslink to the splice sites on DNA but do so indirectly through their interaction with the elongating transcript and CTD of RNAPII during cotranscriptional splicing (Herzel et al., 2017; Nojima et al., 2018). We used a more robust crosslinking approach, disuccinimidyl glutarate (DSG) together with formaldehyde, to detect indirect interactions of splicing factors on splice sites of DNA as described in GRID-Seq approach (Li et al., 2017). Our results show that Rat1 crosslinks to the promoter and terminator regions of *APE2* (Figure 2.10, primers pairs A and E) in accordance with the published results (Baejen et al., 2017; Kim et al., 2004b). More importantly, Rat1 also localized to the intronic splice sites of *APE2* (Figure 2.10, primer pair C). The crosslinking of Rat1 to the splice sites suggest that Rat1 may have a direct role in splicing of *APE2* mRNA.

We next examined if other intron containing genes similarly exhibit crosslinking of Rat1 to the intron. We used ChIP-Seq data from Baejen et.al., (2017) for this analysis. In this study, ChIP-Seq data for transcription factors was examined using 2501 genes. Of these, 109 genes were intron-containing genes whereas remaining 2392 genes were non-intronic. Therefore, to examine occupancy of Rat1 selectively on intronic genes, we selected 280 genes in our analysis that are characterized as intron-containing in Saccharomyces Genome Database (SGD). Rat1 occupancy examination on these 280 intron-containing genes, aligned to TSS and TES ± 400 bp, showed that Rat1 is localized

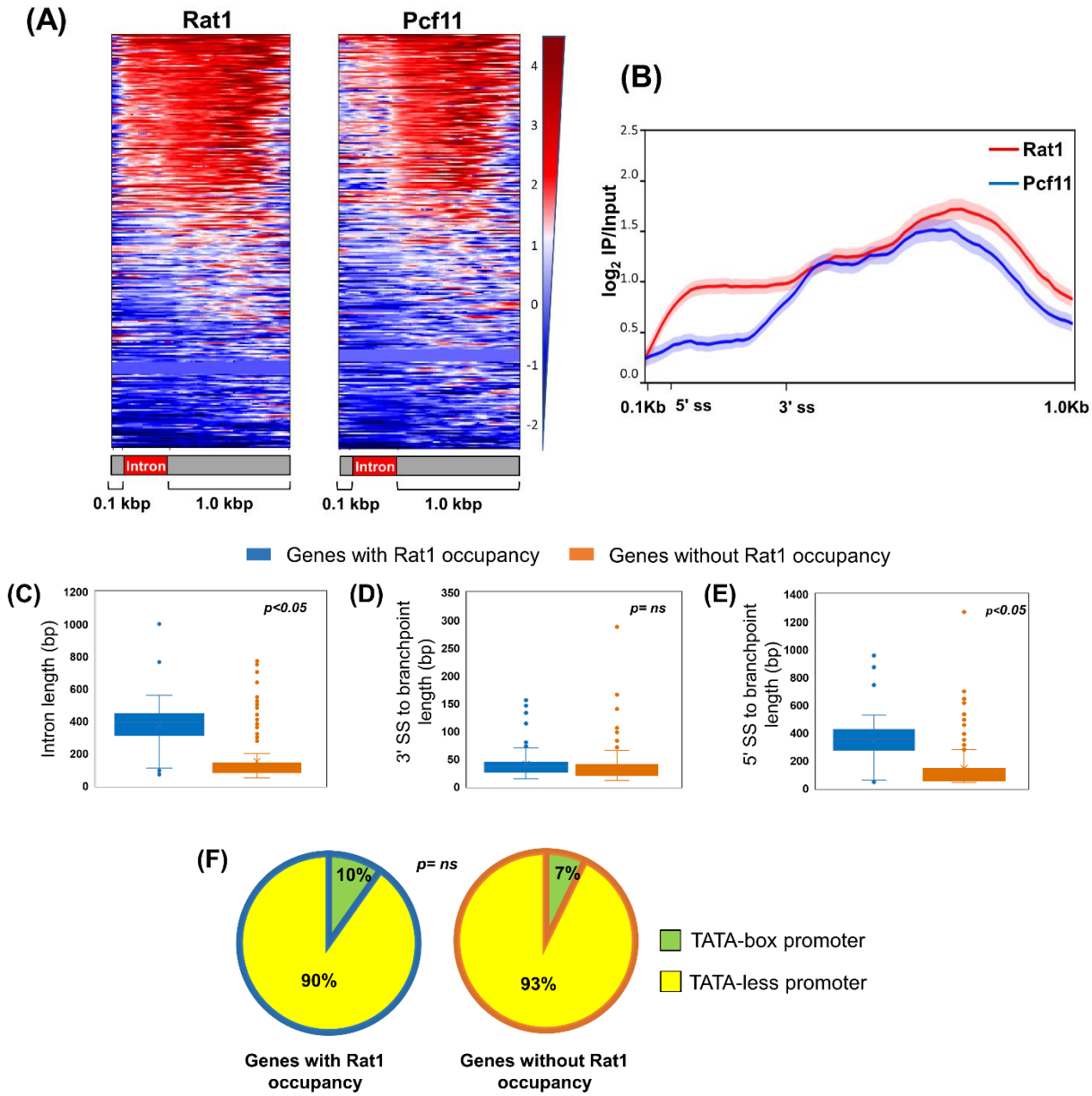

Figure 2. 12 Rat1 is selectively enriched in the intronic region of the genes. (A) Heatmaps of Rat1 and Pcf11 occupancy levels calculated as the $\log_2$ IP/Input ratio from ChIP-Seq experiments performed in Rat1-TAP and Pcf11-TAP strains, respectively. Genes are sorted in the descending order of the occupancy. The y- axis indicates individual genes from the 280 intron containing genes and the x-axis shows relative position of 0.1 Kbp upstream and 1.0 Kbp downstream around 5' ss and 3' ss intronic sites. (B) Averaged normalized occupancy profiles from the ChIP-Seq data of Rat1 (red) and Pcf11 (blue). The metagene plot was designed with an average intron size of 300 bp and plotting 0.1 Kbp upstream and 1Kbp downstream. The profiles are aligned at both the 5' ss and 3' ss of 280 intron containing genes. The shaded areas around the ChIP-Seq profile represent standard deviation ±1. Genes with the presence and absence of Rat1 on intron are sorted using box-whisker plot according to intron length (C), distance between 3' SS to branchpoint (D), distance between 5' SS to branchpoint (E), and the promoter element TATA-box and TATA-less (F). Mann-Whitney U test was used to calculate *p value* for C, D, and E. Fisher's exact test was used to calculate *p value* for F.

at the terminator and promoter regions of genes (Figure 2.11 A). It is possible that it is not

merely Rat1 but all 3' end processing-termination factors that display similar crosslinking to the intron. We therefore analyzed Pcf11 occupancy on introns using the data from the same genomewide study. Pcf11 and Rat1 occupancy profile were comparable at the termination region. They, however, differed considerably near the promoter region. Rat1 occupancy was significantly higher than that of Pcf11 just downstream of the promoter

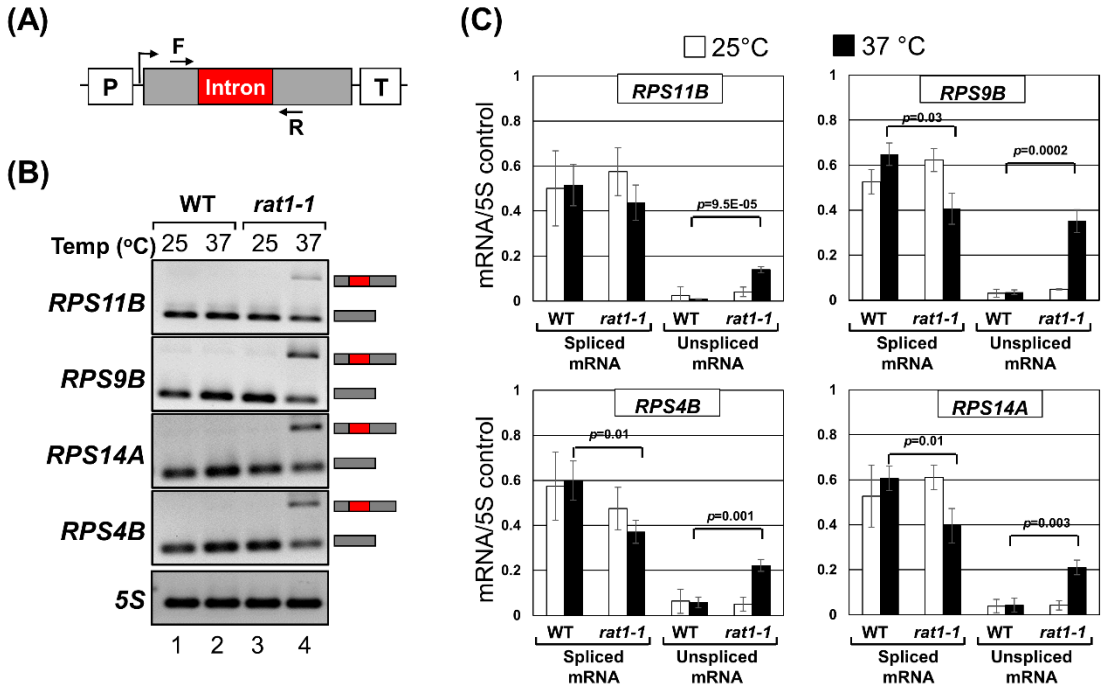

Figure 2. 13 Rat1 intronic occupancy correlates with Rat1-dependent accumulation of unspliced transcripts in *rat1-1* mutant at non-permissive temperature. (A) Schematic depiction of a gene showing the position of primers F and R used in RT-PCR analysis. (B) Gel pictures showing RT-PCR products for the indicated genes in wild type (WT) and *rat1-1* mutant cells at the indicated temperatures. (C) Quantification of data shown in (B). 5S rRNA was used as normalization control.

region (Figure 2.11 A). It is known that most introns in budding yeast are positioned near the TSS and thus, this increase in the chromatin occupancy of Rat1 over Pcf11 near TSS can possibly be due to its localization to the intronic region. Further, Rat1 and Pcf11 occupancy profile of 2392 non-intronic was almost identical (Figure 2.11 B). These results indicate that the Rat1 is preferentially enriched at the TSS, and this enrichment may be due to Rat1 crosslinking to the promoter-proximal intron.

Next, we extracted Rat1-ChIP-Seq data for intron containing genes from this study. Rat1-ChIP-Seq reads were aligned with respect to the 5' and 3' splice sites of introns, encompassing a 100 bp window upstream of the 5' splice sites and a 1 kbp window downstream of the 3' splice sites as shown in Figure 2.12. Out of 280 intron containing genes analyzed in this study, 105 exhibited crosslinking of Rat1 to the intronic sequence. Thus, not all intron containing genes show Rat1 occupancy. Unlike Rat1, Pcf11 does not exhibit significant localization to the intronic region (Figure 2.12, A and B). These results indicate that Rat1 localizes to the introns of a subset of genes in budding yeast. We randomly selected ten intron containing genes from the group of 105 genes that exhibited Rat1-intron occupancy to determine if Rat1 intronic occupancy correlated with Rat1-dependence on splicing. Out of these ten genes, seven genes were dependent on Rat1 for their efficient splicing (Figure 2.13). Thus, not all the genes that exhibit Rat1-intron occupancy are dependent on Rat1 for splicing.

Next we searched among the 105 Rat1-occupied introns for common features, which conferred potential dependence on Rat1 for their splicing. Compared to other introns, Rat1-occupied introns do not display any common predicted structure or enriched sequence motif. Of all the features that we examined, the only one that differed was the size of the intron. On average, Rat1-crosslinked introns were ~400 bp long, while those not exhibiting Rat1 occupancy were ~100 bp in length (Figure 2.12 C). The distance between branchpoint and 3' splice site was the same for all introns (Figure 2.12 D). It was the distance between the 5' splice site and branchpoint that tended to be longer for Rat1-occupied introns (Figure 2.12 E). Furthermore, Rat1 intron occupancy was not correlated to the presence or absence of TATA-box in the promoter of the gene. (Figure 2.12 F).

Ontological analysis revealed that the Rat1-occupied introns were enriched for ribosomal protein genes.

2.3.8 Rat1 interacts with splicing factors

If Rat1 is involved in splicing, it may exhibit stable or transient interaction with the splicing machinery. We therefore examined the interaction of Rat1 with the splicing factors. To identify the splicing proteins that may interact with Rat1, we performed biochemical purification of Rat1 employing Tandem Affinity Purification (TAP) approach. Purification was performed as described in Puig et al., (2001), and double affinity purified fraction was subjected to mass spectrometry. A total of 1220 proteins were identified. Rai1 and Rtt103, as expected, were present in the purified fraction of Rat1. A significant finding was the identification of five splicing proteins, Clf1, Isy1, Yju2, Sub2, and Prp43 in the Rat1 preparation (Table1). Of these; Clf1, Isy1, and Yju2 are the subunits of NineTeen complex (NTC). Clf1 and Isy1 are characterized as a part of the core complex, whereas Yju2 subunit is an accessory protein. Clf1 is an essential splicing factor that serve as a scaffold in spliceosome assembly during assembly of the tri-snRNP (U4 U5.U6). It also interacts with Mud2 and Prp40 that are branchpoint binding proteins (Chung et al., 1999). Isy1 is not an essential splicing factor but contributes to the fidelity of splicing reaction (Villa and Guthrie, 2005). Yju2 is an essential splicing factor that functions after Prp2 to promote the first transesterification reaction (Liu et al., 2007). Sub2, a DEAD-box RNA helicase, is a known to collaborate with Msl5 that interacts with branchpoint binding protein to promote the recruitment of U2 snRNP (Zhang and Green, 2001). Like Sub2, Prp43 is an RNA helicase; however, it functions post splicing in disassembling the U2, U5, and U6 snRNPs (Arenas and Abelson, 1997). The splicing factors, however, were

not present in stoichiometric proportion with respect to Rat1. This data indicates that Rat1 is not in a stable complex with splicing factors, but transiently interacts with them during cotranscriptional splicing.

Protein _ID	Coverage (relative to Rat1)
Rat1	100%
Rai1	39.05%
Sub2	10.73%
Rtt103	3.0%
Isy1	0.86%
Prp43	2.15%
Yju2	0.43%
Clf1	0.86%

Table 1: Affinity purified fraction of Rat1 -TAP shows interaction between Rat1 and splicing factors. Splicing factors (highlighted in bold) and previous known Rat1- interactors are listed relative to the levels of Rat1.

2.3.9 Recruitment of splicing factor Prp2 is compromised in the absence of Rat1

Rat1 interaction with intronic sequences as well as its interaction with Isy1, Prp43, Clft1, Sub2 and Yju2 splicing factors strongly suggested the direct involvement of Rat1 in splicing of primary transcripts. To further probe the role of Rat1 in splicing, we monitored the recruitment of splicing factors on introns. We reasoned that if Rat1 is indeed playing a role in splicing, recruitment of at least some splicing factors on the intron will be

adversely affected in the absence of Rat1. We chose Prp2 for three reasons; (1) it exhibits a genetic interaction with Rat1; (2) it is recruited on the intron before Yju2 which is the Rat1 interacting factor identified above in our analysis; and (3) its recruitment on the intron can be detected by ChIP. We checked binding of Prp2 to the intron of *APE2* by ChIP in the *Rat1-AA* strain in the presence and absence of rapamycin. As expected, Prp2-ChIP signal could be detected on the intron of *APE2* in the absence of rapamycin (Figure 2.14, lane 2). In the presence of rapamycin, when Rat1 is depleted from the nucleus, Prp2-ChIP signal registered a nearly 30% decline. These results are in agreement with the observation that splicing of *APE2* intron is not completely dependent on Rat1 but decreases by about 30-40% in the presence of rapamycin. The results described above strongly suggest that Rat1 has a direct role in splicing. Furthermore, Rat1 is not an

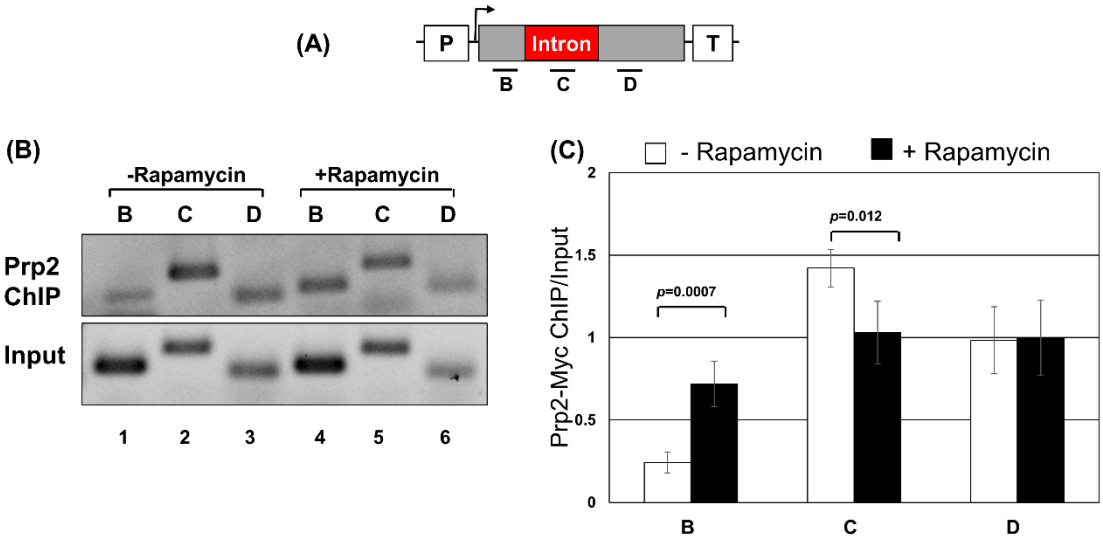

Figure 2. 14 Crosslinking of Prp2 to intron of *APE2* gene. (A) Schematic depiction of *APE2* indicating the position of ChIP primer pairs A, B, and C. (B) ChIP analysis showing crosslinking of Prp2 to different regions of intron-containing *APE2*. The input signal represents DNA prior to immunoprecipitation.

essential splicing factor but enhances the efficiency of splicing of a subset of yeast genes.

2.4 Discussion

Xrn2 has been implicated in degradation of unspliced mRNA in humans (Davidson et al., 2012). A report suggested involvement of Rat1 in degrading unspliced transcripts in budding yeast as well (Bousquet-Antonelli et al., 2000). In a temperature-sensitive mutant of essential splicing factor Prp2 called *prp2-1*, splicing was compromised, and unspliced transcripts could be detected at elevated temperature, as expected. In the double mutant *prp2-1/rat1-1*, there was a further increase in the amount of unspliced transcripts of selected genes. The authors attributed this additional increase in the double mutant to the role of Rat1 in degrading unspliced transcripts by its exoribonuclease activity. Estimating the level of unspliced transcripts in the *rat1-1* mutant alone, which is an important control, was not performed in this study. When we measured splicing in the *rat1-1* mutant, we observed accumulation of unspliced transcripts at the non-permissive temperature even in cells with wild type Prp2. Nuclear depletion of Rat1 by anchor-away approach gave identical results indicating that *rat1-1* mutant does not have a secondary mutation in another splicing factor. We have three pieces of evidence to support that accumulation of unspliced transcripts in the absence of Rat1 activity in yeast cells is not due to the surveillance role of Rat1 in degrading unspliced transcripts. First, unspliced transcripts were capped. Capping of unspliced mRNA was almost to the same extent as that of spliced transcripts. Rat1 can degrade 5′ monophosphorylated transcripts, but not 5′ methylguanosine capped transcripts. Second, Rat1 does not copurify with any known decapping protein. Xrn2 in humans coimmunoprecipitates with three decapping proteins; Edc3, Dcp1a and Dcp2 (Brannan et al., 2012). Our mass spectrometric analysis of Rat1 did not yield any decapping proteins of yeast in affinity purified Rat1 preparation. Third,

deletion of Rat1 interacting protein Rai1, which has been shown to remove the demethylated cap from mRNA in plants, has absolutely no effect on the level of unspliced transcripts in yeast cells. These results rule out the possibility of Rat1 degrading unspliced transcripts in budding yeast.

Rat1-dependent accumulation of unspliced mRNA can also be due to an indirect effect of Rat1 function in termination or elongation of transcription. We have three lines of evidence that unspliced transcripts are not produced because of defective termination in Rat1 mutants. First, unspliced transcripts were polyadenylated in *rat1-1* as well as *Rat1-AA* cells. Second, in the mutant of CF1 subunit Rna14, which affects the cleavage-polyadenylation step of termination, the amount of unspliced transcripts was similar to that in isogenic wild type cells. Third, in the absence of Rat1 interacting partner Rai1, which is not a termination factor itself but affects dissociation of the polymerase from the template by stimulating 5′→3′ exoribonuclease activity of Rat1, there was absolutely no increase in the level of unspliced mRNA over the wild type control. Our results also demonstrate that generation of the unspliced transcript is not the result of increased elongation rate associated with the *rat1-1* mutant. In *Rat1-AA* strain, no reduction of polymerase signal over the intronic region was observed, thereby strongly suggesting that there was no decrease in pausing of the polymerase while transcribing the intronic region when Rat1 is nuclear depleted.

The results discussed above give credence to the hypothesis that Rat1 is playing a direct role in splicing of mRNA in budding yeast. We have three pieces of experimental evidence in support of involvement of Rat1 in splicing. First, genomewide ChIP analysis found Rat1 crosslinked to the intronic region of a number of intron-containing yeast genes.

We expect a factor with a direct role in splicing to contact the intron. Since splicing occurs cotranscriptionally, any protein that is involved in splicing gets indirectly crosslinked to the intronic regions on the gene through its direct contact with the transcribing RNA and the polymerase. Second, mass spectrometric analysis of tandem-affinity-purified Rat1 found at least five splicing factors, Sub2, Isy1, Prp43, Yju2 and Clf1, in the purified preparation. Three of these splicing proteins, Clf1, Isy1 and Yju2, are a part of the NineTeen complex that helps in activation of the assembled spliceosome. Third, recruitment of Prp2, an essential DEAD-box containing splicing factor in yeast, onto the intron is compromised in the absence of a functional Rat1 in the cell. On the basis of these results, we propose that Rat1 has a novel role in splicing of precursor mRNA in budding yeast. Rat1, however, is not an essential splicing factor as splicing of less than 50% of intron containing genes is dependent on Rat1. Rat1 ChIP-Seq analysis revealed 105 out of 280 genes analyzed in the study showing crosslinking of Rat1 to introns. All the genes that exhibited intronic Rat1 occupancy, however, are not dependent on the protein for removal of their introns. We believe that there are other factors that determine if Rat1 localization on the intron will require it to complete the splicing reaction. In *C. elegans*, several genes that exhibited 3' end occupancy of Rat1 did not require Rat1 for termination. Whether Rat1 facilitated termination was dictated by the promoter of the gene. A similar mechanism might be determining if intron-bound Rat1 plays a role in splicing in yeast. It is possible that either the promoter or terminator region of the gene determines the Rat1-dependence on splicing. Even considering the genes that require Rat1 for splicing, there is no complete loss of splicing in the absence of Rat1. On average, there is 40-80% decrease in splicing efficiency in the absence of Rat1 for the genes that require it for removing an intron. These

data suggest that Rat1 is not an essential splicing factor but has a rather stimulatory effect on the splicing process.

We searched for features of introns that made them dependent on Rat1 for efficient splicing. The 105 Rat1-occupying introns, which are relatively more dependent on Rat1 for splicing, were found to be approximately four-times longer in size compared to introns showing no Rat1 occupancy. The longer size of Rat1-crosslinked introns was not due to a longer distance between the branchpoint and 3′ splice site, but due to a longer distance

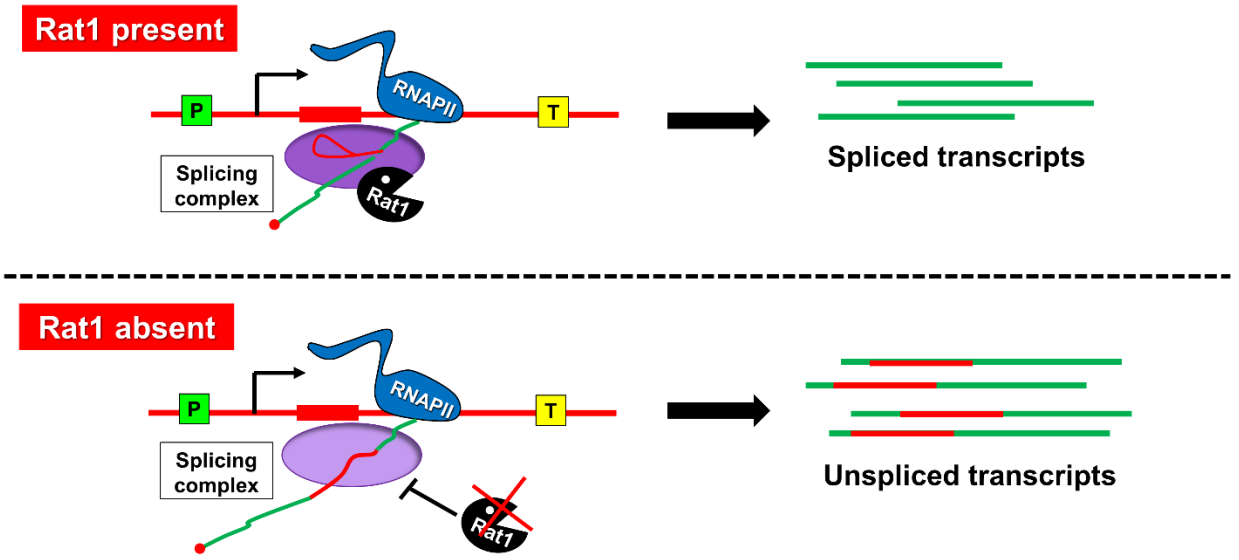

Figure 2. 15 Model for Rat1 role in cotranscriptional splicing. Rat1 promotes cotranscriptional splicing leading to formation of mature mRNA. At intronic sites, Rat1 coordinate with components of NineTeen complex (Clf1, Isy1, Yju2,) and helicase (Sub2 and Prp43) that are known splicing factors. In the absence of Rat1, loss of coordination between Rat1 and splicing factors results in the increase levels of unspliced transcripts.

between the 5′ splice site and branchpoint. These data suggest that Rat1 probably has a role in stabilizing the interaction of the branchpoint with 5′ splice sites of long introns, which is critical for the first transesterification reaction. We propose that Rat1 is able to do so due to its interaction with NineTeen complex as shown in the Figure 2.15.

Apart from Prp2, Rat1 is known to exhibit a genetic interaction with at least four splicing factors; Msl5, Yhc1, Sub2 and Prp46, in yeast thereby corroborating the notion

that Rat1 is a splicing factor (Bousquet-Antonelli et al., 2000; Costanzo et al., 2016; González-Aguilera et al., 2008). A logical question is if Rat1 homolog Xrn2 plays a similar role in splicing in higher eukaryotes. Xrn2 is a component of *in vivo* assembled, purified splicing complexes in human HeLa cells and chicken DT-40 cells (Chen et al., 2007). These purified supraspliceosomes contained all five U-snRNPs and other known splicing proteins. In addition, they identified some novel proteins that are not known to play any role in splicing. Xrn2 was one such protein that was present in both human and chicken supraspliceosomes. The supraspliceosomes did not contain any other termination factor besides Xrn2. The possibility of Xrn2, like Rat1, playing a role in splicing of a subset of introns in higher eukaryotes therefore cannot be ruled out.

CHAPTER 3. COMPREHENSIVE ANALYSIS OF RAT1-INTERACTING PROTEINS IN BUDDING YEAST

3.1 Abstract

Rat1 was initially discovered as a protein involved in transport of poly(A)-containing mRNA from the nucleus to the cytoplasm in budding yeast. Investigations over two decades have implicated Rat1 and its human homolog Xrn2 in several aspects of RNA metabolism. These include steps in RNAPII-mediated transcription such as elongation, termination, promoter-associated premature termination and upstream anti-sense transcription. Furthermore, Rat1/Xrn2 function in the processing of several classes of RNA including mRNA, rRNA, snoRNA, and miRNA. Involvement of Rat1/Xrn2 in a multitude of nuclear processes suggests that the protein must be interacting with a variety of protein partners associated. To understand the comprehensive biological function of Rat1 in budding yeast, we attempted biochemical purification of Rat1 by Tandem Affinity Purification (TAP) approach. Isolation of Rat1 has been attempted before, but a very limited set of factors were identified. Purification of Rat1 using Rtt103 as a bait revealed the existence of Rat1 in a trimeric complex with Rtt103 and Rai1. We attempted purification using Rat1 protein itself as the bait. Tandem-affinity purification followed by mass spectrometry identified several Rat1-interacting proteins that can be broadly divided into six classes. These are, termination factors, initiation factors, elongation factors, RNAPII subunits, rRNA processing factors, and mRNA processing factors. Some of these factors, like Rai1 and Pabp1, exhibit a strong interaction with Rat1 suggesting that Rat1 is in a stable complex with these factors. However, other factors were in low abundance, which suggests that these may be the transiently interacting factors involved in different

steps of transcription and RNA processing. Presence of such a variety of factors in the Rat1 preparation explains the observed diversity in Rat1 function in budding yeast.

3.2 Introduction

In general, the interactions of protein molecules in the cell can be broadly categorized as stable or transient. Inside the cell, most of the proteins tend to bind other molecules that they frequently run into because of their continuous state of random thermal motions (Frauenfelder et al., 1991; Northrup and Erickson, 1992). These collisions can result in the formation of noncovalent bonds between two proteins with weak affinity towards each another, leading these proteins to interact transiently and dissociate rapidly soon after. Many of the cellular processes including transcription, RNA processing, and transport are regulated by such transient interactions (Westermarck et al., 2013). These interactions are often regulated by post-translational modification of proteins, such as phosphorylation, acetylation, methylation, glycosylation, ubiquitinylation as well as by ligand-induced conformational changes in structure of protein (Filtz et al., 2014; Prabakaran et al., 2012). The random collision events can also result in enough noncovalent bonds to form a stable macromolecular complex. For instance, ribosome consists of protein-protein as well as protein-RNA interactions that result in a stable ribonucleoprotein complex (Staley and Woolford, 2009). Proteins that exhibit stable interactions can be purified as multi-subunit complexes from the cell. Ribosomes, RNAPII, TFIID and TFIIH represents a classic example of a stable complex that exists in the cell (Bushnell and Kornberg, 2003).

The function of a protein is influenced by the chemical properties of its surface, which is modulated by the conformation of the protein (Frauenfelder et al., 1991). This

can help to precisely position amino acid residues or side chains that can dictate binding of the protein to its ligand molecule. In addition, such conformation changes can also alter the binding site of the substrate in the active center of the enzyme (Gunasekaran et al., 2004). In order to understand the function of a protein at the system level in a cell, it is crucial to identify its interacting partners or the protein within in the cell. The data from such studies can be key to visualize an appropriate testable model to gain insight into the mechanistic regulation of a step and/or the process, thus defining the function and application of the protein in the cell. Further, proteomic analysis can aid not only to characterize and confirm *in vivo* functional interactors of the protein in a complex but may also help in discovering any unanticipated functional novel of the protein. Any identified protein-protein interaction thus obtained must be verified by additional independent approaches.

Traditionally, most of the studies to identify stable protein complexes often employed multiple chromatographic separation procedures like anion-exchange, cation-exchange, hydrophobic-interaction and size-exclusion chromatography (Coskun, 2016; Porath, 1997). The proteins that copurify through multiple chromatographic steps or sedimentation procedures are considered to be a part of a single complex (Dong et al., 2008). Through these strategies, the protein complexes were identified one at a time and sometimes experimental conditions were applied that were non-physiological, including high salt conditions that restricted its applicability to purify only stably associated proteins in a complex. The advent of gene manipulation and the use of epitope tagging has revolutionized the isolation of protein complexes allowing researchers to identify both stable and transient interacting partners of a protein. Affinity purification coupled with

high-throughput mass spectrometry analysis have immensely aided in identifying protein complexes that are both stable and transient in nature. In particular, TAP approach is one of the most effective tool that currently exists to isolate protein complexes and it has been widely used to predict protein interactome maps in budding yeast and other organisms (Bürckstümmer et al., 2006; Gavin et al., 2002; Rigaut et al., 1999).

A previous report, that employed multiple chromatographic procedures to purify Rat1 from budding yeast, identified a 45 kDa polypeptide called Rai1 as a stable interacting partner of Rat1 (Stevens and Poole, 1995). Later, a large-scale proteomic analysis validated the interaction of Rat1 with Rai1. This study used Rai1 as a bait protein for purification (Gavin et al., 2002). TAP-approach coupled with mass spectrometry was used to develop a proteomic map in budding yeast in this study. An interesting finding of this endeavor was that Rat1 is a part of another complex with twelve proteins when Rpl2A is used as a bait. This complex was predicted to be involved in protein synthesis and turnover. The existence of this complex, however, could not be independently verified by other studies. Further, the study that defined the role of Rat1 in termination in budding yeast identified it in a complex with Rai1 and Rtt103 (Kim et al., 2004b). Rtt103 is a CTD-interaction domain (CID) containing protein and was used as a bait for the affinity purification in this case. The existence of the Rat1-Rai1-Rtt103 trimeric complex was further validated by another large-scale proteomic analysis (Krogan et al., 2006). Extensive investigations over the years have suggested a role for Rat1/Xrn2 in diverse cellular processes that are central to RNA metabolism. The biochemical purification of Rat1 in budding yeast has led to the identification of interacting partners that explains its role in termination and RNA quality control. However, it is still not clear how Rat1 functions

in RNA processing of rRNA, snoRNA, and transcription elongation. It is logical to assume that Rat1 interacts with different proteins as well as substrates in order to achieve these diverse set of functions related to RNA processing and metabolism. Thus, we believe that a detailed proteomic analysis of Rat1 will help to elucidate its function not only for the previously established processes, but also in the newly proposed role in splicing of pre-mRNA in budding yeast.

Xrn2 chromatin occupancy and its enzymatic activity in mammalian cells was found to be regulated by Cdk9 (Sansó et al., 2016). Specifically, threonine-439 was the target site that was found to be phosphorylated in this study. The findings from this study are central to the argument that Rat1, in the context of chromatin, can undergo several modifications which in turn regulate its enzymatic activity and occupancy on the chromatin. Moreover, it can be argued that such posttranslational modifications of Rat1/Xrn2 can also govern its transient association with other proteins, thereby allowing it to function in different cellular processes. It is possible that Rat1 interacts with different set of proteins in different conditions or cellular environment which may extend the functional repertoire of the protein. The knowledge of the interacting proteins of Rat1, especially inside and outside the context of chromatin will help to greatly improve our understanding of how Rat1 activity is finely tuned to adapt to different cellular processes.

The rationale to investigate of interacting partners of Rat1, inside and outside the context of chromatin, has been inspired from the previous studies that have attempted to purify transcription factor TFIIH and Mediator complex. TFIIH is a multi-subunit complex that consists of eleven subunits, of which seven belong to the core module, whereas three

subunits are part of a dissociable TFIIK submodule. An initial attempt to purify native TFIIH using a combination of chromatographic techniques led to the identification of the core proteins, where the association of the kinase module TFIIK was lost (Chang and Kornberg, 2000). This association was later identified when isolation of the complex from the yeast cell extract was adopted under physiological conditions using affinity approach. Specifically, Kin28, Ccl1, and Tfb3 were the three proteins that were identified to be a part of the TFIIK complex (Gibbons et al., 2012). Similarly, the Mediator complex, a coactivator of RNAPII-mediated transcription, was also found to contain a transiently associated kinase subcomplex called the CDK8 module. Specifically, it was identified that Mediator complex, which is comprised of 25 subunits in yeast can be purified in at least two distinct forms, a stable associated structure of 21 bona fide proteins constituting the core complex and another complex that contains both the core complex components with a 4-subunit cyclin-dependent kinase 8 (CDK8) kinase module (Asturias et al., 1999; Davis et al., 2002; Dotson et al., 2000; Taatjes et al., 2002). From these findings, it is clear that the composition of Mediator complex in yeast varies greatly depending on the purification strategy employed. Use of stringent conditions failed to detect transiently associated proteins. Furthermore, these Mediator compositions are based on purification of the protein from the soluble fraction. Mediator complex and some of its subunits strongly interact with chromatin. An attempt to isolate Mediator from the chromatin fraction, without the use of any crosslinking agent, led to the identification termination factors of CF1 and CPF complexes as its novel, unique interacting partners (Chereji et al., 2017). This result is interesting as it provides strong supporting evidence towards the existence of a physical interaction between promoter and termination bound factors, thus supporting the idea that

a gene adopts a looped conformation during transcriptional activation. The functional role of Mediator subunit Srb5 in termination was established long ago, which further supports the data from this high through-put measurement of Mediator complex protein interactors in the context of chromatin (Chereji et al., 2017; Mukundan and Ansari, 2011). In addition, several other transcription factors exhibit functions that cannot be justified by their currently known protein interactors. This demands the need to reexamine interactors of transcription factors keeping in mind their role in the context of chromatin.

In this study we attempted to examine interactors of Rat1 inside and outside the context of chromatin using chromatin and soluble fractions, respectively. Our attempt to identify Rat1 interactors from chromatin fraction failed. From the soluble fractions, we successfully isolated Rat1 interacting partners with high confidence value. In our screen, we identified many of the previously identified interactors of Rat1. These include Rai1, Rtt103, Grc3, and Las1. In addition, we identified several novel interactors including the one involved in different steps of RNAPII cycle. Lastly, we also identified splicing factors Clf1, Yju2, Isy1, Sub2 and Prp43 in association with Rat1 that explains the role of Rat1 in the splicing of pre-mRNA.

3.3 Results and Discussion

3.3.1 Purification of Rat1 complex components by TAP-coupled mass spectrometry

To determine the native composition of the Rat1 complex in *S. cerevisiae*, and to address whether Rat1 interacts and/or exists in multiple complexes, we adopted TAP-approach (Rigaut et al., 1999). The affinity tag used in this procedure is referred to as a TAP-tag as it allows two affinity/elution steps to isolate the protein of interest. In the first

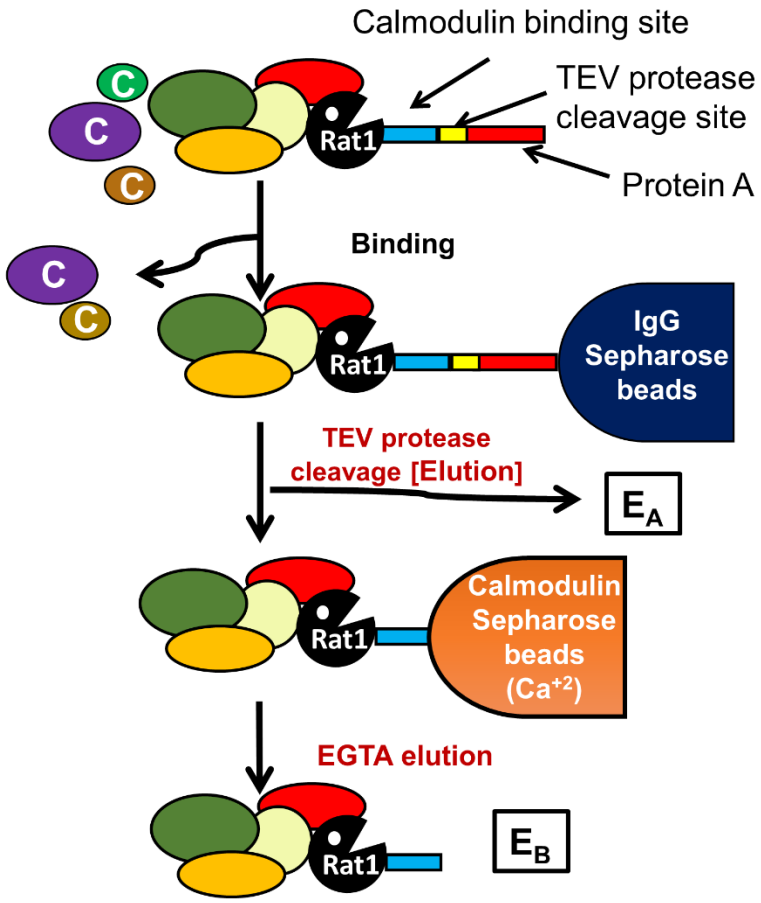

Figure 3. 1 Schematic overview for Tandem affinity purification. Protein A domain of the tag binds to IgG Sepharose and subsequent elution using TEV protease in the first step. CBP (calmodulin binding peptide) binds to Calmodulin Sepharose and the purified complex is eluted with EGTA (adapted from Puig et.al., 2001). The elution at these steps are represented as E_A and E_B, respectively

affinity step, Protein-A tag of TAP-tagged protein is allowed to bind to the IgG-Sepharose beads. The protein along with its interacting partners is eluted from IgG-Sepharose beads by TEV protease. In the subsequent affinity step, CBP part of TAP-tag is allowed to bind to Calmodulin-Sepharose beads in the presence of calcium. The elution from the second affinity step is then achieved with a mild chelator EGTA (Figure 3.1). The endogenous genomic loci of *RAT1* was modified to express Rat1 protein with a C-terminal TAP-tag. The strain producing the TAP-tagged Rat1 exhibits morphology and growth

characteristics akin to the parent strain. The purification of Rat1 from a two-liter culture was attempted and the presence of Rat1 was confirmed at every step using Western blot. The bound proteins were eluted from the Calmodulin-Sepharose column using 3mM EGTA in the form of five fractions of 200 µl each.

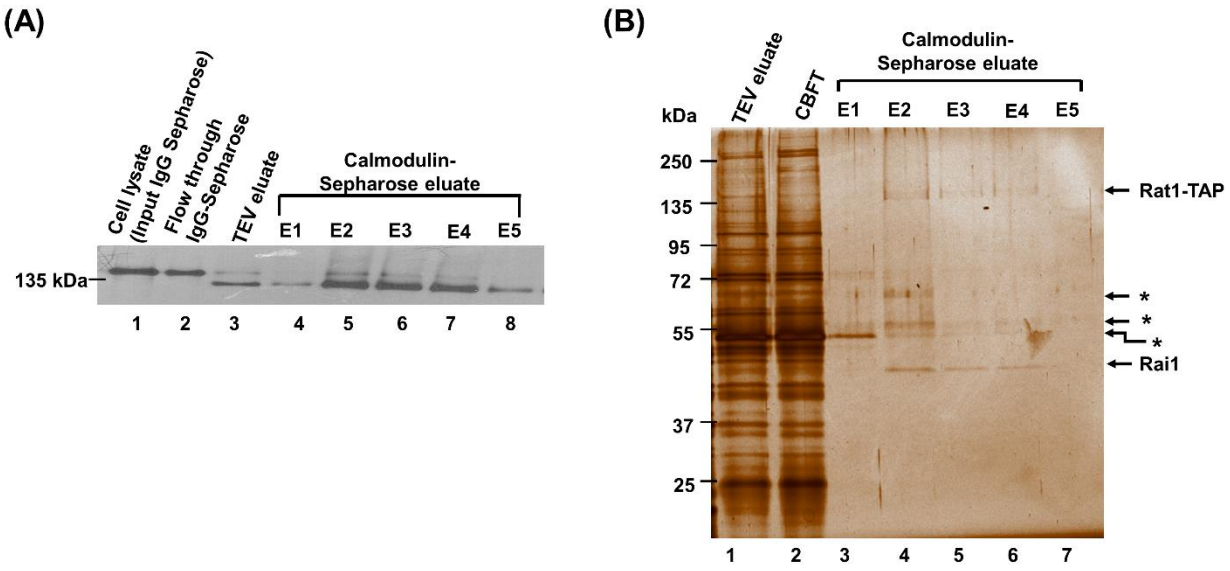

Figure 3. 2 Purification of proteins interactors of Rat1 from yeast strain with TAP tag inserted at the C-terminus of Rat1 gene. (A) Western blot analysis at the different steps of the purification tracks Rat1 before and after TEV cleavage step. Rat1 is higher molecular weight prior TEV cleavage (lane 1 and 2) compare to post TEV cleavage (lane 3-8). (B) Silver stained gel (10% SDS-PAGE) showing the presence of proteins in the after TEV digestion (TEV eluate), flow through from calmodulin binding step (CBFT), and five elution fractions (E1-E5) following calmodulin binding step. High amount of proteins in TEV and CBFT represents considerable contaminant proteins. Eluate fractions E2, E3, E4 represents elution Rat1 and associated proteins in a complex.

In order to examine, how Rat1 can function in different cellular processes, we analyzed Rat1-associated proteins from two different fractions, soluble and chromatin. We reasoned that Rat1 associated factors would be considerably different in the soluble and chromatin fractions, with the possibility that the proteins associated with Rat1 in the chromatin fraction would be more representative of the functionally active complex. We avoided using stringent high salt conditions for elution to prevent the dissociation of

complex subunits during purification, and rather eluted Rat1 complex from the chromatin using DNase. The chromatin eluted fraction was subjected to affinity purification with Rat1 as the bait. Analysis of the protein from the affinity purified sample showed that the coverage of bait protein, Rat1, was significantly lower in the preparation while background was remarkably high making it difficult to identify the true interacting partners. Therefore, we proceeded further with isolation of the Rat1 complex from the soluble fraction. Three independent grown cultures WT Rat1-TAP (test) and WT no-tag (control) were processed and tested using silver staining and western blot analysis. Two preparations were removed from the screen as the fraction obtained from second affinity purification step were not homogenous in their purity determined by silver staining.

Protein _ID	Rat1-TAP		WT (No tag-Control)
	Counts	Coverage (relative to Rat1)	Counts
Termination			
Rat1	233	100 %	0
Rai1	91	39.1%	0
Pab1	83	35.63%	3
Rtt103	7	3.0%	0
Reb1	19	8.2%	0
Polymerase			
Rpb1	27	11.6%	0
Rpb2	21	9.0%	0

Rpb3	2	0.86%	0
RPA190	30	12.9%	7
Rpa135	17	7.3%	3
Elongation			
Paf1	15	6.5%	0
Elp1	13	5.6%	0
Elp3	5	2.15%	0
Elp4	2	0.86%	0
Spt6	21	9.01%	0
Initiation			
Tfg1 (TFIIF)	23	9.87%	2
Tfg2 (TFIIF)	5	2.14%	0
Taf12	20	8.58%	1
Taf5	14	6.0%	2
Taf6	3	1.29%	0
Taf1	4	1.72%	1
Taf10	5	2.15%	
Ssl1	9	3.87%	1
Splicing			
Sub2	25	10.73%	1
Isy1	2	0.86%	

Prp43	5	2.15%	
Yju2	1	0.43%	0
Clf1	2	0.86%	0
rRNA processing			
Las1	1	0.43%	0
Grc3	1	0.43%	0
PNO1	1	0.43%	0
Export			
Dbp5	12	5.15%	0
Sub2	25	10.73%	1

Table 2 List of proteins identified in the affinity purified fraction of Rat1-TAP purification are sorted and characterized in to six different classes that are related to RNA metabolism. The spectral count of proteins is listed from the test (Rat1-TAP strain) and no-tag control (WT strain). Proteins levels relative to Rat1 for the test sample are listed.

The presence of Rat1 was monitored through different steps of purification using Western blot analysis (Figure. 3.2). Simultaneously, the composition of the proteins associated with Rat1 was confirmed by visualizing the eluted proteins by silver staining (Figure 3.2). Mass spectrometry of the trypsin digested peptides from the pooled elution fractions (E2-E4) was then used to identify the composition of the Rat1-complex. Proteins present in the mock experiment, which was carried out in parallel with a strain lacking a TAP tag on Rat1, served as a control to identify any non-specific contaminant. Proteins that were identified by mass spectrometry as copurifying with Rat1 can be sorted into six functional categories that are listed in Table 2. Of these, four functional categories were identified as proteins involved in different steps of the RNAPII transcription cycle,

initiation, elongation, termination, and splicing. The remaining two functional categories identified were rRNA processing and RNA export proteins.

3.3.2 Interaction of Rat1 with termination factors of RNAPII transcription cycle

As mentioned earlier, few studies that investigated the role of Rat1 in termination, identified the protein in a complex with Rai1 and Rtt103 (Kim et al., 2004b; Pearson and Moore, 2013). Validation of these interactions has been done using both *in vivo* as well as *in vitro* approaches. It was proposed that Rtt103 is first recruited at the 3' end of the gene as it contains a CTD interacting domain (CID) that interacts with the CTD of RNAPII (Kim et al., 2004b). Specifically, the phosphorylation of the serine-2 and threonine-4 residues on the CTD signals the recruitment of Rtt103. Rat1 is believed to be recruited at the 3′ end of a gene through the interaction of the Rtt103 subunit with the CTD (Jasnovidova et al., 2017; Kim et al., 2004b). The interaction between Rat1 and Rai1 has been reported with higher confidence in many studies, where it has been shown to stimulate the exoribonuclease activity of Rat1 in degrading the RNA substrate that is 5' monophosphorylated. In our analysis, we detected high peptide counts of 81 for Rai1, whereas corresponding peptide counts for Rtt103 were merely 5. The peptide counts of Rai1 confirms the existence of a stable interaction between Rat1 and Rai1 in the complex. The lower peptide counts of Rtt103 in the purified fraction implies that there is a weak interaction of Rtt103 with the complex. Although our data supports the previous studies that have shown the existence of a trimeric complex between these proteins, it also gives further insights into the strength of interactions that exist between these proteins. This

also explains why many studies were unsuccessful in identifying Rtt103 as an interacting partner as this transient interaction can be easily lost during biochemical purification.

In addition, we identified a novel interactor of Rat1called Pab1 or poly(A)-binding protein which binds poly(A) tail of mRNA immediately after cleavage-polyadenylation step. It needs further experimentation to understand how the Rat1-Pab1 interaction affects 3' end processing of mRNA. Like Rat1, Pab1 contributes to multiple nuclear processes related to mRNA metabolism. In the termination step of transcription, Pab1 interacts with CFI subunit Rna15, which is known to function in the cleavage and polyadenylation of mRNA (Amrani et al., 1997; Minvielle-Sebastia et al., 1997). Furthermore, role of Pab1 in regulating poly(A) tail length *in vitro* can be viewed as an important step for efficient export of mRNA into the cytoplasm (Brune et al., 2005; Dunn et al., 2005). This leaves open the possibility that Rat1 could be coordinating with Pab1 to influence either termination or mRNA export or both.

3.3.3 Interaction of Rat1 with elongation factors of RNAPII transcription cycle

The Y657C mutant of Rat1 exhibits synthetic lethality with the *rpb1-E1103G* mutation, which is responsible for increased RNAPII elongation rate. However, the mutant phenotype of Rat1 is alleviated by *rpb1-N488D and rpb2-10* mutations that lower the elongation rate of RNAPII (Jimeno-González et al., 2010, 2014). In addition, Rat1 and Rai1 were found in the screen to identify the components of the early elongation complex that was purified using the Rpb3 component of RNAPII (Harlen and Churchman, 2017). Our results show the presence of three RNAPII subunits, Rpb1, Rpb2 and Rpb3 in the purified preparation of Rat1. These suggests Rat1 recruitment during transcription can be

through its interaction with RNAPII subunits. This may explain recruitment of Rat1 at the 5' end, gene body and 3' end of the gene in the absence of Rtt103.

Genetic interaction of Rat1 with slow and fast elongating mutants of RNAPII implicated Rat1 in elongation step of transcription. There is, however, no evidence of a direct physical interaction of the protein with elongation factors. Here, we demonstrate for the first time, the interaction of Rat1 with elongation factors Paf1, Elp1, Elp3, and Elp4. Paf1 copurifies with RNAPII and has been shown to play an important role in elongation of transcription. Paf1 is known to interact with four other proteins, Cdc73p, Ctr9p, Leo1p, and Rtf1p (Krogan et al., 2006; Mueller and Jaehning, 2002). Paf1 functions with Rtf1 to modulate histone modifications on the chromatin template. Specifically, Paf1 and Rtf1 can mono-ubiquitinylate H2B on K123 in yeast and introduce di- and trimethylation of H3-K4 and H3-K79 (Laribee et al., 2005; Ng et al., 2003; Wood et al., 2003). Furthermore, a component of COMPASS, Set2, associates with RNAPII during elongation in a Paf1-dependent manner (Krogan et al., 2003). Like Paf1, we found interaction of Rat1 with elongator complex subunits as well. Elongator complex is comprised of a core subunits Elp1, Elp2, Elp3, and a subcomplex of Elp4, Elp5, and Elp6. Elp3 is the histone acetyltransferase enzyme of this complex. (Dong et al., 2015; Wittschieben et al., 2000). Rat1 interacts with Elp1, Elp3 and Elp4. The functional significance of Rat1-elongator interaction needs further research.

In addition to its role in transcription elongation, Paf1 has also been implicated in termination step of transcription (Nordick et al., 2008). Specifically, Paf1 makes contact with multiple cleavage and polyadenylation factors, and accordingly was shown to play a role in the 3′ end formation of mRNA (Van Oss et al., 2017). Our data showing the

association of Rat1 with Paf1 would be important in understanding whether Rat1 is involved in influencing elongation of transcription with Paf1 or alternatively, Paf1 may influence elongation which can affect the Rat1 role in termination.

3.3.4 Interaction of Rat1 with initiation factors of RNAPII transcription cycle

The crosstalk between the 5' and 3' end of the gene mediated through gene looping can explain the physical association of promoter-bound factors with the termination factors at the 3' end of the gene (O'Sullivan et al., 2004). However, it is still unclear how these processes are interconnected. Studies have clearly shown that Rat1 and its homolog Xrn2 are not only present at the 3' end of the gene, but also at the 5' end of the gene (Baejen et al., 2017; Brannan et al., 2012; Fong et al., 2015; Kim et al., 2004b; Nojima et al., 2015). There have been a few speculations about the role Rat1 can play at the 5' end of the gene. Rat1 and Xrn2 both have been reported to degrade uncapped RNA. In fact, Xrn2 is reported to associate with decapping factors Edc3, Dcp1 and Dcp2 in the mammalian system (Brannan et al., 2012). In addition, it has been proposed that Xrn2 is involved in the regulation of promoter-associated antisense transcription. The mechanistic details of the process, however, are unclear. Although both Rat1 and Xrn2 have been found localized to the 5' end of the gene, molecular interactions that facilitate recruitment of Rat1 to the promoter are absolutely lacking. Here we present evidence, for the first time, that shows Rat1 interacts with subunits of general transcription factors TFIIF, TFIID, and TFIIH. Specifically, we identified interactions of Rat1 with Tfg1 and Tfg2 subunits of TFIIF; Taf1, Taf5, Taf6, Taf10, and Taf12 subunits of TFIID; and the Ssl1 subunit of TFIIH in our screen. These interactions will help elucidate mechanistic details

underlying the role of Rat1 in limiting upstream anti-sense transcription and abortive transcription of uncapped transcripts.

3.3.5 Interaction of Rat1 with splicing factors

In yeast, Rat1 exhibits genetic interactions with five different factors that are involved in splicing; Prp2, Prp46, Msl5, Yhc1 and Sub2 (Bousquet-Antonelli et al., 2000; Costanzo et al., 2016; González-Aguilera et al., 2008). All these proteins are required for the formation of the catalytic spliceosomal complex, which implies that these factors contribute to the first step of the splicing reaction. Yhc1 is a subunit of U1 snRNP that is bound on the 5' SS. Msl5 is known to exist in a heterodimer with Msl2, and together they are known to interact with U1 snRNP and participate in the recruitment of U2 snRNP at the branch point (Neubauer et al., 1997; Tang et al., 1997; Zhang and Rosbash, 1999). The Msl5-Mud2 heterodimer is proposed to facilitate the recruitment of the DEAD-box RNA helicase, Sub2. In coordination with another helicase Prp5, Sub2 functions to promote the recruitment of U2 snRNP at the branchpoint region (Zhang and Green, 2001). Prp46 is one of the core components of the NineTeen complex that associates during the first step of the splicing reaction (Albers et al., 2003). Prp2 is an ATPase, which is required for spliceosome activation during the first transesterification step of the splicing process (Kim and Lin, 1996). It has been demonstrated that unspliced transcripts in a Prp2 mutant are stabilized to a greater extent when combined with Rat1 mutation (Bousquet-Antonelli et al., 2000). This study, however, did not perform a critical control, that is, monitoring the presence of unspliced transcripts in the Rat1 mutant alone, which we have done in our study. In addition, metazoan Xrn2 has been shown to physically associate with the splicing complex purified from human HeLa cells and chicken DT-40

cells (Chen et al., 2007). In our analysis, we identified interaction of Rat1 with five splicing factors; Clf1, Yju2, Isy1, Sub2, and Prp43. Clf1, Yju2, and Isy1 are the subunits of NineTeen complex (NTC). Clf1 and Isy1 are components of the core, whereas the Yju2 subunit is an accessory protein of NTC. Clf1 is an essential splicing factor that serves as a scaffold in spliceosome assembly during assembly of the tri-snRNP (U4 U5.U6) (Chung et al., 1999). Like Clf1, Yju2 is also an essential splicing factor that functions during the first transesterification step upon activation of the spliceosome. Shortly after the execution of the first step of splicing, Yju2 has been shown to dissociate from the spliceosome (Liu et al., 2007). The Isy1 subunit of NTC is involved in maintaining splicing fidelity, both during 3′ splice site and branch site selection (Villa and Guthrie, 2005). Although NTC associates with the spliceosome before the first step of splicing, it functions in both steps of splicing through its core and accessory proteins (Fabrizio et al., 2009). This is interesting as it provides some leads that can be investigated to understand how Rat1 is functioning in the process of splicing. The verification of the components identified in our screen is currently under progress.

3.3.6 Interaction of Rat1 with rRNA processing factors

Rat1 has also been shown to function in the processing of rRNA in budding yeast. Rat1 has been identified previously to exist in a tetrameric complex with Las1, Grc3 and Rai1 (Gasse et al., 2015). Las1 is an endoribonuclease that associates with the polynucleotide kinase, Grc3, to function in processing the ITS2 site of pre-rRNA (Pillon et al., 2017). Las1 and Grc3 exist in a stable complex and their interaction has been shown to persist under high salt conditions. However, when the purification of the Las1-Grc3 complex was performed under less stringent conditions, two additional proteins Rat1 and

Rai1 were identified (Gasse et al., 2015). Notably in this study, purification was attempted for all four proteins after overexpression. Thus, it can be argued that the endonuclease complex of Las1-Grc3 detected with exoribonuclease complex of Rat1-Rai1 may not exist *in vivo*. In our mass spectrometry data, we identified only one peptide hit for both Grc3 and Las1, supporting the view that these proteins may weakly interact with Rat1 and Rai1 and together they coordinate pre-rRNA processing in budding yeast.

3.3.7 Interaction of Rat1 with proteins involved in RNA export

The first phenotype observed in the *rat1-1* mutant was nuclear retention of poly(A) containing RNA (Amberg et al., 1992). This suggests Rat1 has a role in RNA trafficking, however, the molecular mechanism behind this function has not been explored to date. It is possible that Rat1 might affect binding of messenger ribonucleoproteins (mRNPs) that are involved in export. Our data identified two DEAD-box helicases, Dbp5 and Sub2, that are known to function in mRNA export. Dbp5 is an evolutionarily conserved protein that functions in the rearranging and remodeling of mRNPs through displacement of the poly(A) binding protein, Nab2 (Tran et al., 2007). In addition to this, Dbp5 has been shown to promote the release of Mex67 from the export mRNPs (Lund and Guthrie, 2005). Like Dbp5, Sub2 has also been implicated in nuclear mRNA export. The TREX (transcription and/or export) complex is composed of Sub2, Yra1 and the components of the THO (Tho2, Hpr1, Mft1, and Thp2) subcomplex, which is loaded onto RNA during transcription (Sträßer et al., 2002). The interaction of Yra1 and Sub2 is essential for mRNA export (Strässer et al., 2000). However, the precise mechanism behind the regulation of this process is not clear. Further investigation is needed to establish a potential link between Rat1 and these helicases in regulating mRNA export.

In conclusion, the proteomic analysis of Rat1 has led to the identification of diverse set of stable and transient protein interactors that are involved in the multitude of nuclear processes. The potential interacting partners of Rat1 identified from this study can serve as a useful link that can be investigated in future to determine the molecular mechanism of these interactions.

APPENDIX A: EXPERIMENTAL PROCEDURES

A.1 Cell Culture

A 5 ml culture was started in Yeast-Peptone-Dextrose (YPD) medium using colonies from a freshly streaked plate. The culture was grown overnight at 30ºC with constant shaking at 250 rpm. All *S. cerevisiae* cell cultures were grown in YPD medium unless otherwise stated. All strains except the temperature-sensitive mutants were incubated at 30°C and 250 rpm. Next morning, the overnight grown cultures were diluted 1:100 in YPD broth and allowed to grow at 30°C with constant shaking until the desired A_{600} was reached. Cells were centrifuged at 1,860x g for 3 mins at 4°C and the cell pellets were used further for the experiment.

A.1.1 Cell culture for Anchor-away experiment:

The starter culture Rat1-FRB-tagged strain was grown in YPD broth whereas strains with plasmids expressing wild type (pRS415-WT-Rat1) or catalytically inactive mutant of Rat1 (pRS415-D235A-Rat1) were cultivated in synthetic complete media without leucine. The cells were incubated overnight at 30°C and 250 rpm in an orbital shaker. The cells from the starter cultures were added in 1:100 dilution to the respective culture broths. Cultures were grown until A_{600} 0.5-0.7 was reached. At this point, one-half of the culture was treated with 1 mg/ml rapamycin and the other was left untreated. Both the samples were grown for another 60 min. Cells were collected by centrifugation as described above and used for RT-PCR, TRO and ChIP analysis.

A.1.2 Cell culture of temperature-sensitive strains

Temperature-sensitive mutants of Rat1 (*rat1-1*) and Rna14 *(rna14-1)*, were used in this study. The overnight culture of these cells was obtained as described above except

that cells were grown at 25°C. Next morning, the overnight culture was diluted 1:50, and not 1:100 as is normally done, in YPD broth. The cells were grown with gentle shaking at permissive temperature (25°C) until the A_{600} reached 0.25 and 0.4 for *rat1-1* and *rna14-1* strains respectively. Cells were then transferred to non-permissive temperature (37°C) for 3 hours (*rat1-1)* or 1 hr (*rna14-1*). Cells were harvested at A_{600} between 0.5-0.7 by centrifugation at 3000 rpm for 3 mins at 4°C. Another cell culture of these strains was run in parallel, incubated at permissive temperature (25°C) until desired A_{600} 0.5-0.7 reached. An isogenic wild type strain was subjected to similar growth conditions and the final cell density was normalized considering the mutants strain cell density. The harvested cells were used for RT-PCR, TRO and RIP analysis and were processed as described below in the respective experimental sections.

A.2 Transformation:

For deleting genes, epitope tagging with either TAP or FRB or MYC, transformation was carried out using heat shock, lithium acetate, polyethylene glycol (PEG) solution, and dimethyl sulfoxide (DMSO) in order to achieve high efficiency. Cells were grown to an A_{600} between 0.3-0.5, centrifuged at 1,860x g for 3 mins at 4°C and washed with 10 ml of lithium acetate buffer. Cells were resuspended in 200 µl of lithium acetate buffer. 100 µl of cells in lithium acetate buffer was transferred to microfuge tube with 5 µg plasmid DNA or 20 µg PCR product. The DNA was mixed with cells by gentle tapping and incubated for 5 mins at 25°C. The PCR product was generated using Advantage polymerase enzyme and gene specific primers coupled with a tag or knock-out plasmid listed in appendix C. After adding 280 µl of PEG solution, contents were mixed by vigorous shaking and incubated for 45 mins at 25°C. 43 µl DMSO was added to the cells and briefly

vortexed. Cells were subjected to heat shocked for 5 mins at 42°C and immediately transferred to ice for at least 2 mins. Cells were then centrifuged at 1,377 x g for 30 sec at 25°C. The resultant cell pellet was washed with 500 µl of sterile water. Cells were then resuspended and plated on appropriate selection plates. Colonies that grew on the selective media were further screened by colony PCR for positive transformants using primers listed in appendix C. For transformation where kanamycin was used as a selective marker, cells were incubated for two generation time in YPD at 30°C prior to plating on selective media. A 5' gene-specific primer and a 3' reverse primer within the tag was used to confirm that the DNA coding for the tag was inserted at the correct genomic position. For validation of a gene knockout, forward and reverse primer designed in the upstream and downstream regions of the open reading frame were used.

A.3 Cloning

pRS314 WT Rat1 and pRS314 D235A Rat1 were kindly gifted by Dr. Claire Moore. The vector backbone contained a *TRP* selection marker, which restricts the usability of this vector in the anchor away strain. Therefore, we cloned WT Rat1 and D235A Rat1 from the vector pRS314 to pRS415 vector which carries a *LEU* marker. Restriction enzymes XhoI and NotI were used to digest both vectors. The WT Rat1 and D235A Rat1 fragments were excised from pRS314 WT Rat1 and pRS314 D235A Rat1 vectors and purified from the gel using NucleoSpin Gel Clean-up kit (Clonetech). The gel purified WT/D235A RAT1 fragments were ligated to Xho1-Not1 digested pRS415 by Quick ligase (NEB). The ligation mixture was transformed into *E. Coli* DH5α and transformants were selected on 2xYT-ampicillin plates. The positive transformants were confirmed by sequencing and restriction digestion.

A.4 RT-PCR

Cells were grown and processed as discussed in the section A.1. Cells were washed with 5 ml DEPC (Diethylpyrocarbonate) treated water. Cells were resuspended in 500 µl of Trizol and transferred to the lock top tubes. To this tube, 250 µl of acid washed glass beads were added. Cells were lysed on a vortex machine for 40 minutes at 4°C. Using a 22-gauge needle, the tubes were then punctured at the bottom and drained into 2 ml tubes by centrifugation at 8,600 x g for 1 min at 4°C. To the cell lysate, 500 µl of Trizol was added and incubated for 5 minutes at room temperature. 200 µl of chloroform was added and samples were vortexed briefly. Samples were centrifuged at 12,000 x g for 15 minutes at 4°C. The aqueous layer containing RNA was carefully transferred to a new 1.5 ml eppendorf tube. 500 µl of isopropanol was added and tubes were incubated for 10 minutes at room temperature. RNA was pelleted by centrifugation at 12,000 × g for 10 minutes at 4°C. The RNA pellet was washed with 75% ethanol by centrifugation at 12,000 × g at 4°C for 10 mins. The pellet containing RNA was air dried for 5 mins, resuspended in 195 µl of 1x DNase buffer. The samples were treated with 50 U DNase for 30 mins at 37°C. Sample were digested with 100 µg Proteinase-K for 30 mins at 37°C. RNA was purified from this sample using phenol chloroform (pH 4.2) extraction. The RNA was precipitated by ethanol precipitation using glycogen as a carrier. RNA pellet was collected by centrifugation at 16,168 x g for 10 minutes at 4°C. RNA was air dried for 5 mins and resuspended in 52 µl of RNase-free water. The concentration of RNA was measured using a nanodrop. 2 µg/µl of RNA was used for cDNA synthesis using M-MuLV reverse transcriptase with oligo-dT and 5S 3' primers, respectively. cDNA was diluted ten times prior to step of PCR amplification by Taq DNA polymerase. The gene specific

primers used for PCR reaction are listed in appendix C. Each PCR was normalized against the 5S ribosomal RNA control. In parallel, a negative control without reverse transcriptase enzyme was run to ensure DNA contamination is not contributing to any RT-PCR signal.

A.5 Transcription Run-On Assay (TRO)

Cells were grown and processed as discussed in the section A.1. The transcription run-on assay was performed as described with few modifications. Briefly, the cell pellet was washed with 10 ml of ice cold TMN buffer and resuspended in 940 µl of ice-cold DEPC treated water. Cells were permeabilized by using 1% sarkosyl and incubated in an ice pack on a nutator for 25 minutes at 4°C. Cells were pelleted by centrifugation at 1,200 x g for 6 min at 4°C. *In vivo* transcription elongation was then performed in 150 µl of transcription run-on buffer at 30 °C for 5 minutes. The transcription reaction was immediately stopped by the addition of 500 µl of ice-cold Trizol reagent. Cells were then lysed with 250 µl acid-washed glass beads for 30 mins using a vortex at 4 °C. The tubes were punctured at the bottom using a 22-gauge needle and the lysate was collected into a 2.0 ml tube by centrifugation at 8,600 x g for 1 min at 4°C. 500 µl of Trizol was added and the lysate was incubated at 25°C for 5 mins. 200 µl of chloroform was added and the tubes were vortexed briefly and centrifuged at 12,000 x g for 15 mins at 4°C. Approximately 700 µl of aqueous layer containing RNA was transferred to a new microfuge tube. Equal volume of phenol/chloroform/isoamyl alcohol (125:24:1) pH 4-5 was added to the samples. Samples were vortexed briefly and centrifuged at 16,168 x g for 15 mins at 4°C. The aqueous layer containing RNA was transferred to a new microfuge tube. RNA was precipitated with 1/16th volume of 5M NaCl and 2.5 volumes of 100%

ethanol and incubating the samples overnight at -20°C. RNA from the samples was collected by centrifugation at 16,168 x g for 15 mins at 4°C. RNA pellet was washed once with chilled 70% ethanol and resuspended in 100 µl of RNase-free water. The RNA obtained that represents total RNA was purified using the Qiagen RNeasy kit and eluted twice with 50 µl of RNase-free water as per manufacturer protocol. The eluted RNA was incubated at 65°C for 5 mins to remove secondary structure. Following which, it was cooled on ice for at least 2 mins and used in immunoprecipitation step below in order to selectively purify BrdU-labelled nascent transcripts from the total RNA.

For immunoprecipitation step, a bed volume of 25 µl Anti-BrdU conjugated agarose beads was washed thrice with 500 µl of binding buffer by centrifugation at 200 x g for 30 sec at 4 °C. Supernatant was removed and discarded. To reduce non-specific binding, beads were incubated with 500 µl of blocking buffer on nutator for 1 hr at 4 °C. Beads were washed twice using 500 µl binding buffer. Beads were then resuspended in 400 µl of binding buffer. The total RNA purified described above was added to the Anti-BrdU conjugated agarose beads and incubated for 2 hours at 4 °C. The unbound RNA was removed by centrifugation at 200 x g for 30 sec at 4 °C. Beads were sequentially washed with 500 µl binding buffer, 500 µl low salt buffer, 500 µl of high salt buffer and 500 µl TET buffer. RNA was eluted with 250 µl of elution buffer twice by incubating at 42°C for 5 mins. Samples were subjected to phenol chloroform extraction, centrifuged at 16,168 x g for 15 mins at 4°C to obtain an aqueous layer containing RNA. The aqueous layer was transferred to a new 1.5 ml microfuge and RNA was precipitated with 0.1 volume LiCl and 2.5 volume of 100% ethanol overnight. Samples were centrifuged at 16,168 x g for 20

minutes at 4°C to collect RNA in the pellet. RNA was resuspended in 26 µl of RNase-free water and the concentration of the RNA was determined using nanodrop.

500 ng/µl of RNA was utilized for cDNA synthesis with Superscript IV reverse transcriptase with oligo-dT and 5S 3' primers. cDNA was diluted five times prior to use as a step in PCR amplification using Platinum SuperFi DNA polymerase. The gene specific primers used for PCR reaction are listed in appendix C. Each PCR was normalized against the 5S ribosomal RNA control. A sample without reverse transcriptase enzyme was run in parallel as a negative control.

A.6 Chromatin Immunoprecipitation (ChIP)

100 ml of cell cultures were treated with 1% formaldehyde and 20mM DSG in 1X PBS with constant shaking for 20 minutes at 25°C. Crosslinking was quenched by the addition of 125 mM glycine and the samples were incubated for 5 minutes at 25°C. Cells were transferred to 50 ml conical tubes and pelleted by centrifugation at 1,860 x g for 3 minutes at 4°C. Cells were washed with 10 ml of TBST and the cell pellets were resuspended in 500 µl of FA-lysis buffer. The resultant mixture was transferred to 1.5 ml lock top tubes, flash-frozen in liquid nitrogen, and stored at – 80°C. Cell were lysed using ~250 µl volume of acid washed glass beads and vortexed for 40 minutes at 4 °C. The tubes were punctured with a 22-gauge needle and lysate was drained into a 2.0 ml tube by centrifugation at 8,600 x g for 1 minute at 4°C. Samples were centrifuged at 16,168 x g for 30 minutes at 4°C to collect the chromatin pellet. The pellet was washed with 500 µl of FA-lysis buffer. The pellet was then resuspended in 1 ml FA-lysis buffer, transferred to a 15 ml tube, and then the volume was adjusted to 4 ml using FA-lysis buffer. Chromatin was fragmented to yield DNA of ~400 bp by sonication (Branson digital sonifier) in an ice

water bath at 25% duty cycle using 20 second pulses followed by 20 second rest, for a total sonication time of 15 minutes. The sonicated lysate was then transferred into 2.0 ml tubes and centrifuged at 16,168 x g for 15 minutes at 4°C. The resultant supernatant that represents solubilized chromatin was pooled together and used for the immunoprecipitation in the next step. At this point 50 µl of solubilized chromatin was set aside as input.

Immunoprecipitation of Rat1-TAP was performed using IgG Sepharose 6 Fast Flow for 3 hours at 4°C. 20 ul of IgG Sepharose were washed twice with 500 ul of FA-lysis buffer. For Immunoprecipitation of Rat1-TAP, 500 ul of the solubilized chromatin sample was used and binding step. Incubation was carried on a nutator for 3 hours at 4°C. The samples were centrifuged at 1000 rpm for 30 sec at 4°C. The supernatant was then removed and the beads with immunoprecipitated chromatin was subjected to a series of washing steps.

For the immunoprecipitation of Prp2-Myc, 400 µl of the solubilized chromatin was pre-incubated with 5 µl Anti-Myc antibody (RN016M) in a 1.5 mL lock top tube for 2 hours at 4°C. 20 µl of Magnetic Dynabeads Protein A was prewashed two times prior to use. Antibody-antigen complex was captured on magnetic Dynabeads (1001D) by incubating at room temperature for 15 minutes. The tubes were placed on a Thermo Scientific Magjet rack, and supernatant containing unbound antigen and antibody was discarded. After removing the supernatant, the beads were subjected to a series of washing steps.

The washing of the immunoprecipitated sample was performed using 500 µl of the following buffers for 1 min in sequential order: FA-1 lysis buffer, FA-2 lysis buffer, FA-3 lysis buffer, ChIP wash buffer and 1x TE. After removing the supernatant from the final

wash, beads were resuspended in 100 µl of 1x TE and transferred to a new 1.5 ml lock-top microfuge tube. The supernatant was separated from beads by centrifugation (Sepharose beads) or on magnetic rack (Dynabeads). Bait protein bound to the DNA complex was eluted twice for 5 mins at 65°C with 150 µl of elution buffer.

At this step, input fraction and the elution from the immunoprecipitated chromatin were processed as described below. Eluted samples and input were subjected to 5 µl RNase A and incubated at 37 °C for 30 mins. The proteins in the samples were digested using 100 µg Proteinase K at 37°C for 30 mins. Crosslinking was reversed by incubating the samples in a 65°C water bath overnight. Phenol chloroform (pH 7.5) extraction was performed twice to purify DNA. After phenol chloroform precipitation was completed, ethanol precipitation was performed to precipitate the DNA. The resultant DNA pellet was air dried for 5 minutes. The IP samples were resuspended in 52 µl of 1x TE and the input samples were resuspended in 102 µl of 1x TE. Purified DNA from the input and IP fraction were used as a template for PCR using the primers that were selected based on the region of the *APE2* gene. The primers used for ChIP-PCR analysis are listed in appendix C. The association of a protein to a given genomic region was expressed as ChIP over input ratios.

A.7 Data mining and re-analysis of published datasets

Published ChIP-Seq analysis dataset from the study by Baejen et al. (2017) was obtained through the Gene Expression Omnibus (GEO), GSE79222. Raw reads were further processed and analyzed using the Galaxy web platform. Raw reads (50 bp paired end reads) were aligned to the *S. cerevisiae* genome (sacCer3, version 64.2.1) using the Bowtie (version 2.3.4.2) (Langmead and Salzberg, 2012) using the parameters described

in Baejen et.al., (2017). To normalize and compare two BAM files and obtain a $\log_2$ ratio of IP/Input in the form of a bigwig file, bamCompare (Galaxy Version 3.1.2.0.0) in deepTools (Ramirez et al. 20116) was used. The aligned reads with MAPQ smaller than 7 (-q 7) were skipped. Signal extraction scaling (SES) factor was used for scaling and establishing ChIP enrichment profile over $\log_2$ scale with the options: –l 100 –n 100000 and the bin size of 20. The subsequent data in the bigwig file was selectively enriched for the genomic regions in the BED file using computeMatrix function (Galaxy Version 3.3.0.0.0). Specifically, three BED files were used in this study, 280 intron containing genes with TSS and TES coordinates, 280 intron containing genes with intron start and intron end coordinates, and 2392 non-intronic genes with TSS and TES coordinates. Heatmap was plotted to visualize the score distributions of the transcription factors or RNAPII enrichment associated with genomic regions specified in the BED files using the plotHeatmap (Galaxy Version 3.3.0.0.1). The metagene plot of the transcription factors and RNAPII was drawn with the plotProfile function (Galaxy Version 3.3.0.0.0).

A.8 RNA Immunoprecipitation

Mutant strain *rat1-1* was cultured and processed as discussed in the section A.1. RNA was isolated from these cells as described in the section A.3. 2 µg/µl of RNA was set aside as input that was used to quantify the total amount of spliced and unspliced RNA in the samples. Briefly, cDNA synthesis of 2 µg/µl of RNA was performed using Superscript IV reverse transcriptase and oligo-dT 3' primer. cDNA was diluted five times prior to use as a step in PCR amplification using Platinum SuperFi DNA polymerase. The gene specific primers used for PCR reaction are listed in appendix C.

The immunoprecipitation of the capped transcripts was performed by incubating 4 µl of Anti-7-methylguanosine (m^7G)-Cap mAb (RN016M, MBL) with 40 µg of total RNA adjusted to the total volume of 400 µl using RNA-IP binding buffer. 15 µl of magnetic Dynabeads (1001D) were prewashed with RNA-IP binding buffer twice prior to use. The immunoprecipitation complex was added to the beads and incubated on a nutator for 12 minutes at room temperature. The beads were separated from the sample using Thermo Scientific Magjet rack and supernatant was discarded. The beads were washed three times with 400 µl RNA-IP binding buffer. After the final wash, 100 µl of binding buffer was added, mixed to resuspend the beads, and transferred to a new 1.5 mL lock top tube. The buffer was removed as described above. Capped RNA bound to the Anti-7-methylguanosine (m^7G)-Cap mAb was released by 100 µl of elution buffer. Tubes were kept on a nutator for 5 minutes at room temperature, the beads were separated using magnetic rack and the supernatant containing capped RNA was transferred into a new tube, subjected to phenol-chloroform (pH 4.2) extraction. RNA was precipitated using 0.1 volume of 3M sodium acetate, 2.5 volume of 100 % ethanol and glycogen as a carrier. The RNA pellet was collected by centrifugation at 16,168 x g for 10 minutes at 4 °C, air-died and resuspended in 16 µl of RNase-free water. The concentration of RNA was determined using nanodrop. Around 200 ng/µl of RNA was used for cDNA synthesis by M-MuLV reverse transcriptase and 3' oligo-dT primer. cDNA was diluted two times prior to PCR amplification. The gene specific PCR primers used are listed in appendix C. A negative control without reverse transcriptase enzyme was carried out in parallel to rule out the possibility of DNA contamination.

A.9 Quantification:

The quantification was performed with gel densitometry as described in (El Kaderi et al., 2009).

A.10 Preparation of soluble and chromatin-bound protein extracts

The Rat1-TAP and an isotype no-tag control strains were cultured in YPD media to an A_{600} of 4. Two liters of each cultures were harvested by centrifugation at 5000 rpm for 5 minutes at 4°C. Cell pellets were washed with 1xTBS and resuspended in 15 ml of lysis buffer. The resultant cell suspension was added dropwise to liquid nitrogen filled 50 ml conical tube and stored at -80 °C.

Lysis of the frozen cell beads was performed in Retsch MM301 mixer mill. The mixer mill chamber and stainless-steel ball were chilled in liquid nitrogen before use. The frozen cell beads and stainless-steel ball were transferred to chambers, secured properly and placed on the holder on the mixer mill. The mixer mill was set for 3 mins cycles at 15 Hz. Lysis was performed by repeating this cycle for 15 times. After each cycle, the chambers were unfastened from the holder and submerged into liquid nitrogen to ensure that frozen cells are not thawing. After completion of 15 lysis cycles, lysed cells were scrapped from the chambers and transferred to a 50 ml conical tube. An additional 15 ml of lysis buffer was added to obtain a paste of lysed yeast cells. This suspension was centrifuged at 5000 rpm for 10 mins at 4°C to remove cell debris. Supernatant represents the whole cell extract. The segregation of soluble extract and chromatin-bound protein extract from the whole cell extract was performed with slight modification of the procedure described in Svejstrup et al., (2003). The whole cell extract was centrifuged in Ti45 rotor at 42,000 rpm for 90 mins at 4°C. The supernatant is carefully transferred to a new tube and labelled as ‘soluble fraction’ and stored at -80°C until further use. The grayish brown

pellet is resuspended in 15 ml of lysis buffer by gentle pipetting without introducing air bubbles. 2mM $CaCl_2$ and 300 U DNase was added, and the sample was incubated on ice for 2 hours. After DNase digestion, samples were centrifuged in Ti45 rotor at 42,000 rpm for 90 minutes at 4°C. The supernatant composed of 'chromatin fraction' was frozen at -80°C until further use. The pellet is composed of insolubilized matter and is discarded.

A.11 Tandem Affinity Purification

The Rat1 complex was purified as described in Puig et al., 2001 with slight modifications. Immunoprecipitation was carried out in the Poly-Prep® Chromatography Columns (BIORAD). In the first affinity purification step, 600 µl of IgG Sepharose® 6 Fast Flow bead suspension was taken into a column and washed once with 10 ml of IP binding buffer. The soluble or chromatin fraction from the step described above in A.10 is added to the column and incubated with the beads on a nutator for 3 hours at 4°C. Washing and elution steps were performed by gravity flow. The unbound sample was drained and a 100 µl aliquot of this sample is set aside as flow through. Protein and its interacting partners bound to the beads are washed thrice with 10 ml wash buffer to ensure that any unbound proteins are not carried forward in the analysis. A wash with 10 ml of TEV cleavage buffer was followed to equilibrate beads with the buffer for TEV cleavage step. Cleavage at TEV site in TAP tag was carried out using 160 units of TEV protease (MC Labs) in 1 ml of TEV cleavage buffer. The column was incubated at 16°C for 2 hrs. The eluate from this enzymatic elution step is collected into a new tube. An additional 500 µl of TEV cleavage buffer was added to collect 1.5 ml of the total eluate from this step.

300 µl of Calmodulin-Sepharose 4B (GE healthcare) bead suspension was equilibrated with 10 ml of calmodulin binding buffer. To the previously collected TEV

eluate, 4.5 ml of calmodulin binding buffer and 4 mM $CaCl_2$ was added. This sample is then transferred to a column that contains equilibrated Calmodulin-Sepharose 4B beads. The binding step was carried out for 2 hours at 4°C. Unbound proteins were drained from the column under gravity and 500 µl was set aside as flow through from calmodulin binding step. Protein captured on Calmodulin-Sepharose 4B were washed thrice with 10 ml of Calmodulin binding buffer. The bound proteins and their associated partners are eluted in five fractions of 200 µl using calmodulin elution buffer.

A.12 TCA precipitation

Trichloroacetic acid (TCA) precipitation is a method used to concentrate proteins in a dilute solution. To the eluate fractions from the Calmodulin-Sepharose 4B column, TCA was added to a final concentration of 30%. The sample was incubated at 4°C for 2 hours on ice. Samples were then centrifuged at 15,700x g for 30 minutes at 4°C to pellet proteins. The supernatant was carefully removed, and any traces of TCA were removed by spinning again at 15,700x g for 1 minute at 4°C. The TCA precipitated protein pellet was washed with 500 µl of 100% acetone. The acetone washed pellet was air-dried at room temperature for 10 minutes and then resuspended in 40 µl of resuspension buffer (100 mM Tris-HCl pH 8.8, 1% SDS).

A.13 SDS -PAGE and WESTERN BLOTTING

Protein samples containing 1x Laemmli buffer were heated at 95°C for 5 mins prior to loaded onto 8% or 10% polyacrylamide gels. The proteins were separated based on the molecular weight by SDS-PAGE at 220 volts. Proteins from the gel were transferred to a polyvinylidene difluoride (PVDF) membrane using a Trans-Blot Turbo Transfer System (BIO-RAD). Following transfer, the membrane was submerged in 100% methanol

and then dried at room temperature. The non-specific protein binding sites on the membrane were blocked by incubating in 5% nonfat dry milk in 1xTBST for 1 hr. The membrane was probed for the bait protein with specific primary antibody for 1hr at room temperature. Primary antibody solution was prepared in 1xTBST containing 5% nonfat dry milk as per the manufacturer recommended dilutions. Following this step, membrane was washed thrice with 1xTBST for 5 mins. A secondary antibody conjugated to HRP-enzyme that can detect primary antibody was incubated with the membrane for 1 hr at room temperature. Secondary antibody was prepared in 1xTBST, 5 % non-fat dry milk with 3 µl secondary antibody. The membrane was then washed thrice with 30 ml of 1x TBST for 5 mins. Each membrane was then covered in SuperSignal West Pico Chemiluminescent Substrate (Thermo Scientific) for 5 mins. The membrane was removed, placed in plastic wrap and sealed. The blots were exposed to X-ray films for different time and the signals were recorded.

A.14 Silver Staining

Silver staining was performed to visualize protein bands present in the eluate fraction after the second affinity purification step in the Tandem Affinity Purification approach (TAP). The enrichment of the bait protein, Rat1, and its associated interacting partners were monitored after each step in the two steps of affinity purification. The electrophoresed gel was fixed in 150 ml solution of 50% methanol and 5% acetic acid by gently shaking the gel at room temperature for 20 min. The gel was then washed with 150 ml 50% methanol, followed by washing with water for 10 mins. Both steps were carried out on a shaker at room temperature. The gel was sensitized in a solution containing 0.02% sodium thiosulfate for 1 min at room temperature. The gel was rinsed for 1 min

with 150 ml of water and incubated in 0.1% silver nitrate and 0.08% formaldehyde for 20 mins on a shaker. The gel was further rinsed twice with 150 ml of water to remove unbound silver nitrate. The protein bands were visualized using a developer solution that comprised of 2% sodium carbonate and 0.04% formalin (37%). The developer was replaced once solution turns yellow. The bands were allowed to develop until the desired intensity was obtained. The developing reaction was stopped by washing the gel for 10 min with 150 ml 5% acetic acid. The gel was washed with water once before recording the image on ChemiDoc MP (*BIO-RAD*). Three independent grown cultures WT Rat1-TAP (test) and WT no-tag (control) were processed and tested using silver staining and western blot analysis. Two preparations were removed from the screen as the fraction obtained from second affinity purification step were not homogenous in their purity determined by silver staining.

A.15 Mass Spectrometry analysis

Three fractions of the proteins eluted from the Calmodulin-Sepharose column were pooled together and submitted to the proteomic core facility at Wayne State University. The processing of samples from the WT Rat1-TAP (test) and WT no-tag (control) samples were performed at the proteomic facility as described below. Further, in order to gauge the background contamination due to sample processing, a background (bkg) tube without protein was processed in parallel with the sample tubes. Samples were precipitated with two volumes of ice-cold 100% methanol in 1 mM acetic acid at -20°C overnight. The next day, samples were spun at 17,000x g for 20 min at 4°C and the supernatant was removed. The resulting pellets were rinsed with 100 µl methanol-acetic acid. Samples were dried in Speed-Vac for 3 min. The pellets were solubilized in 50 mM

triethylammonium bicarbonate (TEAB) using Qsonica sonicator, then reduced with 5 mM DTT and alkylated with 15 mM iodoacetamide (IAA) under standard conditions. Excess IAA was quenched with an additional 5 mM DTT. The samples were then digested overnight in 50 mM TEAB with sequencing-grade trypsin (Promega). The next day, digests were acidified with 1% formic acid and a 10% aliquot of the supernatant was analyzed.

The peptides were separated by reversed-phase chromatography (Acclaim PepMap100 C18 column, Thermo Scientific), followed by ionization with the Nanospray Flex Ion Source (Thermo Scientific), and introduced into a Q Exactive™ mass spectrometer (Thermo Scientific). Abundant species were fragmented with high-energy collision-induced dissociation (HCID). Data analysis was performed using Proteome Discoverer 2.1 (Thermo Scientific) which incorporated the SEQUEST algorithm (Thermo Scientific). The Uniprot_Yeast_Compl_20160407 database was searched for yeast protein sequences and a reverse decoy protein database was run simultaneously for false discovery rate (FDR) determination. The data files were loaded into Scaffold (Proteome Software) for distribution. SEQUEST was searched with a fragment ion mass tolerance of 0.02 Da and a parent ion tolerance of 10 PPM. Carbamidomethylation of cysteine was specified in SEQUEST as a fixed modification. Deamidation of asparagine and glutamine, oxidation of methionine, and acetylation of the N-terminus were specified in SEQUEST as variable modifications. Minimum protein identification probability was set at <= 1.0% FDR with 2 unique peptides at <= 1.0% FDR minimum peptide identification probability. (0.5% protein decoy FDR, 0.16% peptide decoy FDR).

A.16 Spot Assay

YFR1321 and YF1321 Rat1-FRB strains were incubated overnight at 30ºC to saturation and then diluted to the same A_{600} (0.5 – 1.0). Serial dilutions were then made before being spotted on drug-free or 1 µg/µl rapamycin YPD agar plates. These plates were incubated for 48 hours at 30°C. The observation of the growth was recorded using Kodak Gel Logic 200 system.

APPENDIX B: LIST OF STRAINS

Strain	Genotype	Reference
FY23	*MATa ura3-52 trp1Δ63 leu2Δ1*	[Winston et al. Yeast 11:53-56 (1995)]
rat1-1	*MATa ura3-52 leu2Δ1 trplΔ163 rat1-1*	Amberg et al., 1992
BY4733	*MATa his3Δ200 trp1Δ63 leu2Δ0 met15Δ0 ura3Δ0*	
rna14-1	*MATa ura3-1 trp1-1 ade2-1 leu2-3,112 his3-11,15 rna14-1*	
ZA1	*MATa ura3-52 trp1Δ63 leu2Δ1 rtt103Δ::TRP1*	This study
ZA2	*MATa ura3-52 trp1Δ63 leu2Δ1 rai1Δ::TRP1*	This study
ZD7	*MATa ura3-52 trp1Δ63 leu2Δ1 RAT1-TAP::TRP1*	This study
HHY168	*MATα ade2-1 trp1-1 can1-100 leu2-3,112 his3-11,15 ura3 tor1-1 fpr1::NAT RPL13A-2xFKBP12::TRP1*	Haruki et al., 2008
ZD42	*MATα ade2-1 trp1-1 can1-100 leu2-3,112 his3-11,15 ura3 tor1-1 fpr1::NAT RPL13A-2xFKBP12::TRP1 RAT1-FRB::kanMX6 [pFA6a-FRB-KanMX6]*	This study
ZD48	*MATα ade2-1 trp1-1 can1-100 leu2-3,112 his3-11,15 ura3 tor1-1 fpr1::NAT RPL13A-2xFKBP12::TRP1 RAT1-FRB::kanMX6 [pRS415, CEN LEU2 RAT1-WT]*	This study
ZD49	*MATα ade2-1 trp1-1 can1-100 leu2-3,112 his3-11,15 ura3 tor1-1 fpr1::NAT RPL13A-2xFKBP12::TRP1 RAT1-FRB::kanMX6 [pRS415, CEN LEU2 RAT1-D235A]*	This study

ZD47	*MATα ade2-1 trp1-1 can1-100 leu2-3,112 his3-11,15 ura3 tor1-1 fpr1::NAT RPL13A-2xFKBP12::TRP1 RAT1-FRB::kanMX6 RRP6KO::HIS3*	This study
ZD63	*MATα ade2-1 trp1-1 can1-100 leu2-3,112 his3-11,15 ura3 tor1-1 fpr1::NAT RPL13A-2xFKBP12::TRP1 RAT1-FRB::kanMX6 XRN1KO::HIS3*	This study
ZD64	*MATα ade2-1 trp1-1 can1-100 leu2-3, 112 his3-11, 15 ura3 tor1-1 fpr1::NATRPL113A-2xFKBP1112::TRP1 PRP2-13MYC::HIS*	This study

APPENDIX C: LIST OF PRIMERS

C.1. RT Primers

MRK1-F	TGCAGTTGATCACATTGAAC
MRK1-R	TACAACACCAAATGAACCATG
NMD2-F	TACATTGGACAGAAATTATGG
NMD2-F	ACCGTCACTATTATCTCTGATAG
APE2-F	AATAGACGCCCTCTGTTCACAG
APE2-R	GGTGTTCAATGTTACCGTATCA
STO- F	GAAGGACAGAAGGTATATTGC
STO1-R	ACCAATAGTGCGAATATCGG
LSB4-F	TGGGTATCAATAATCCAATTCC
LSB4-R	TGCTCTCCCTGAAAACAAG
RPS4B-F	AGATACACCACTATTGAGG
RPS4B-R	TGTGGACCAGCAGATGGTC
RPS14A-F	ACAAGAACCCGCCATGTC
RPS14A-R	ACCACCAGTAACTCTGGC
RPS9B-F	ACTAAGCAACAATGCCAAG
RPS9B-R	ACAAGTCTCTGGCAGCAC
RPS11B-F	ACTGTTCAATCTGAAAGAGC
RPS11-R	TACGGTGCATCTTGGTGG
5S rRNA-F	GGTTGCGGCCATATCTAC
5S rRNA-R	TGAGTTTCGCGTATGGTC
Oligo dT	TTTTTTTTTTTTTTTTTTTTTTTT

5S rRNA cDNA	TGAGTTTCGCGTATGGTC

C.2 ChIP primers

5′ APE2 A	GGCTTGGAACACCTAAGAAGA
3′ APE2 A	GATCGAAGTACCGTAGCACATC
5′ APE2 B	TCTAATAGGCAAAGCTCACC
3′ APE2 B	ACAGGCGTGACTAGTTTC
5′ APE2 C	AGCTTAACAAAGAAACAAGG
3′ APE2 C	AACTCAAAGGTCGGTAAC
5′ APE2 D	CTCTGGCTTTACATCAATGC
3′ APE2 D	TCAAATACTTGTTGACGACATC
5′ APE2 E	ATTTTGCCTTTTTATATAGTCAAGT
3′ APE2 E	TGTGCACGGGTTGAAATAAATG

C.3 Primers used in transformation

5′ RAT1-C-TAP	CTCGGAATAACAAGCAAAGTCGGTATGACAATTCAAGAGC AAATAGGCGTTCCATGGAAAAGAGAAGATG
3′ RAT1-C-TAP	AATTTGCGAAAACCTAAATTTACCATAAAATAAAATGCGCA CGAGTAGTTTAATACGACTCACTATAGGG
5′ RAT1-FRB	CTCGGAATAACAAGCAAAGTCGGTATGACAATTCAAGAGC AAATAGGCGT GGATCCCCGGGTTAATTAAC
3′ RAT1-FRB	AATTTGCGAAAACCTAAATTTACCATAAAATAAAATGCGCA CGAGTAGTT GAATTCGAGCTCGTTTAAAC

5' RTT103-KO	AGAGGTAGAAAATTTGAAGAAAGCAATAATCCAAGATTAAA ATAGACGGTGCGGCATCAGAGCAGATTG
3' RTT103-KO	ATATATTTGTATAAGTTATCTCCTTGTTTTCTTTTTACTCAA CCATCATACTGTGCGGTATTTCACACCG
5' RAI1-KO	GTAATATGGTGAAAGAATAGCGAAATATTAGACCAACATA GTGTATCCCATGTACTGAGAGTGCACCATA
3' RAI1-KO	GATCCATACGTCGATGAGGATATGCGCAGGAAAGACATAA AGGAATATTGGCAAGTGCACAAACAATACT
5' RRP6-KO	TAGACGAAATAGGAACAACAAACAGCTTATAAGCACCCAA TAAGTGCGTTGCATCAGAGCAGATTGTACTGAG
3' RRP6-KO	ATGAAAATTACCATAATTTATAAATAAAAAAATACGCTTGTT TTACATAATGTGCGGTATTTCACACCG
5' XRN1-KO	ATCAACACTTGTAACAACAGCAGCAACAAATATATATCAGT ACGGTGCATCAGAGCAGATTGTACTGAG
3' XRN1-KO	GATATACTATTAAAGTAACCTCGAATATACTTCGTTTTTAGT CGTATGTTTGTGCGGTATTTCACACCG
5' PRP2-MYC	CACAAATCTTTAAAGATTTAATTGACGATAAAACAAATAGG GGGAGGCGGCGGATCCCCGGGTTAATTAA
3' PRP2-MYC	AGAATGGAGCCTGCGTTTCTAGCAATACACATACACCTGT CAAAAAACCTGAATTCGAGCTCGTTTAAAC

C.4 Primers used in identifying correct transformants

RAT1-F	ATAATGTCCAACCCGCCC

RAT1-R	ATAGAGTTCGGTCTATTGGC
3' UNIVERSAL TAP	GTTGAATTTGTTGTCTACTTTCGG
3' UNIVERSAL FRB	TTGGAGGAGGTCCTTGAC
RTT103-F	TCGGTCTTTCCACTCCCTTG
RTT103-R	TGCATAGGTGTCCATGTAG
RAI1-F	GAGTAAGCATTCGGGTAAATTG
RAI1-R	TCGGTTCGCCATGCATAG
XRN1-F	GGATACTGTCTTCTTCCGTAC
XRN1-R	GGACGGTGTCCACAGATC
RRP6-F	CAAAAATATGAGGGCATCGG
RRP6-R	CACGCACAATATTCGAGCG
3' pRS 414	GAATCTAGAGCACATTCTGC
3' pRS 313	ACCCTATACCTGTGTGGAC
PRP2-F	TGATGCTAACGTCGAAGG
PRP2-R	AGAGAGACGTCTTGGATG

APPENDIX D: BUFFERS AND SOLUTION

D1. Buffers for Transcription run-on assay

Buffer	Concentration	Composition
TMN buffer	10 mM 5 mM 100 mM	Tris-HCl pH7.5 NaCl $MgCl_2$
Run-On buffer	50 mM 100 mM 10 mM 0 mM 0.75 mM	Tris-HCl pH7.5 KCl $MgCl_2$ DTT ATP, CTP, GTP and Br-UTP RNase Inhibitor
Binding Buffer	0.25X 1 mM 0.05% 37.5mM	SSPE buffer EDTA Tween NaCl
Blocking Buffer	1X 0.1% 1 mg/ml	Binding buffer PVP Ultra pure BSA
Low Salt Buffer	0.2X 1 mM 0.05%	SSPE buffer EDTA Tween 20
High Salt Buffer	0.25X 1 mM 0.05% 100 mM	SSPE buffer EDTA Tween 20 NaCl
Elution Buffer	20mM 150mM 50mM	DTT NaCl Tris-HCl pH 7.5

	1mM 0.1%	EDTA SDS Rnase Inhibitor
TET Buffer	1X 0.05%	Tris-EDTA (TE) Tween 20

D2. Buffers for Chromatin immunoprecipitation

Buffer	Concentration	Composition
FA lysis	50 mM 140 mM 1 mM 1 % 0.1 % 1 mM	HEPES-KOH (pH 7.9) NaCl EDTA Triton X-100 (v/v) Sodium Deoxycholate PMSF
FA Lysis Buffer + 500 mM NaCl	50 mM 1 % SDS 10 mM	Tris-HCl (pH 7.5 to 8) SDS EDTA pH 8.0
ChIP Wash Buffer	10 mM 250 mM 0.5 % 1 mM 0.5 % 0.1 %	Tris-HCl (pH 7.5 to 8) LiCl Triton X-100 (v/v) EDTA pH 8.0 Sodium Deoxycholate SDS
ChIP Elution Buffer	50 mM 1 % 10 mM	Tris-HCl (pH 7.5 to 8) SDS EDTA pH 8.0

D3. Buffers for RNA immunoprecipitation

Buffer	Concentration	Composition
RIP-Binding Buffer	20 mM	HEPES-KOH

	10 mM	$MgCl_2$
	150 mM	NaCl
	0.5 mM	DTT
		RNase inhibitor
RIP-Elution buffer	20 mM	HEPES-KOH
	1%	SDS
	2 mM	EDTA

D3. Buffers for Tandem Affinity purification

Buffer	Concentration	Composition
Lysis Buffer	20 mM	HEPES-KOH
	10 mM	$MgCl_2$
	0.5 mM	DTT
	1 mM	PMSF
	150 mM	NaCl
	10%	Glycerol (v/v)
Wash Buffer	20 mM	HEPES-KOH
	150 mM	NaCl
	10 mM	$MgCl_2$
	0.5 mM	DTT
	1 mM	PMSF
	10%	Glycerol (v/v)
1x TEV wash buffer	20 mM	Tris-HCl pH 8.0
	150 mM	NaCl
Calmodulin Binding buffer	50 mM	Tris-HCl pH 8.0
	150 mM	NaCl
	1 mM	$(CH_3COO)_2Mg$
	1 mM	Immidazole
	2 mM	β-mercaptoethanol
	3 mM	$CaCl_2$

Calmodulin Elution buffer	50 mM 150 mM 1 mM 2 mM 3 mM	Tris-HCl pH 8.0 NaCl $(CH_3COO)_2Mg$ β-mercaptoethanol EGTA

D4. Buffers for Western Blot

Buffer	Concentration	Composition
TGS buffer	25 mM 250 mM 0.1 %	Tris base Glycine SDS (w/v)
5x Laemeli buffer	250 mM 50 % 10 % 0.02 % 2.8 M	Tris-HCl (pH 6.8) Glycerol (v/v) SDS (w/v) (w/v) Bromophenol blue β-mercaptoethanol
30% acrylamide:bis Solution	29.22 % 0.78 %	Acrylamide Bisacrylamide
Resolving Gel (10%)	3.3 ml 2.5 ml 0.1 ml 0.1 ml 6 µl	30% acrylamide:bis solution 1.5 mM Tris-HCl (pH8.8) 10% SDS 10% Ammonium persulfate TEMMED Add water to make final volume 10 ml
Resolving Gel (8%)	2.7 ml 2.5 ml 0.1 ml 0.1 ml 6 µl	30% acrylamide:bis solution 1.5 mM Tris-HCl (pH8.8) 10% SDS 10% Ammonium persulfate TEMMED Add water to make final volume 10 ml

Stacking gel	125 mM 0.1% 5% 0.1	acrylamide:bis solution Tris-HCl (pH6.8) SDS Ammonium persulfate
Anode I buffer	300 mM 10 %	Tris base Methanol Adjust pH to 10.4
Anode II buffer	25 mM 10 %	Tris base Methanol Adjust pH to 10.4
Cathode buffer	25 mM 40 mM 10 %	Tris base Glycine Methanol Adjust pH to 9.4

D5. Buffers for Transformation

Buffer	Concentration	Composition
LiAc buffer	0.1M 10 mM 1 mM	LiAc Tris-HCl (pH=8.0) EDTA
PEG solution	50 % 0.1M 10 mM 1 mM	PEG (M.W. = 4000) (w/v) LiAOAc Tris-HCl (pH=8.0) EDTA

D6. Solutions

Buffer	Concentration	Composition

TBS buffer (10x)	200 mM 2 M	Tris-HCl (pH 7.5) NaCl For 1X solution, mix 1 part of 10X with 9 parts ddH_2O and adjust to pH=7.5
1X TBS with Triton-X 100	1X 1%	TBS Triton-X 100
TE buffer (10x)	100 mM 10 mM	Tris-HCl (pH 8) EDTA pH 8.0 For 1X solution, mix 1 part of 10X with 9 parts ddH_2O and adjust to pH= 8

D7. Agarose gel electrophoresis

Buffer	Concentration	Composition
Tris-acetate (TAE)	40 mM 1 mM	Tris-acetate EDTA (pH 8.0)
0.8 or 1.5% Agarose	0.8% or 1.5%	Agarose in 1x TAE

APPENDIX E: MEDIA

Media	Concentration	Composition
YPD medium	1 %	Yeast extract (w/v)
	2 %	Peptone (w/v)
	2 %	Dextrose (w/v)
for YPD agar	plus 2 %	Agar (w/v)
YPD agar with Rapamycin	1 µg/ml	YPD agar as mentioned above plus Rapamycin
YPD agar with G418	400 µg/ml	YPD agar as mentioned above plus G-418
2X YT medium	1 %	Yeast extract (w/v)
	1.6 %	Tryptone (w/v)
	2 %	NaCl (w/v)
for 2X YT agar	2 %	Agar (w/v)
2X YT agar with Ampicillin	100 µg/ml	2X YT medium as mentioned above plus Ampicillin
2X YT agar with Kanamycin	50 µg/ml	2X YT medium as mentioned above plus Kanamycin
Minimal medium	6.9 g/l	Yeast Nitrogen Base
	1 g/L	Amino acid drop-out mix
	2 %	Dextrose
	2 %	Agar
Amino acid drop-out mix		Adenine, Uracil 20 mg/l L-Methionine, L-Histidine HCl, L-Methionine

		50 mg/l L-Arginine, L-Isoleucine, L-Lysine HCl, L-Tryptophan, L-Tyrosine, L-Phenylalanine 80 mg/l L-Aspartic acid 100 mg/l L-Leucine, L-Threonine, 4-Thiouracil 140 mg/l Valine 2 % Glucose (w/v)

REFERENCES

Albers, M., Diment, A., Muraru, M., Russell, C.S., and Beggs, J.D. (2003). Identification and characterization of Prp45p and Prp46p, essential pre-mRNA splicing factors. RNA *9*, 138–150.

Aldrich, T.L., Di Segni, G., McConaughy, B.L., Keen, N.J., Whelen, S., and Hall, B.D. (1993). Structure of the yeast TAP1 protein: dependence of transcription activation on the DNA context of the target gene. Mol. Cell. Biol. *13*, 3434–3444.

de Almeida, R.A., and O'Keefe, R.T. (2015). The NineTeen Complex (NTC) and NTC-associated proteins as targets for spliceosomal ATPase action during pre-mRNA splicing. RNA Biol *12*, 109–114.

Amberg, D.C., Goldstein, A.L., and Cole, C.N. (1992). Isolation and characterization of RAT1: an essential gene of Saccharomyces cerevisiae required for the efficient nucleocytoplasmic trafficking of mRNA. Genes Dev. *6*, 1173–1189.

Arenas, J.E., and Abelson, J.N. (1997). Prp43: An RNA helicase-like factor involved in spliceosome disassembly. PNAS *94*, 11798–11802.

Ares, M., Grate, L., and Pauling, M.H. (1999). A handful of intron-containing genes produces the lion's share of yeast mRNA. RNA *5*, 1138–1139.

Asturias, F.J., Jiang, Y.W., Myers, L.C., Gustafsson, C.M., and Kornberg, R.D. (1999). Conserved structures of mediator and RNA polymerase II holoenzyme. Science *283*, 985–987.

Baejen, C., Andreani, J., Torkler, P., Battaglia, S., Schwalb, B., Lidschreiber, M., Maier, K.C., Boltendahl, A., Rus, P., Esslinger, S., et al. (2017). Genome-wide Analysis of RNA Polymerase II Termination at Protein-Coding Genes. Molecular Cell *66*, 38-49.e6.

Baptista, T., Grünberg, S., Minoungou, N., Koster, M.J.E., Timmers, H.T.M., Hahn, S., Devys, D., and Tora, L. (2017). SAGA Is a General Cofactor for RNA Polymerase II Transcription. Mol. Cell *68*, 130-143.e5.

Barrass, J.D., Reid, J.E.A., Huang, Y., Hector, R.D., Sanguinetti, G., Beggs, J.D., and Granneman, S. (2015). Transcriptome-wide RNA processing kinetics revealed using extremely short 4tU labeling. Genome Biol *16*, 282.

Bataille, A.R., Jeronimo, C., Jacques, P.-É., Laramée, L., Fortin, M.-È., Forest, A., Bergeron, M., Hanes, S.D., and Robert, F. (2012). A universal RNA polymerase II CTD cycle is orchestrated by complex interplays between kinase, phosphatase, and isomerase enzymes along genes. Mol. Cell *45*, 158–170.

Belotserkovskaya, R., Oh, S., Bondarenko, V.A., Orphanides, G., Studitsky, V.M., and Reinberg, D. (2003). FACT facilitates transcription-dependent nucleosome alteration. Science *301*, 1090–1093.

Belshaw, P.J., Ho, S.N., Crabtree, G.R., and Schreiber, S.L. (1996). Controlling protein association and subcellular localization with a synthetic ligand that induces heterodimerization of proteins. PNAS *93*, 4604–4607.

Bienroth, S., Wahle, E., Suter-Crazzolara, C., and Keller, W. (1991). Purification of the cleavage and polyadenylation factor involved in the 3'-processing of messenger RNA precursors. J. Biol. Chem. *266*, 19768–19776.

Bousquet-Antonelli, C., Presutti, C., and Tollervey, D. (2000). Identification of a regulated pathway for nuclear pre-mRNA turnover. Cell *102*, 765–775.

Braberg, H., Jin, H., Moehle, E.A., Chan, Y.A., Wang, S., Shales, M., Benschop, J.J., Morris, J.H., Qiu, C., Hu, F., et al. (2013). From Structure to Systems: High-Resolution, Quantitative Genetic Analysis of RNA Polymerase II. Cell *154*, 775–788.

Brannan, K., Kim, H., Erickson, B., Glover-Cutter, K., Kim, S., Fong, N., Kiemele, L., Hansen, K., Davis, R., Lykke-Andersen, J., et al. (2012). mRNA decapping factors and the exonuclease Xrn2 function in widespread premature termination of RNA polymerase II transcription. Mol Cell *46*, 311–324.

Brody, E., and Abelson, J. (1985). The "spliceosome": yeast pre-messenger RNA associates with a 40S complex in a splicing-dependent reaction. Science *228*, 963–967.

Brow, D.A. (2002). Allosteric cascade of spliceosome activation. Annu. Rev. Genet. *36*, 333–360.

Brune, C., Munchel, S.E., Fischer, N., Podtelejnikov, A.V., and Weis, K. (2005). Yeast poly(A)-binding protein Pab1 shuttles between the nucleus and the cytoplasm and functions in mRNA export. RNA *11*, 517–531.

Buratowski, S., and Kim, T. (2010). The role of cotranscriptional histone methylations. Cold Spring Harb. Symp. Quant. Biol. *75*, 95–102.

Bürckstümmer, T., Bennett, K.L., Preradovic, A., Schütze, G., Hantschel, O., Superti-Furga, G., and Bauch, A. (2006). An efficient tandem affinity purification procedure for interaction proteomics in mammalian cells. Nat. Methods *3*, 1013–1019.

Bushnell, D.A., and Kornberg, R.D. (2003). Complete, 12-subunit RNA polymerase II at 4.1-A resolution: implications for the initiation of transcription. Proc. Natl. Acad. Sci. U.S.A. *100*, 6969–6973.

Butler, J.S., and Platt, T. (1988). RNA processing generates the mature 3' end of yeast CYC1 messenger RNA in vitro. Science *242*, 1270–1274.

Candelli, T., Challal, D., Briand, J.-B., Boulay, J., Porrua, O., Colin, J., and Libri, D. (2018). High-resolution transcription maps reveal the widespread impact of roadblock termination in yeast. EMBO J. *37*.

Carrillo Oesterreich, F., Preibisch, S., and Neugebauer, K.M. (2010). Global analysis of nascent RNA reveals transcriptional pausing in terminal exons. Mol. Cell *40*, 571–581.

Chalamcharla, V.R., Folco, H.D., Dhakshnamoorthy, J., and Grewal, S.I.S. (2015). Conserved factor Dhp1/Rat1/Xrn2 triggers premature transcription termination and nucleates heterochromatin to promote gene silencing. Proc. Natl. Acad. Sci. U.S.A. *112*, 15548–15555.

Chan, S.-P., and Cheng, S.-C. (2005). The Prp19-associated complex is required for specifying interactions of U5 and U6 with pre-mRNA during spliceosome activation. J. Biol. Chem. *280*, 31190–31199.

Chanarat, S., and Sträßer, K. (2013). Splicing and beyond: The many faces of the Prp19 complex. Biochimica et Biophysica Acta (BBA) - Molecular Cell Research *1833*, 2126–2134.

Chanfreau, G., Rotondo, G., Legrain, P., and Jacquier, A. (1998). Processing of a dicistronic small nucleolar RNA precursor by the RNA endonuclease Rnt1. EMBO J. *17*, 3726–3737.

Chang, W.-H., and Kornberg, R.D. (2000). Electron Crystal Structure of the Transcription Factor and DNA Repair Complex, Core TFIIH. Cell *102*, 609–613.

Chang, J.H., Jiao, X., Chiba, K., Oh, C., Martin, C.E., Kiledjian, M., and Tong, L. (2012). Dxo1 is a new type of eukaryotic enzyme with both decapping and 5’-3’ exoribonuclease activity. Nat. Struct. Mol. Biol. *19*, 1011–1017.

Chatterjee, S., and Grosshans, H. (2009). Active turnover modulates mature microRNA activity in Caenorhabditis elegans. Nature *461*, 546–549.

Chen, J., Zheng, X.F., Brown, E.J., and Schreiber, S.L. (1995). Identification of an 11-kDa FKBP12-rapamycin-binding domain within the 289-kDa FKBP12-rapamycin-associated protein and characterization of a critical serine residue. PNAS *92*, 4947–4951.

Chen, Y.-I.G., Moore, R.E., Ge, H.Y., Young, M.K., Lee, T.D., and Stevens, S.W. (2007). Proteomic analysis of in vivo-assembled pre-mRNA splicing complexes expands the catalog of participating factors. Nucleic Acids Res *35*, 3928–3944.

Chereji, R.V., Bharatula, V., Elfving, N., Blomberg, J., Larsson, M., Morozov, A.V., Broach, J.R., and Björklund, S. (2017). Mediator binds to boundaries of chromosomal interaction domains and to proteins involved in DNA looping, RNA metabolism, chromatin remodeling, and actin assembly. Nucleic Acids Res *45*, 8806–8821.

Chernyakov, I., Whipple, J.M., Kotelawala, L., Grayhack, E.J., and Phizicky, E.M. (2008). Degradation of several hypomodified mature tRNA species in Saccharomyces cerevisiae is mediated by Met22 and the 5’-3’ exonucleases Rat1 and Xrn1. Genes Dev. *22*, 1369–1380.

Cho, E.J., Takagi, T., Moore, C.R., and Buratowski, S. (1997). mRNA capping enzyme is recruited to the transcription complex by phosphorylation of the RNA polymerase II carboxy-terminal domain. Genes Dev. *11*, 3319–3326.

Chung, S., McLean, M.R., and Rymond, B.C. (1999). Yeast ortholog of the Drosophila crooked neck protein promotes spliceosome assembly through stable U4/U6.U5 snRNP addition. RNA *5*, 1042–1054.

Combs, D.J., Nagel, R.J., Ares, M., and Stevens, S.W. (2006). Prp43p is a DEAH-box spliceosome disassembly factor essential for ribosome biogenesis. Mol. Cell. Biol. *26*, 523–534.

Connelly, S., and Manley, J.L. (1988). A functional mRNA polyadenylation signal is required for transcription termination by RNA polymerase II. Genes Dev. *2*, 440–452.

Corden, J.L. (1990). Tails of RNA polymerase II. Trends Biochem. Sci. *15*, 383–387.

Corden, J.L. (2013). RNA polymerase II C-terminal domain: Tethering transcription to transcript and template. Chem. Rev. *113*, 8423–8455.

Corden, J.L., Cadena, D.L., Ahearn, J.M., and Dahmus, M.E. (1985). A unique structure at the carboxyl terminus of the largest subunit of eukaryotic RNA polymerase II. Proc. Natl. Acad. Sci. U.S.A. *82*, 7934–7938.

Cordin, O., and Beggs, J.D. (2013). RNA helicases in splicing. RNA Biol *10*, 83–95.

Core, L.J., Waterfall, J.J., and Lis, J.T. (2008). Nascent RNA sequencing reveals widespread pausing and divergent initiation at human promoters. Science *322*, 1845–1848.

Coskun, O. (2016). Separation techniques: Chromatography. North Clin Istanb *3*, 156–160.

Costanzo, M., VanderSluis, B., Koch, E.N., Baryshnikova, A., Pons, C., Tan, G., Wang, W., Usaj, M., Hanchard, J., Lee, S.D., et al. (2016). A global genetic interaction network maps a wiring diagram of cellular function. Science *353*.

Crick, F.H. (1958). On protein synthesis. Symp. Soc. Exp. Biol. *12*, 138–163.

Davidson, L., Kerr, A., and West, S. (2012). Co-transcriptional degradation of aberrant pre-mRNA by Xrn2. EMBO J *31*, 2566–2578.

Davis, J.A., Takagi, Y., Kornberg, R.D., and Asturias, F.A. (2002). Structure of the yeast RNA polymerase II holoenzyme: Mediator conformation and polymerase interaction. Mol. Cell *10*, 409–415.

Deng, W., and Roberts, S.G.E. (2005). A core promoter element downstream of the TATA box that is recognized by TFIIB. Genes Dev. *19*, 2418–2423.

Deng, W., and Roberts, S.G.E. (2007). TFIIB and the regulation of transcription by RNA polymerase II. Chromosoma *116*, 417–429.

Dengl, S., and Cramer, P. (2009). Torpedo Nuclease Rat1 Is Insufficient to Terminate RNA Polymerase II in Vitro. J Biol Chem *284*, 21270–21279.

Dhoondia, Z., Tarockoff, R., Alhusini, N., Medler, S., Agarwal, N., and Ansari, A. (2017). Analysis of Termination of Transcription Using BrUTP-strand-specific Transcription Run-on (TRO) Approach. J Vis Exp.

Di Segni, G., McConaughy, B.L., Shapiro, R.A., Aldrich, T.L., and Hall, B.D. (1993). TAP1, a yeast gene that activates the expression of a tRNA gene with a defective internal promoter. Mol. Cell. Biol. *13*, 3424–3433.

Dieci, G., and Sentenac, A. (1996). Facilitated recycling pathway for RNA polymerase III. Cell *84*, 245–252.

Dong, C., Lin, Z., Diao, W., Li, D., Chu, X., Wang, Z., Zhou, H., Xie, Z., Shen, Y., and Long, J. (2015). The Elp2 subunit is essential for elongator complex assembly and functional regulation. Structure *23*, 1078–1086.

Dong, M., Yang, L.L., Williams, K., Fisher, S.J., Hall, S.C., Biggin, M.D., Jin, J., and Witkowska, H.E. (2008). A “tagless” strategy for identification of stable protein complexes genome-wide by multidimensional orthogonal chromatographic separation and iTRAQ reagent tracking. J. Proteome Res. *7*, 1836–1849.

Donner, A.J., Ebmeier, C.C., Taatjes, D.J., and Espinosa, J.M. (2010). CDK8 is a positive regulator of transcriptional elongation within the serum response network. Nat Struct Mol Biol *17*, 194–201.

Dotson, M.R., Yuan, C.X., Roeder, R.G., Myers, L.C., Gustafsson, C.M., Jiang, Y.W., Li, Y., Kornberg, R.D., and Asturias, F.J. (2000). Structural organization of yeast and mammalian mediator complexes. Proc. Natl. Acad. Sci. U.S.A. *97*, 14307–14310.

Dunn, E.F., Hammell, C.M., Hodge, C.A., and Cole, C.N. (2005). Yeast poly(A)-binding protein, Pab1, and PAN, a poly(A) nuclease complex recruited by Pab1, connect mRNA biogenesis to export. Genes Dev. *19*, 90–103.

Eaton, J.D., Davidson, L., Bauer, D.L.V., Natsume, T., Kanemaki, M.T., and West, S. (2018). Xrn2 accelerates termination by RNA polymerase II, which is underpinned by CPSF73 activity. Genes Dev.

El Hage, A., Koper, M., Kufel, J., and Tollervey, D. (2008). Efficient termination of transcription by RNA polymerase I requires the 5’ exonuclease Rat1 in yeast. Genes Dev. *22*, 1069–1081.

El Kaderi, B., Medler, S., Raghunayakula, S., and Ansari, A. (2009). Gene looping is conferred by activator-dependent interaction of transcription initiation and termination machineries. J. Biol. Chem. *284*, 25015–25025.

Fabrizio, P., Dannenberg, J., Dube, P., Kastner, B., Stark, H., Urlaub, H., and Lührmann, R. (2009). The evolutionarily conserved core design of the catalytic activation step of the yeast spliceosome. Mol. Cell *36*, 593–608.

Fang, F., Phillips, S., and Butler, J.S. (2005). Rat1p and Rai1p function with the nuclear exosome in the processing and degradation of rRNA precursors. RNA *11*, 1571–1578.

Filtz, T.M., Vogel, W.K., and Leid, M. (2014). Regulation of transcription factor activity by interconnected post-translational modifications. Trends in Pharmacological Sciences *35*, 76–85.

Flanagan, P.M., Kelleher, R.J., Sayre, M.H., Tschochner, H., and Kornberg, R.D. (1991). A mediator required for activation of RNA polymerase II transcription in vitro. Nature *350*, 436–438.

Fong, N., Brannan, K., Erickson, B., Kim, H., Cortazar, M.A., Sheridan, R.M., Nguyen, T., Karp, S., and Bentley, D.L. (2015). Effects of Transcription Elongation Rate and Xrn2 Exonuclease Activity on RNA Polymerase II Termination Suggest Widespread Kinetic Competition. Mol. Cell *60*, 256–267.

Frauenfelder, H., Sligar, S.G., and Wolynes, P.G. (1991). The energy landscapes and motions of proteins. Science *254*, 1598–1603.

Gasse, L., Flemming, D., and Hurt, E. (2015). Coordinated Ribosomal ITS2 RNA Processing by the Las1 Complex Integrating Endonuclease, Polynucleotide Kinase, and Exonuclease Activities. Mol. Cell *60*, 808–815.

Gavin, A.-C., Bösche, M., Krause, R., Grandi, P., Marzioch, M., Bauer, A., Schultz, J., Rick, J.M., Michon, A.-M., Cruciat, C.-M., et al. (2002). Functional organization of the yeast proteome by systematic analysis of protein complexes. Nature *415*, 141–147.

Geerlings, T.H., Vos, J.C., and Raué, H.A. (2000). The final step in the formation of 25S rRNA in Saccharomyces cerevisiae is performed by 5'-->3' exonucleases. RNA *6*, 1698–1703.

Ghosh, A., and Lima, C.D. (2010). Enzymology of RNA cap synthesis. Wiley Interdiscip Rev RNA *1*, 152–172.

Giardina, C., and Lis, J.T. (1993). DNA melting on yeast RNA polymerase II promoters. Science *261*, 759–762.

Gibbons, B.J., Brignole, E.J., Azubel, M., Murakami, K., Voss, N.R., Bushnell, D.A., Asturias, F.J., and Kornberg, R.D. (2012). Subunit architecture of general transcription factor TFIIH. Proc Natl Acad Sci U S A *109*, 1949–1954.

González-Aguilera, C., Tous, C., Gómez-González, B., Huertas, P., Luna, R., and Aguilera, A. (2008). The THP1-SAC3-SUS1-CDC31 complex works in transcription elongation-mRNA export preventing RNA-mediated genome instability. Mol. Biol. Cell *19*, 4310–4318.

Grabowski, P.J., Seiler, S.R., and Sharp, P.A. (1985). A multicomponent complex is involved in the splicing of messenger RNA precursors. Cell *42*, 345–353.

Green, M.R. (2005). Eukaryotic transcription activation: right on target. Mol. Cell *18*, 399–402.

Gross, S., and Moore, C. (2001). Five subunits are required for reconstitution of the cleavage and polyadenylation activities of Saccharomyces cerevisiae cleavage factor I. Proc. Natl. Acad. Sci. U.S.A. *98*, 6080–6085.

Grünberg, S., and Hahn, S. (2013). Structural insights into transcription initiation by RNA polymerase II. Trends Biochem. Sci. *38*, 603–611.

Gunasekaran, K., Ma, B., and Nussinov, R. (2004). Is allostery an intrinsic property of all dynamic proteins? Proteins *57*, 433–443.

Hahn, S. (2004). Structure and mechanism of the RNA polymerase II transcription machinery. Nat. Struct. Mol. Biol. *11*, 394–403.

Hahn, S., and Young, E.T. (2011). Transcriptional regulation in Saccharomyces cerevisiae: transcription factor regulation and function, mechanisms of initiation, and roles of activators and coactivators. Genetics *189*, 705–736.

Hardin, J.W., Warnasooriya, C., Kondo, Y., Nagai, K., and Rueda, D. (2015). Assembly and dynamics of the U4/U6 di-snRNP by single-molecule FRET. Nucleic Acids Res *43*, 10963–10974.

Harlen, K.M., and Churchman, L.S. (2017). Subgenic Pol II interactomes identify region-specific transcription elongation regulators. Mol Syst Biol *13*.

Haruki, H., Nishikawa, J., and Laemmli, U.K. (2008). The anchor-away technique: rapid, conditional establishment of yeast mutant phenotypes. Mol. Cell *31*, 925–932.

Henry, Y., Wood, H., Morrissey, J.P., Petfalski, E., Kearsey, S., and Tollervey, D. (1994). The 5’ end of yeast 5.8S rRNA is generated by exonucleases from an upstream cleavage site. EMBO J. *13*, 2452–2463.

Herzel, L., Ottoz, D.S.M., Alpert, T., and Neugebauer, K.M. (2017). Splicing and transcription touch base: co-transcriptional spliceosome assembly and function. Nat. Rev. Mol. Cell Biol. *18*, 637–650.

Hsin, J.-P., and Manley, J.L. (2012). The RNA polymerase II CTD coordinates transcription and RNA processing. Genes Dev. *26*, 2119–2137.

Huisinga, K.L., and Pugh, B.F. (2004). A genome-wide housekeeping role for TFIID and a highly regulated stress-related role for SAGA in Saccharomyces cerevisiae. Mol. Cell *13*, 573–585.

Jasnovidova, O., Krejcikova, M., Kubicek, K., and Stefl, R. (2017). Structural insight into recognition of phosphorylated threonine-4 of RNA polymerase II C-terminal domain by Rtt103p. EMBO Rep. *18*, 906–913.

Jenny, A., Minvielle-Sebastia, L., Preker, P.J., and Keller, W. (1996). Sequence Similarity Between the 73-Kilodalton Protein of Mammalian CPSF and a Subunit of Yeast Polyadenylation Factor I. Science *274*, 1514–1517.

Jeronimo, C., Bataille, A.R., and Robert, F. (2013). The writers, readers, and functions of the RNA polymerase II C-terminal domain code. Chem. Rev. *113*, 8491–8522.

Jiao, X., Xiang, S., Oh, C., Martin, C.E., Tong, L., and Kiledjian, M. (2010). Identification of a quality-control mechanism for mRNA 5′-end capping. Nature *467*, 608–611.

Jiao, X., Chang, J.H., Kilic, T., Tong, L., and Kiledjian, M. (2013). A Mammalian Pre-mRNA 5′ End Capping Quality Control Mechanism and an Unexpected Link of Capping to Pre-mRNA Processing. Molecular Cell *50*, 104–115.

Jimeno-González, S., Haaning, L.L., Malagon, F., and Jensen, T.H. (2010). The yeast 5'-3' exonuclease Rat1p functions during transcription elongation by RNA polymerase II. Mol. Cell *37*, 580–587.

Jimeno-González, S., Schmid, M., Malagon, F., Haaning, L.L., and Jensen, T.H. (2014). Rat1p maintains RNA polymerase II CTD phosphorylation balance. RNA *20*, 551–558.

Jin, D.J., Burgess, R.R., Richardson, J.P., and Gross, C.A. (1992). Termination efficiency at rho-dependent terminators depends on kinetic coupling between RNA polymerase and rho. PNAS *89*, 1453–1457.

Johnson, A.W. (1997). Rat1p and Xrn1p are functionally interchangeable exoribonucleases that are restricted to and required in the nucleus and cytoplasm, respectively. Mol Cell Biol *17*, 6122–6130.

Kastenmayer, J.P., and Green, P.J. (2000). Novel features of the XRN-family in Arabidopsis: evidence that AtXRN4, one of several orthologs of nuclear Xrn2p/Rat1p, functions in the cytoplasm. Proc. Natl. Acad. Sci. U.S.A. *97*, 13985–13990.

Kawauchi, J., Mischo, H., Braglia, P., Rondon, A., and Proudfoot, N.J. (2008). Budding yeast RNA polymerases I and II employ parallel mechanisms of transcriptional termination. Genes Dev. *22*, 1082–1092.

Keller, W., Bienroth, S., Lang, K.M., and Christofori, G. (1991). Cleavage and polyadenylation factor CPF specifically interacts with the pre-mRNA 3' processing signal AAUAAA. EMBO J. *10*, 4241–4249.

Kenna, M., Stevens, A., McCammon, M., and Douglas, M.G. (1993). An essential yeast gene with homology to the exonuclease-encoding XRN1/KEM1 gene also encodes a protein with exoribonuclease activity. Molecular and Cellular Biology *13*, 341–350.

Kessler, M.M., Zhao, J., and Moore, C.L. (1996). Purification of the Saccharomyces cerevisiae cleavage/polyadenylation factor I. Separation into two components that are required for both cleavage and polyadenylation of mRNA 3' ends. J. Biol. Chem. *271*, 27167–27175.

Kim, S.H., and Lin, R.J. (1993). Pre-mRNA splicing within an assembled yeast spliceosome requires an RNA-dependent ATPase and ATP hydrolysis. Proc. Natl. Acad. Sci. U.S.A. *90*, 888–892.

Kim, S.H., and Lin, R.J. (1996). Spliceosome activation by PRP2 ATPase prior to the first transesterification reaction of pre-mRNA splicing. Mol Cell Biol *16*, 6810–6819.

Kim, M., Ahn, S.-H., Krogan, N.J., Greenblatt, J.F., and Buratowski, S. (2004a). Transitions in RNA polymerase II elongation complexes at the 3' ends of genes. EMBO J. *23*, 354–364.

Kim, M., Krogan, N.J., Vasiljeva, L., Rando, O.J., Nedea, E., Greenblatt, J.F., and Buratowski, S. (2004b). The yeast Rat1 exonuclease promotes transcription termination by RNA polymerase II. Nature *432*, 517–522.

Kim, M., Suh, H., Cho, E.-J., and Buratowski, S. (2009). Phosphorylation of the yeast Rpb1 C-terminal domain at serines 2, 5, and 7. J. Biol. Chem. *284*, 26421–26426.

Krogan, N.J., Kim, M., Tong, A., Golshani, A., Cagney, G., Canadien, V., Richards, D.P., Beattie, B.K., Emili, A., Boone, C., et al. (2003). Methylation of histone H3 by Set2 in Saccharomyces cerevisiae is linked to transcriptional elongation by RNA polymerase II. Mol. Cell. Biol. *23*, 4207–4218.

Krogan, N.J., Cagney, G., Yu, H., Zhong, G., Guo, X., Ignatchenko, A., Li, J., Pu, S., Datta, N., Tikuisis, A.P., et al. (2006). Global landscape of protein complexes in the yeast Saccharomyces cerevisiae. Nature *440*, 637–643.

Kubik, S., Bruzzone, M.J., and Shore, D. (2017). TFIID or not TFIID, a continuing transcriptional SAGA. EMBO J. *36*, 248–249.

Kulish, D., and Struhl, K. (2001). TFIIS enhances transcriptional elongation through an artificial arrest site in vivo. Mol. Cell. Biol. *21*, 4162–4168.

Lacadie, S.A., Tardiff, D.F., Kadener, S., and Rosbash, M. (2006). In vivo commitment to yeast cotranscriptional splicing is sensitive to transcription elongation mutants. Genes Dev. *20*, 2055–2066.

Lagrange, T., Kapanidis, A.N., Tang, H., Reinberg, D., and Ebright, R.H. (1998). New core promoter element in RNA polymerase II-dependent transcription: sequence-specific DNA binding by transcription factor IIB. Genes Dev. *12*, 34–44.

Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nat. Methods *9*, 357–359.

Laribee, R.N., Krogan, N.J., Xiao, T., Shibata, Y., Hughes, T.R., Greenblatt, J.F., and Strahl, B.D. (2005). BUR kinase selectively regulates H3 K4 trimethylation and H2B ubiquitylation through recruitment of the PAF elongation complex. Curr. Biol. *15*, 1487–1493.

Larochelle, M., Robert, M.-A., Hébert, J.-N., Liu, X., Matteau, D., Rodrigue, S., Tian, B., Jacques, P.-É., and Bachand, F. (2018). Common mechanism of transcription termination at coding and noncoding RNA genes in fission yeast. Nature Communications *9*, 1–15.

Leeds, N.B., Small, E.C., Hiley, S.L., Hughes, T.R., and Staley, J.P. (2006). The Splicing Factor Prp43p, a DEAH Box ATPase, Functions in Ribosome Biogenesis. Molecular and Cellular Biology *26*, 513–522.

Lescoute, A., and Westhof, E. (2006). Topology of three-way junctions in folded RNAs. RNA *12*, 83–93.

Li, X., Zhou, B., Chen, L., Gou, L.-T., Li, H., and Fu, X.-D. (2017). GRID-seq reveals the global RNA-chromatin interactome. Nat. Biotechnol. *35*, 940–950.

Li, X.-Y., Bhaumik, S.R., Zhu, X., Li, L., Shen, W.-C., Dixit, B.L., and Green, M.R. (2002). Selective recruitment of TAFs by yeast upstream activating sequences. Implications for eukaryotic promoter structure. Curr. Biol. *12*, 1240–1244.

Liu, X., Bushnell, D.A., Wang, D., Calero, G., and Kornberg, R.D. (2010). Structure of an RNA polymerase II-TFIIB complex and the transcription initiation mechanism. Science *327*, 206–209.

Liu, Y., Warfield, L., Zhang, C., Luo, J., Allen, J., Lang, W.H., Ranish, J., Shokat, K.M., and Hahn, S. (2009). Phosphorylation of the transcription elongation factor Spt5 by yeast Bur1 kinase stimulates recruitment of the PAF complex. Mol. Cell. Biol. *29*, 4852–4863.

Liu, Y.-C., Chen, H.-C., Wu, N.-Y., and Cheng, S.-C. (2007). A novel splicing factor, Yju2, is associated with NTC and acts after Prp2 in promoting the first catalytic reaction of pre-mRNA splicing. Mol. Cell. Biol. *27*, 5403–5413.

Logan, J., Falck-Pedersen, E., Darnell, J.E., and Shenk, T. (1987). A poly(A) addition site and a downstream termination region are required for efficient cessation of transcription by RNA polymerase II in the mouse beta maj-globin gene. Proc. Natl. Acad. Sci. U.S.A. *84*, 8306–8310.

Lund, M.K., and Guthrie, C. (2005). The DEAD-Box Protein Dbp5p Is Required to Dissociate Mex67p from Exported mRNPs at the Nuclear Rim. Molecular Cell *20*, 645–651.

Luo, W., Johnson, A.W., and Bentley, D.L. (2006). The role of Rat1 in coupling mRNA 3'-end processing to transcription termination: implications for a unified allosteric-torpedo model. Genes Dev. *20*, 954–965.

Lykke-Andersen, S., and Jensen, T.H. (2007). Overlapping pathways dictate termination of RNA polymerase II transcription. Biochimie *89*, 1177–1182.

Malagon, F., Kireeva, M.L., Shafer, B.K., Lubkowska, L., Kashlev, M., and Strathern, J.N. (2006). Mutations in the Saccharomyces cerevisiae RPB1 gene conferring hypersensitivity to 6-azauracil. Genetics *172*, 2201–2209.

Martinez-Rucobo, F.W., Sainsbury, S., Cheung, A.C., and Cramer, P. (2011). Architecture of the RNA polymerase–Spt4/5 complex and basis of universal transcription processivity. EMBO J *30*, 1302–1310.

Mayer, A., Heidemann, M., Lidschreiber, M., Schreieck, A., Sun, M., Hintermair, C., Kremmer, E., Eick, D., and Cramer, P. (2012). CTD tyrosine phosphorylation impairs termination factor recruitment to RNA polymerase II. Science *336*, 1723–1725.

Meisterernst, M., and Roeder, R.G. (1991). Family of proteins that interact with TFIID and regulate promoter activity. Cell *67*, 557–567.

Miki, T.S., Rüegger, S., Gaidatzis, D., Stadler, M.B., and Großhans, H. (2014a). Engineering of a conditional allele reveals multiple roles of XRN2 in Caenorhabditis elegans development and substrate specificity in microRNA turnover. Nucleic Acids Res. *42*, 4056–4067.

Miki, T.S., Richter, H., Rüegger, S., and Großhans, H. (2014b). PAXT-1 promotes XRN2 activity by stabilizing it through a conserved domain. Mol. Cell *53*, 351–360.

Miki, T.S., Carl, S.H., and Großhans, H. (2017). Two distinct transcription termination modes dictated by promoters. Genes Dev. *31*, 1870–1879.

Miller, T., Krogan, N.J., Dover, J., Erdjument-Bromage, H., Tempst, P., Johnston, M., Greenblatt, J.F., and Shilatifard, A. (2001). COMPASS: A complex of proteins associated with a trithorax-related SET domain protein. Proc Natl Acad Sci U S A *98*, 12902–12907.

Minvielle-Sebastia, L., Preker, P.J., Wiederkehr, T., Strahm, Y., and Keller, W. (1997). The major yeast poly(A)-binding protein is associated with cleavage factor IA and functions in premessenger RNA 3′-end formation. PNAS *94*, 7897–7902.

Mischo, H.E., and Proudfoot, N.J. (2013). Disengaging polymerase: terminating RNA polymerase II transcription in budding yeast. Biochim. Biophys. Acta *1829*, 174–185.

Mueller, C.L., and Jaehning, J.A. (2002). Ctr9, Rtf1, and Leo1 are components of the Paf1/RNA polymerase II complex. Mol. Cell. Biol. *22*, 1971–1980.

Mukundan, B., and Ansari, A. (2011). Novel Role for Mediator Complex Subunit Srb5/Med18 in Termination of Transcription. J. Biol. Chem. *286*, 37053–37057.

Murakami, K., Elmlund, H., Kalisman, N., Bushnell, D.A., Adams, C.M., Azubel, M., Elmlund, D., Levi-Kalisman, Y., Liu, X., Gibbons, B.J., et al. (2013). Architecture of an RNA polymerase II transcription pre-initiation complex. Science *342*, 1238724.

Murakami, K., Mattei, P.-J., Davis, R.E., Jin, H., Kaplan, C.D., and Kornberg, R.D. (2015). Uncoupling Promoter Opening from Start-Site Scanning. Mol. Cell *59*, 133–138.

Nagarajan, V.K., Jones, C.I., Newbury, S.F., and Green, P.J. (2013). XRN 5'→3' exoribonucleases: structure, mechanisms and functions. Biochim. Biophys. Acta *1829*, 590–603.

Neubauer, G., Gottschalk, A., Fabrizio, P., Séraphin, B., Lührmann, R., and Mann, M. (1997). Identification of the proteins of the yeast U1 small nuclear ribonucleoprotein complex by mass spectrometry. Proc. Natl. Acad. Sci. U.S.A. *94*, 385–390.

Neuvéglise, C., Marck, C., and Gaillardin, C. (2011). The intronome of budding yeasts. C. R. Biol. *334*, 662–670.

Ng, H.H., Dole, S., and Struhl, K. (2003). The Rtf1 Component of the Paf1 Transcriptional Elongation Complex Is Required for Ubiquitination of Histone H2B. J. Biol. Chem. *278*, 33625–33628.

Nojima, T., Gomes, T., Grosso, A.R.F., Kimura, H., Dye, M.J., Dhir, S., Carmo-Fonseca, M., and Proudfoot, N.J. (2015). Mammalian NET-Seq Reveals Genome-wide Nascent Transcription Coupled to RNA Processing. Cell *161*, 526–540.

Nojima, T., Rebelo, K., Gomes, T., Grosso, A.R., Proudfoot, N.J., and Carmo-Fonseca, M. (2018). RNA Polymerase II Phosphorylated on CTD Serine 5 Interacts with the Spliceosome during Co-transcriptional Splicing. Mol. Cell *72*, 369-379.e4.

Nordick, K., Hoffman, M.G., Betz, J.L., and Jaehning, J.A. (2008). Direct interactions between the Paf1 complex and a cleavage and polyadenylation factor are revealed by dissociation of Paf1 from RNA polymerase II. Eukaryotic Cell *7*, 1158–1167.

Northrup, S.H., and Erickson, H.P. (1992). Kinetics of protein-protein association explained by Brownian dynamics computer simulation. Proc Natl Acad Sci U S A *89*, 3338–3342.

Ohi, M.D., Vander Kooi, C.W., Rosenberg, J.A., Chazin, W.J., and Gould, K.L. (2003). Structural insights into the U-box, a domain associated with multi-ubiquitination. Nat Struct Biol *10*, 250–255.

Ooi, S.L., Samarsky, D.A., Fournier, M.J., and Boeke, J.D. (1998). Intronic snoRNA biosynthesis in Saccharomyces cerevisiae depends on the lariat-debranching enzyme: intron length effects and activity of a precursor snoRNA. RNA *4*, 1096–1110.

Osheim, Y.N., Miller, O.L., and Beyer, A.L. (1985). RNP particles at splice junction sequences on Drosophila chorion transcripts. Cell *43*, 143–151.

O'Sullivan, J.M., Tan-Wong, S.M., Morillon, A., Lee, B., Coles, J., Mellor, J., and Proudfoot, N.J. (2004). Gene loops juxtapose promoters and terminators in yeast. Nat. Genet. *36*, 1014–1018.

Padgett, R.A., Konarska, M.M., Grabowski, P.J., Hardy, S.F., and Sharp, P.A. (1984). Lariat RNA's as intermediates and products in the splicing of messenger RNA precursors. Science *225*, 898–903.

Park, J., Kang, M., and Kim, M. (2015). Unraveling the mechanistic features of RNA polymerase II termination by the 5'-3' exoribonuclease Rat1. Nucleic Acids Res. *43*, 2625–2637.

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

Parker, R., Siliciano, P.G., and Guthrie, C. (1987). Recognition of the TACTAAC box during mRNA splicing in yeast involves base pairing to the U2-like snRNA. Cell *49*, 229–239.

Pearson, E.L., and Moore, C.L. (2013). Dismantling Promoter-driven RNA Polymerase II Transcription Complexes in Vitro by the Termination Factor Rat1. J. Biol. Chem. *288*, 19750–19759.

Petfalski, E., Dandekar, T., Henry, Y., and Tollervey, D. (1998). Processing of the precursors to small nucleolar RNAs and rRNAs requires common components. Mol. Cell. Biol. *18*, 1181–1189.

Pillon, M.C., Sobhany, M., Borgnia, M.J., Williams, J.G., and Stanley, R.E. (2017). Grc3 programs the essential endoribonuclease Las1 for specific RNA cleavage. Proc Natl Acad Sci U S A *114*, E5530–E5538.

Porath, J. (1997). From gel filtration to adsorptive size exclusion. J. Protein Chem. *16*, 463–468.

Prabakaran, S., Lippens, G., Steen, H., and Gunawardena, J. (2012). Post-translational modification: nature's escape from genetic imprisonment and the basis for dynamic information encoding. Wiley Interdiscip Rev Syst Biol Med *4*, 565–583.

Proudfoot, N.J. (1986). Transcriptional interference and termination between duplicated alpha-globin gene constructs suggests a novel mechanism for gene regulation. Nature *322*, 562–565.

Proudfoot, N.J. (1989). How RNA polymerase II terminates transcription in higher eukaryotes. Trends Biochem. Sci. *14*, 105–110.

Qu, L.-H., Henras, A., Lu, Y.-J., Zhou, H., Zhou, W., Zhu, Y.-Q., Zhao, J., Henry, Y., Caizergues-Ferrer, M., and Bachellerie, J.-P. (1999). Seven Novel Methylation Guide Small Nucleolar RNAs Are Processed from a Common Polycistronic Transcript by Rat1p and RNase III in Yeast. Mol Cell Biol *19*, 1144–1158.

Richard, P., and Manley, J.L. (2009). Transcription termination by nuclear RNA polymerases. Genes Dev. *23*, 1247–1269.

Richter, H., Katic, I., Gut, H., and Großhans, H. (2016). Structural Basis and Function of XRN2-Binding by XTB Domains. Nat Struct Mol Biol *23*, 164–171.

Rigaut, G., Shevchenko, A., Rutz, B., Wilm, M., Mann, M., and Séraphin, B. (1999). A generic protein purification method for protein complex characterization and proteome exploration. Nat. Biotechnol. *17*, 1030–1032.

Rosonina, E., Kaneko, S., and Manley, J.L. (2006). Terminating the transcript: breaking up is hard to do. Genes Dev. *20*, 1050–1056.

Ruby, S.W., and Abelson, J. (1988). An early hierarchic role of U1 small nuclear ribonucleoprotein in spliceosome assembly. Science *242*, 1028–1035.

Ruskin, B., Krainer, A.R., Maniatis, T., and Green, M.R. (1984). Excision of an intact intron as a novel lariat structure during pre-mRNA splicing in vitro. Cell *38*, 317–331.

Sansó, M., Levin, R.S., Lipp, J.J., Wang, V.Y.-F., Greifenberg, A.K., Quezada, E.M., Ali, A., Ghosh, A., Larochelle, S., Rana, T.M., et al. (2016). P-TEFb regulation of transcription termination factor Xrn2 revealed by a chemical genetic screen for Cdk9 substrates. Genes Dev *30*, 117–131.

Schroeder, S.C., Schwer, B., Shuman, S., and Bentley, D. (2000). Dynamic association of capping enzymes with transcribing RNA polymerase II. Genes Dev. *14*, 2435–2440.

Schwer, B., and Gross, C.H. (1998). Prp22, a DExH-box RNA helicase, plays two distinct roles in yeast pre-mRNA splicing. EMBO J. *17*, 2086–2094.

Schwer, B., and Guthrie, C. (1991). PRP16 is an RNA-dependent ATPase that interacts transiently with the spliceosome. Nature *349*, 494–499.

Schwer, B., and Shuman, S. (1996). Conditional inactivation of mRNA capping enzyme affects yeast pre-mRNA splicing in vivo. RNA *2*, 574–583.

Seila, A.C., Calabrese, J.M., Levine, S.S., Yeo, G.W., Rahl, P.B., Flynn, R.A., Young, R.A., and Sharp, P.A. (2008). Divergent transcription from active promoters. Science *322*, 1849–1851.

Shobuike, T., Sugano, S., Yamashita, T., and Ikeda, H. (1995). Characterization of cDNA encoding mouse homolog of fission yeast dhp1+ gene: structural and functional conservation. Nucleic Acids Res. *23*, 357–361.

Spingola, M., Grate, L., Haussler, D., and Ares, M. (1999). Genome-wide bioinformatic and molecular analysis of introns in Saccharomyces cerevisiae. RNA *5*, 221–234.

Staley, J.P., and Guthrie, C. (1998). Mechanical Devices of the Spliceosome: Motors, Clocks, Springs, and Things. Cell *92*, 315–326.

Staley, J.P., and Woolford, J.L. (2009). Assembly of ribosomes and spliceosomes: complex ribonucleoprotein machines. Curr. Opin. Cell Biol. *21*, 109–118.

Stevens, A., and Poole, T.L. (1995). 5'-exonuclease-2 of Saccharomyces cerevisiae. Purification and features of ribonuclease activity with comparison to 5'-exonuclease-1. J. Biol. Chem. *270*, 16063–16069.

Strässer, K., Bassler, J., and Hurt, E. (2000). Binding of the Mex67p/Mtr2p heterodimer to FXFG, GLFG, and FG repeat nucleoporins is essential for nuclear mRNA export. J. Cell Biol. *150*, 695–706.

Sträßer, K., Masuda, S., Mason, P., Pfannstiel, J., Oppizzi, M., Rodriguez-Navarro, S., Rondón, A.G., Aguilera, A., Struhl, K., Reed, R., et al. (2002). TREX is a conserved complex coupling transcription with messenger RNA export. Nature *417*, 304–308.

Sugano, S., Shobuike, T., Takeda, T., Sugino, A., and Ikeda, H. (1994). Molecular analysis of the dhp1+ gene of Schizosaccharomyces pombe: an essential gene that has homology to the DST2 and RAT1 genes of Saccharomyces cerevisiae. Mol. Gen. Genet. *243*, 1–8.

Sun, Z.W., and Hampsey, M. (1996). Synthetic enhancement of a TFIIB defect by a mutation in SSU72, an essential yeast gene encoding a novel protein that affects transcription start site selection in vivo. Mol Cell Biol *16*, 1557–1566.

Swanson, M.S., Malone, E.A., and Winston, F. (1991). SPT5, an essential gene important for normal transcription in Saccharomyces cerevisiae, encodes an acidic nuclear protein with a carboxy-terminal repeat. Mol. Cell. Biol. *11*, 3009–3019.

Taatjes, D.J., Näär, A.M., Andel, F., Nogales, E., and Tjian, R. (2002). Structure, function, and activator-induced conformations of the CRSP coactivator. Science *295*, 1058–1062.

Takahashi, H., Parmely, T.J., Sato, S., Tomomori-Sato, C., Banks, C.A.S., Kong, S.E., Szutorisz, H., Swanson, S.K., Martin-Brown, S., Washburn, M.P., et al. (2011). Human mediator subunit MED26 functions as a docking site for transcription elongation factors. Cell *146*, 92–104.

Tang, J., Abovich, N., Fleming, M.L., Seraphin, B., and Rosbash, M. (1997). Identification and characterization of a yeast homolog of U1 snRNP-specific protein C. EMBO J. *16*, 4082–4091.

Tosi, A., Haas, C., Herzog, F., Gilmozzi, A., Berninghausen, O., Ungewickell, C., Gerhold, C.B., Lakomek, K., Aebersold, R., Beckmann, R., et al. (2013). Structure and subunit topology of the INO80 chromatin remodeler and its nucleosome complex. Cell *154*, 1207–1219.

Tran, E.J., Zhou, Y., Corbett, A.H., and Wente, S.R. (2007). The DEAD-box protein Dbp5 controls mRNA export by triggering specific RNA:protein remodeling events. Mol. Cell *28*, 850–859.

Tucker, J.F., Ohle, C., Schermann, G., Bendrin, K., Zhang, W., Fischer, T., and Zhang, K. (2016). A Novel Epigenetic Silencing Pathway Involving the Highly Conserved 5'-3' Exoribonuclease Dhp1/Rat1/Xrn2 in Schizosaccharomyces pombe. PLoS Genet. *12*, e1005873.

Van Oss, S.B., Cucinotta, C.E., and Arndt, K.M. (2017). Emerging Insights into the Roles of the Paf1 Complex in Gene Regulation. Trends Biochem Sci *42*, 788–798.

Villa, T., and Guthrie, C. (2005). The Isy1p component of the NineTeen Complex interacts with the ATPase Prp16p to regulate the fidelity of pre-mRNA splicing. Genes Dev. *19*, 1894–1904.

Vincenti, S., Chiara, V.D., Bozzoni, I., and Presutti, C. (2007). The position of yeast snoRNA-coding regions within host introns is essential for their biosynthesis and for efficient splicing of the host pre-mRNA. RNA *13*, 138–150.

Wagner, J.D., Jankowsky, E., Company, M., Pyle, A.M., and Abelson, J.N. (1998). The DEAH-box protein PRP22 is an ATPase that mediates ATP-dependent mRNA release from the spliceosome and unwinds RNA duplexes. EMBO J *17*, 2926–2937.

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

Wang, M., and Pestov, D.G. (2011). 5′-end surveillance by Xrn2 acts as a shared mechanism for mammalian pre-rRNA maturation and decay. Nucleic Acids Res *39*, 1811–1822.

Warfield, L., Ramachandran, S., Baptista, T., Devys, D., Tora, L., and Hahn, S. (2017). Transcription of Nearly All Yeast RNA Polymerase II-Transcribed Genes Is Dependent on Transcription Factor TFIID. Mol. Cell *68*, 118-129.e5.

Warner, J.R. (1999). The economics of ribosome biosynthesis in yeast. Trends in Biochemical Sciences *24*, 437–440.

Weinmann, R., and Roeder, R.G. (1974). Role of DNA-dependent RNA polymerase 3 in the transcription of the tRNA and 5S RNA genes. Proc. Natl. Acad. Sci. U.S.A. *71*, 1790–1794.

Weinmann, R., Raskas, H.J., and Roeder, R.G. (1974). Role of DNA-dependent RNA polymerases II and III in transcription of the adenovirus genome late in productive infection. Proc. Natl. Acad. Sci. U.S.A. *71*, 3426–3439.

Weinstein, L.B., and Steitz, J.A. (1999). Guided tours: from precursor snoRNA to functional snoRNP. Curr. Opin. Cell Biol. *11*, 378–384.

West, S., Gromak, N., and Proudfoot, N.J. (2004). Human 5’ --> 3’ exonuclease Xrn2 promotes transcription termination at co-transcriptional cleavage sites. Nature *432*, 522–525.

Westermarck, J., Ivaska, J., and Corthals, G.L. (2013). Identification of Protein Interactions Involved in Cellular Signaling. Mol Cell Proteomics *12*, 1752–1763.

Will, C.L., and Lührmann, R. (2011). Spliceosome structure and function. Cold Spring Harb Perspect Biol *3*.

Wilson, C.J., Chao, D.M., Imbalzano, A.N., Schnitzler, G.R., Kingston, R.E., and Young, R.A. (1996). RNA polymerase II holoenzyme contains SWI/SNF regulators involved in chromatin remodeling. Cell *84*, 235–244.

Wittschieben, B.O., Fellows, J., Du, W., Stillman, D.J., and Svejstrup, J.Q. (2000). Overlapping roles for the histone acetyltransferase activities of SAGA and elongator in vivo. EMBO J. *19*, 3060–3068.

Wood, A., and Shilatifard, A. (2006). Bur1/Bur2 and the Ctk complex in yeast: the split personality of mammalian P-TEFb. Cell Cycle *5*, 1066–1068.

Wood, A., Schneider, J., Dover, J., Johnston, M., and Shilatifard, A. (2003). The Paf1 complex is essential for histone monoubiquitination by the Rad6-Bre1 complex, which signals for histone methylation by COMPASS and Dot1p. J. Biol. Chem. *278*, 34739–34742.

Woolford, J.L., and Peebles, C.L. (1992). RNA splicing in lower eukaryotes. Curr. Opin. Genet. Dev. *2*, 712–719.

Xiang, S., Cooper-Morgan, A., Jiao, X., Kiledjian, M., Manley, J.L., and Tong, L. (2009). Structure and function of the 5’-->3’ exoribonuclease Rat1 and its activating partner Rai1. Nature *458*, 784–788.

Xue, Y., Bai, X., Lee, I., Kallstrom, G., Ho, J., Brown, J., Stevens, A., and Johnson, A.W. (2000). Saccharomyces cerevisiae RAI1 (YGL246c) is homologous to human DOM3Z and encodes a protein that binds the nuclear exoribonuclease Rat1p. Mol. Cell. Biol. *20*, 4006–4015.

Zakrzewska-Placzek, M., Souret, F.F., Sobczyk, G.J., Green, P.J., and Kufel, J. (2010). Arabidopsis thaliana XRN2 is required for primary cleavage in the pre-ribosomal RNA. Nucleic Acids Res. *38*, 4487–4502.

Zawel, L., and Reinberg, D. (1992). Advances in RNA polymerase II transcription. Curr. Opin. Cell Biol. *4*, 488–495.

Zhang, D., and Rosbash, M. (1999). Identification of eight proteins that cross-link to pre-mRNA in the yeast commitment complex. Genes Dev. *13*, 581–592.

Zhang, M., and Green, M.R. (2001). Identification and characterization of yUAP/Sub2p, a yeast homolog of the essential human pre-mRNA splicing factor hUAP56. Genes Dev. *15*, 30–35.

Zhang, G., Campbell, E.A., Minakhin, L., Richter, C., Severinov, K., and Darst, S.A. (1999a). Crystal structure of Thermus aquaticus core RNA polymerase at 3.3 A resolution. Cell *98*, 811–824.

Zhang, H., Rigo, F., and Martinson, H.G. (2015). Poly(A) Signal-Dependent Transcription Termination Occurs through a Conformational Change Mechanism that Does Not Require Cleavage at the Poly(A) Site. Mol. Cell *59*, 437–448.

Zhang, M., Yu, L., Xin, Y., Hu, P., Fu, Q., Yu, C., and Zhao, S. (1999b). Cloning and Mapping of the XRN2 Gene to Human Chromosome 20p11.1–p11.2. Genomics *59*, 252–254.

Zhang, Z., Fu, J., and Gilmour, D.S. (2005). CTD-dependent dismantling of the RNA polymerase II elongation complex by the pre-mRNA 3′-end processing factor, Pcf11. Genes Dev *19*, 1572–1580.

Zhao, J., Kessler, M., Helmling, S., O'Connor, J.P., and Moore, C. (1999). Pta1, a component of yeast CF II, is required for both cleavage and poly(A) addition of mRNA precursor. Mol. Cell. Biol. *19*, 7733–7740.

Zhou, K., Kuo, W.H.W., Fillingham, J., and Greenblatt, J.F. (2009). Control of transcriptional elongation and cotranscriptional histone modification by the yeast BUR kinase substrate Spt5. Proc. Natl. Acad. Sci. U.S.A. *106*, 6956–6961.

ABSTRACT

TERMINATION INDEPENDENT ROLE OF RAT1 IN COTRANSCRIPTIONAL SPLICING IN BUDDING YEAST

by

ZUZER HAKIMUDDIN DHOONDIA

MAY 2020

Advisor: Dr. Athar Ansari

Major: Biological Sciences

Degree: Doctor of Philosophy

Rat1 is a 5′→3′ exoribonuclease in budding yeast belonging to the XRN-family of nucleases. It is a highly conserved protein with homologs being present in fission yeast, flies, worms, mice and humans. Rat1 and its homolog in metazoan have been shown to function in multiple facets of RNA metabolism. In this study, we report a novel role of Rat1 in splicing of pre-mRNA in budding yeast. In the absence of the functional Rat1 in the nucleus, an increase in the level of unspliced transcripts was observed in yeast cells. Strand-specific TRO analysis revealed that the accumulation of unspliced transcripts upon nuclear depletion of Rat1 was not due to stabilization of intron-containing transcripts. The unspliced transcripts were capped at their 5' end, suggesting that Rat1 nuclear RNA surveillance function is not responsible for the phenotype. Furthermore, altered elongation rate of RNAPII in Rat1 mutant was not linked to the accumulation of unspliced transcripts. Inhibition of termination by inactivation of Rna14, a component of CF1A complex as well as deletion of Rat1 termination complex components, Rai1 and Rtt103, did not affect the level of unspliced transcripts. These results strongly suggest a

novel role of Rat1 in splicing, which is independent of its termination function. The interaction of Rat1 with the splicing competent introns, and its association with spliceosomal components; Clf1, Isy1, Yju2, Sub2 and Prp43 suggest that accumulation of unspliced transcripts in the Rat1 mutant is due to a direct role of Rat1 in splicing and not due to lack of degradation of unspliced transcripts in rat1 mutant. Lastly, the recruitment of Prp2 splicing factor on the intron was compromised in the Rat1 mutant suggesting that Rat1 is a critical factor in the dynamics of spliceosome assembly in budding yeast.

AUTOBIOGRAPHICAL STATEMENT

ZUZER HAKIMUDDIN DHOONDIA

EDUCATION

2013-2020 Ph.D. Biological Sciences, Wayne State University,

2008-2010 M.Sc. Microbiology, University of Mumbai

2004-2006 B.Sc. Microbiology University of Mumbai

FELLOWSHIP AND AWARDS

2018 Recipient of **'Ellis Award'** for graduate research support and travel, Stanley Ellis, Ph.D. Endowed Memorial

2017-2018 Elected for **'Thomas C. Rumble University Graduate Fellowship'**, Wayne State University

2017 Awarded **'Certificate of Recognition'** for exceptional work in research, Wayne State University

2017 Best poster award, 'An investigation into the potentially novel role of Rat1p in transcription', at Biological Sciences Department Retreat, Wayne State University

2017 Recipient of **'Wilhelmine Haley Award'** for the research support by the Department of Biological Sciences, Wayne State University

2007 Awarded Science Honors in Microbiology, K.C College, Mumbai, India

PUBLICATIONS

Dhoondia. Z, Elewa. H, Arif. Z, Malik. M, Pique-Regi, R, Ansari. A. A novel role of Rat1 in co-transcriptional splicing in budding yeast. (Manuscript under submission).

Allepuz-Fuster. P; O'Brien. M, González-Polo. N, Pereira. B, **Dhoondia. Z**, Ansari, A, Calvo. O. RNA polymerase II plays an active role in the formation of gene loops through the Rpb4 subunit. Nucleic Acid Research, 2019, 47, (17), 8975–8987.

Dhoondia. Z, Tarockoff. R, Al Husini. N, Medler. S, Agarwal. N, Ansari. A. Analysis of Termination of Transcription Using BrUTP-strand-specific Transcription Run-on (TRO) Approach. J. Vis. Exp, 2017, (121), e55446, doi:10.3791/55446.

Cloutier. SC, Wang. S, Ma. WK, Al Husini. N, **Dhoondia. Z**, Ansari. A, Pascuzzi. PE, Tran. EJ. Regulated formation of lncRNA-DNA hybrids enables faster transcriptional induction and environmental adaptation. Molecular Cell, 2016, 61 (3), 393-404.

Dhoondia. Z, Chakraborty. H. *Lactobacillus* mediated synthesis of silver oxide nanoparticles. Nanomaterials and Nanotechnology, 2012, Vol 2. ISBN: 1847-9804, InTech, DOI: 10.5772/55741.