Dissertation zur Erlangung des Doktorgrades

der Fakultät für Chemie und Pharmazie

der Ludwig-Maximilians-Universität München

The functions of Mud2 and Prp19C components in transcription and TREX occupancy

Rashmi Minocha

aus

Hansi, India

2018

Erklärung

Diese Dissertation wurde im Sinne von § 7 der Promotionsordnung vom 28. November 2011 von Frau **Prof. Dr. Katja Sträßer** betreut.

Eidesstattliche Versicherung

Diese Dissertation wurde eigenständig und ohne unerlaubte Hilfe erarbeitet.

München, 01-10-2018

RASHMI MINOCHA

Dissertation eingereicht am: 02-07-2018

1. Gutachterin / 1. Gutachter: **Prof. Dr. Katja Sträßer**

2. Gutachterin / 2. Gutachter: **PD Dr. Dietmar E. Martin**

Mündliche Prüfung am: 18-09-2018

Acknowledgements

First of all, I would like to express my sincere gratitude to my supervisor **Prof. Dr. Katja Sträßer** for giving me this wonderful and exciting opportunity to work in her lab. I am grateful to her for constant support and guidance, for all the scientific discussions and exchange, for mentoring and advising me during difficult times, for providing me the opportunity to attend RNA and EMBO conferences, and for allowing me to pursuit my own ideas. I especially enjoyed and learnt a lot from her during the time of writing our manuscript.

I am grateful to **Dr. Dietmar E. Martin** for being so humble to be the second examiner of my thesis and for all the scientific discussions and interactions during my time at the Gene Center. I would also like to thank all other members of my thesis evaluation committee for their time.

I would like to thank **CIPSMwomen** for funding a part of my research and for the meetings they organized.

I would like to thank **Dr. Cornelia Burkert-Kautzsch**, a former PhD student of our lab and my 'first' German friend, who helped me a lot when I came for the first time to a foreign land to start my PhD. Thanks for taking care of all the official formalities after my arrival and helping me to settle down in a new place, and also for teaching me my first experiments in the lab. I can't thank enough for your generous, friendly and overwhelming nature, and for always being there for me.

I would like to thank **Dr. Anja Drescher** for helping me to find accommodation when our lab moved from Munich to Giessen, for giving me the opportunity to be a part of IRTG, for providing me a nice company during the late working hours in the lab, for introducing me to the German prehistoric dance styles and your dance group, for showing me your wonderful paintings and for always being there for 'little refreshing talks' in between the experiments as well as 'movie hours' at your place.

I would like to thank **Dr. Wolfgang Wende** for helping me in the structural predictions of Mud2, for all the scientific discussions, for looking into my computer problems, and for always being there for any kind of help. Your benevolent, generous, composed and pleasant personality is so inspiring and I always have a lot to learn from you.

I would like to thank all the **members of the Sträßer group**, past and present, for providing a very friendly and helpful working atmosphere in the lab and for also giving me this wonderful opportunity to learn more about Germany and German culture from you. A very special thanks to Dr. Max Reuter, Dr. Dominik Meinel, Dr. Sittinan Chanarat, Dr. Cornelia Kilchert, Birte

Keil, Christoph Wierschem, Maya Monakhova, Philipp Keil, Anja Riesterer, Lisa Leib, Kristin Grüllich, Olha Storozhuk and Laura Henke. Thanks to Max for being a great colleague and a good friend over the past years. Thanks to Dominik and Sittinan for always being so kind to reply to my experimental questions even after leaving the lab. Thanks to Cornelia for all the scientific discussions, especially while working on the Mud2 paper. Thanks to Birte for being a great friend, for comforting me during my low moments and for always motivating me. Thanks to Chris for never saying 'no' to any help I asked you ever, and for always being a nice and funny bench neighbor. Thanks to Maya for giving me a chance to learn about Russian culture and for all the friendly meetings. Thanks to Philipp for adding to the cheerful and humorous environment of the lab and for all the experimental or non-lab related help. Thanks to Anja for all the joyful and exciting conversations we had about various topics and for organizing the memorable lab retreat. Thanks to Lisa for cross-linking my cells and for always being there to help, even over the weekends and especially during my stressful phase while working on the revision of the paper. Thanks to Kristin for being a great bench neighbor, for all the lab or non-lab related conversations and for motivating me during my sad times. Thanks to Olha for your charming, optimistic and helpful personality and for giving me company during night hours in the lab. And thanks to Laura for taking over the project and for being a great bench neighbor during my last working days in this lab. I also want to thank all the bachelor and master students I supervised for giving me a chance to improve my teaching abilities and for contributing to this project.

I am grateful to my **family**, as without their constant support, I would not have reached this far! I want to especially thank to my parents for always believing in me, for your endless unconditional love and care, for always encouraging me to give my best efforts and for showing me the clear vision when the paths were blurry. I want to thank you, my little sister Nidhi, for being my best friend and my best companion, for your charming nature and high-spirited attitude, for your boundless love and understanding, and for always being there for me to share and inspire, despite of the time differences between India and Germany. Thank you so much! I also want to thank my little brother Mohit, who also helped me a lot to make this journey possible. Thank you for all the jubilant conversations we had, for coming to visit me, for our exciting trip together and for always trusting in me.

Last but not the least, I want to express my sincere thanks and gratitude to all my **friends**, in Germany and in India, for your never-ending support, for cheering me up during my 'experiment not working' times, for inspiring me through so many different ways and for always being there for me. Thank you!

Summary

Different steps in gene expression from transcription to mRNA processing and nuclear mRNP export are connected intimately for quality control. The TREX complex is a key player that couples transcription elongation to nuclear mRNA export (Strasser, Masuda et al. 2002). The Prp19 complex (Prp19C, also known as NTC complex), originally known for its role in splicing, plays a role in stabilizing TREX complex at transcribed genes. Prp19C interacts with the TREX *in vivo* and is also required for efficient transcription (Chanarat, Seizl et al. 2011). Furthermore, it has been reported in the mammalian cells that Prp19C is recruited to transcriptional machinery by the early splicing factor U2AF65 (David, Boyne et al. 2011). U2AF65 interacts directly with Prp19C and the phosphorylated CTD, leading to an increased recruitment of both U2AF65 and Prp19C in a CTD-dependent manner, which enhances the splicing reaction.

In *Saccharomyces cerevisiae*, Mud2 is the potential homologue of U2AF65. In this study, a novel function of Mud2 in the gene expression, namely in transcription, was identified. Using chromatin immunoprecipitation (ChIP), it was shown that Mud2 is recruited to both intron-containing and intronless genes in a transcription-dependent manner. Mud2 binds directly to the S2-phosphorylated CTD and S2 phosphorylation is needed for its full occupancy at transcribed genes. Furthermore, Mud2 coimmunoprecipitates Prp19C *in vivo* in an RNA-independent manner and is also required for maintaining Prp19C and TREX occupancy at transcribed genes. Using *in vivo* and *in vitro* transcription assays, it was shown that Mud2 is necessary for full transcriptional activity by RNA Polymerase II (RNAPII). Taken together, Mud2 can be classified as a novel transcription factor that is necessary for the recruitment of mRNA-binding proteins to transcribed genes.

In addition, a direct interaction between the Prp19C and THO, a sub-complex of TREX, was shown. Prp19C consists of the essential subunits Prp19, Cef1, Clf1, Syf1, Cwc2 and Prp46 and the non-essential subunits Ntc20, Snt309, Isy1 and Syf2. However, the molecular mechanism of Prp19C function and the roles of its individual subunits in maintaining TREX at transcribed genes and in transcription have remained unknown. In this study, the three non-essential subunits of Prp19C, Isy1, Ntc20 and Syf2, were investigated for their roles in Prp19C-TREX interaction and transcription. Additionally, an NTC related protein (NTR), which is a component of a complex containing Cef1 and hence named Cwc15 (complexed with Cef1), was also analyzed. The functions of these four proteins in Prp19C and TREX occupancy as well as Prp19C-TREX, Prp19C-Mud2 and Prp19C-RNAPII interaction were determined. Taken together, this analysis shows that specific Prp19C subunits play differential roles in gene expression.

Publication

Substantial parts of this thesis have been published in the following journal:

Rashmi Minocha[1], Varvara Popova[2], Daria Kopytova[2], Danny Misiak[3], Stefan Hüttelmaier[3], Sofia Georgieva[2] and Katja Sträßer[1] (2018). ‘‘Mud2 functions in transcription by recruiting the Prp19 and TREX complexes to transcribed genes’’, Nucleic Acids Research, gky782. https://doi.org/10.1093/nar/gky782

Contents

1. Introduction

1.1 Gene Expression

In all living cells, genetic information is stored in the form of DNA. Transfer of this information from DNA to RNA and from RNA to proteins is called gene expression and is one of the most fundamental processes occurring in every living cell. The expression of protein coding genes in eukaryotic cells is a highly coordinated and complex process involving a multitude of interconnected and highly coupled steps. All eukaryotic cells are divided into nuclear and cytoplasmic compartments separated from each other by a nuclear envelope. The synthesis of nascent pre-mRNAs takes place in the nucleus by the process of transcription. The nascent pre-mRNAs are further processed by capping, splicing, 3'-end processing and polyadenylation, and finally packaged by some RNA-binding proteins and other factors into the export-competent messenger ribonucleoprotein particles (mRNPs). The fully mature mRNPs are exported from the nucleus to cytoplasm for translation into proteins.

1.1.1 Transcription

The synthesis of RNA from DNA is called transcription, a process which marks the beginning of gene expression. During expression of a protein coding gene, transcription is facilitated by a multisubunit enzymatic complex called RNA polymerase II (RNAPII) along with a plethora of other proteins and transcription factors. The process of transcription is sub-divided into initiation, elongation, termination and RNAPII recycling, for reviews, see (Svejstrup 2004, Shandilya and Roberts 2012, Harlen and Churchman 2017).

1.1.1.1 Transcription initiation

The first step required to initiate transcription is formation of a pre-initiation complex (PIC). To begin its formation, specific DNA elements (also called core promoter elements: CPEs) located upstream of transcription start site (TSS) need to be recognized. One such widely

studied element is TATA box, which is usually located few bases upstream of TSS. TATA box is bound by the TATA binding protein (TBP) along with TBP-associated factors, which together form TFIID complex. This acts as a platform for binding of other general transcription factors (GTFs). The binding of TFIIA further stabilizes the interaction of TBP with DNA. TFIIB binds afterwards to form a ternary complex, which promotes the binding of hypophosphorylated form of RNAPII along with TFIIF, stabilizing the formation of PIC. The general transcription factors TFIIH and TFIIE are then recruited to the gene promoter along with another complex called mediator complex. TFIIH is a multisubunit complex with helicase and kinase activity. The helicase activity is required to unwind the DNA while the kinase activity is important to phosphorylate the heptapeptide residues in C-terminal domain (CTD) of Rpb1, the largest subunit of RNAPII (for details on Rpb1 CTD, please refer to Section 1.1.4). All these series of events lead to the ATP-dependent transition from closed form of PIC to open form of PIC, so that RNAPII can bind to the promoter and finally begin transcription (Cooper 2000). After synthesis of few bases, RNAPII loses contact with the promoter (called promoter clearance) leaving behind several GTFs as 'scaffold complex' and traverses ahead to enter into the mode of elongation (Hahn 2004). The RNAPII bound factors at this point form RNAPII transcript elongation complex (TEC) and RNAPII becomes hyperphosphorylated (Pokholok, Hannett et al. 2002).

1.1.1.2 Transcription elongation

Immediately after the start of transcription, a variety of transcription elongation factors are recruited to transcriptional machinery, which regulate this highly complex process (Sims, Belotserkovskaya et al. 2004, Zhou, Li et al. 2012, Jonkers and Lis 2015). As DNA is packaged into condensed chromatin form in the cell, RNAPII needs to overcome this barrier to access the DNA in nucleosomes. Various chromatin remodeling factors like Paf1 complex and Elf1 are recruited during the elongation phase to facilitate this process (Sims, Belotserkovskaya et al. 2004). Apart from this, a variety of other factors also regulate transcription elongation. For example, the negative elongation factor (NELF) and the DRB-sensitivity inducing factor (DSIF) are both negative transcription factors, which inhibit elongation by pausing RNAPII. P-TEFb, on the other hand, is a positive transcription elongation factor, which eliminates negative effects of NELF and DSIF, so that RNAPII can engage into productive elongation. This is achieved by phosphorylating the Spt5 subunit of DSIF and also the subunits of NELF complex by the Cdk9 kinase activity of P-TEFb (Peterlin and Price 2006). Apart from this, the

transcription factor TFIIF, besides its function in initiation, also plays an important role in stimulating elongation. The elongin complex (SIII) and the protein encoded by ELL (eleven-nineteen lysine-rich leukemia) gene further help to stimulate mRNA synthesis (Sims, Belotserkovskaya et al. 2004). The histone chaperone complex called FACT complex (facilitates chromatin transcription) and elongation factors TFIIS and Spt5 regulate processivity of RNAPII (Shandilya and Roberts 2012). Also, mRNA processing events such as 5'capping and splicing start cotranscriptionally during the elongation phase (Lee and Tarn 2013).

1.1.1.3 Transcription termination and RNAPII recycling

Transcription termination marks the release of nascent pre-mRNA and dissociation of RNAPII from DNA thereby ceasing transcription. It is signaled by the presence of specific poly(A) signals downstream of the gene. When RNAPII encounters a poly(A) signal, its processivity decreases causing a break in its activity. Two models are accepted to explain the transcription termination (Loya and Reines 2016). The first model is the 'allosteric model' according to which, presence of a poly(A) signal along with the recruitment of some other factors such as cleavage and polyadenylation factor (CPF) causes conformational changes in the transcription elongation complex leading to RNAPII pausing and termination. The other model is called 'torpedo model', which is based on the degradation of remaining 3' RNA by 5′-3′ exonuclease Rat1. The entry site for Rat1 is created by cleavage at the poly(A) signal. This degradation causes the release of RNAPII and various other factors from the termination site (Richard and Manley 2009). After the release, pre-mRNA is further processed before exporting to the cytoplasm while RNAPII undergoes several recycling events such as reverting back to the hypophosphorylated state (Shandilya and Roberts 2012).

1.1.2 mRNA processing

1.1.2.1 Capping

The first mRNA processing event is capping, which takes place as soon as newly generated mRNA is emerged from the transcription site. In this process, m^7Gppp residue called as 'cap' is added to the terminal or first nucleotide of nascent mRNA. This is achieved by mRNA capping enzyme in a series of events – first, terminal 5′-phosphate group is removed by the

action of an RNA 5'-triphosphatase Cet1, a subunit of mRNA capping enzyme. After this, a GTP is added by guanylyltransferase Ceg1, another subunit of mRNA capping enzyme. Abd1, a methyltransferase then transfers a methyl group to this Gppp (Topisirovic, Svitkin et al. 2011). Finally, the formed cap structure is bound by nuclear cap binding complex (CBC), which is composed of the subunits Cbp20 and Cbp80. mRNA capping is not only required for protecting the nascent transcript from degradation, but it also plays an important role in further events of mRNA processing and export (Meinel and Sträßer 2015).

1.1.2.2 Splicing

The next mRNA processing event is splicing i.e. removal of non-coding regions (called introns) from mRNA. It is a two-step based transesterification reaction catalyzed by a large ribonucleoprotein complex called spliceosome. The spliceosome consists of five nuclear ribonucleoproteins (snRNPs), containing U1, U2, U4, U5 and U6 snRNAs associated with various proteins. Splicing is carried out in a stepwise manner and requires sequential assembly and interaction of various snRNP and non-snRNP proteins, which leads to the activation of spliceosome. The activated spliceosome is then ready to catalyze the splicing reaction. There are several evidences which state that splicing is highly coupled to transcription and other mRNA processing events (Reed 2003).

The spliceosomal assembly takes place in a stepwise and ordered manner (Figure 1). First, the 5’ splice site is recognized by U1 snRNP while the 3’ splice site is recognized by two non-snRNP proteins: Msl5, which binds to the branchpoint sequence in the intron and Mud2 (Section 1.3), which binds to the polypyrimidine tract. Mud2 also interacts with U1 snRNP and Msl5 in a branchpoint dependent manner. This interaction recruits U2 snRNP to the branch point, leading to the displacement of Msl5 and Mud2 and thereby forming complex A at the splice site. The formation of complex A weakens the binding of U2 snRNP with splicing apparatus as well as recruits the preformed U5/U4/U6 tri-snRNP complex to splice site, forming complex B. This further leads to the release of U1 and U4 snRNP and formation of complex C, an activated form of the spliceosome. The activated spliceosome then performs the transesterification reaction, leading to the removal of intron and ligation of two exonic regions (Chanarat and Strasser 2013). The resulting post-splicing complex is then disassembled to release the spliced mRNA followed by the recycling of snRNPs (de Almeida and O'Keefe 2015). A non-snRNP splicing complex called Prp19 complex or Prp19C (Section 1.4) is also

recruited to splicing machinery during or after the dissociation of U4 snRNP and helps in the formation of activated spliceosome.

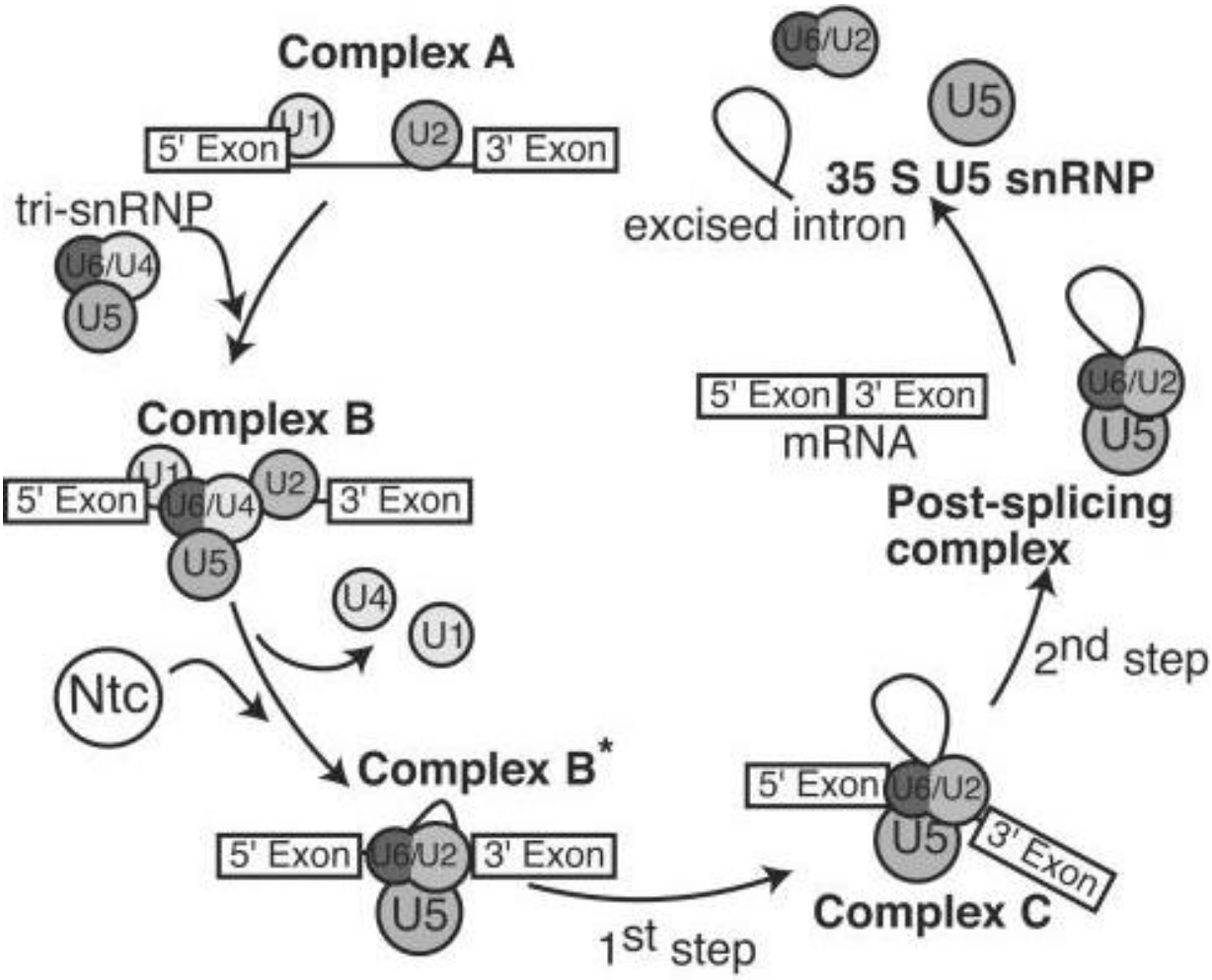


Figure 1. A schematic showing pre-mRNA splicing in yeast (Ohi, Vander Kooi et al. 2005).
The beginning of splicing is marked by the recognition of 5' and 3' of the splice site by U1 and U2 snRNAs, forming the Complex A. It is then followed by the association of U5/U4/U6 tri-snRNP forming the Complex B, which after a series of rearrangements leads to the formation of an active spliceosome intermediate (Complex C). After the splicing reaction, the post-splicing complex and the spliced mRNA are released and factors involved in the process are recycled.

1.1.2.3 3' end processing: cleavage and polyadenylation

All eukaryotic cells undergo processing of mRNA at 3' end, to enable the formation of mature and functional export competent mRNAs, which can be translated into proteins in the cytoplasm. 3′ end processing of a newly transcribed pre-mRNA takes place in two highly

coupled steps: cleavage of pre-mRNA and synthesis of a poly(A) tail, and requires presence of specific cleavage and polyadenylation signals (pA signals) upstream and downstream of the region to be cleaved. In mammalian cells, pA signals consist of A(A/U)UAAA hexameric sequence upstream of the cleavage site and a U and/or G/U rich sequence downstream of the cleavage site. The 3' end machinery in mammalian cells consists of the cleavage and polyadenylation specificity factor (CPSF), cleavage factors I and II, and cleavage stimulation factor (CstF) with some additional proteins. In *S. cerevisiae*, instead of A(A/U)UAAA element, there is a degenerate A-rich sequence, a UA-rich element and a U-rich element upstream and/or downstream of the cleavage site. The core 3' end machinery in yeast mainly consists of highly conserved cleavage and polyadenylation factor (CPF) and cleavage factors IA and IB (CF IA and CF IB). CPF is a large protein complex composed of 15 different subunits. CF IA is a complex of four proteins: Rna14, Rna15, Clp1, and Pcf11 while CF IB is a single protein (also called Nab4/Hrp1). All these factors together provide a complex 3' processing machinery leading to the endonucleolytic cleavage of mRNA. After the cleavage, mRNA processing machinery adds poly(A) tail to the transcript. The final number of adenosines added can vary from 70-90 in yeast to around 200 in humans. Immediately after synthesis, the poly(A) tail is bound by poly(A) binding proteins. The poly(A) tail helps to protect the mRNA from exonucleases and is also vital for efficient mRNA export (Mariconti, Loll et al. 2010, Chan, Choi et al. 2011, Neve, Patel et al. 2017).

An important RNA binding zinc finger protein involved in the regulation of poly(A) length is Nab2 (nuclear abundant poly(A) RNA-binding protein 2). Nab2 is localized both in the nucleus and cytoplasm but is more abundant in the nucleus, hence the name. It interacts with another poly(A) binding protein Pab1 in order to control poly(A) length (Batisse, Batisse et al. 2009). Furthermore, Nab2 is also required for mRNA export (Green, Marfatia et al. 2002, Schmid, Olszewski et al. 2015). It binds through its N-terminal domain to the nuclear pore associated protein Mlp1 (Fasken, Stewart et al. 2008). Deletion of this domain causes severe growth defects and bulk accumulation of poly(A) RNA in the nucleus. Both functions of Nab2 - poly(A) length control and mRNA export, are essential in yeast (Anderson, Wilson et al. 1993).

1.1.3 mRNP formation and nuclear export

During the process of transcription and pre-mRNA processing, various proteins bind to mRNA co-transcriptionally and post-transcriptionally. After the processed form of mRNA leaves the

transcription site, it is associated with a number of mRNA binding proteins, which together assemble and package the mRNA into export competent mRNPs. These mRNPs then enter into interchromatin environment, where they may be further remodeled to form mature mRNPs. The remodeled and processed form of mRNPs then interact with export machinery to finally perforate through the nuclear membrane to the cytoplasm, where translation begins (Bjork and Wieslander 2017).

Nuclear pore complexes (NPC) are the channels which exist across the nuclear membrane and are responsible for regulated exchange of materials between nucleus and cytoplasm. Nucleoporins are the proteins which make the building blocks of these complexes. The dense network of nucleoporins provides a barrier for improper flow of macromolecules through the membrane. The mRNPs ready for nuclear export need to cross this barrier in order to be transported across the nuclear membrane (Katahira 2015). This is achieved by the special heterodimeric export receptor Mex67-Mtr2 in yeast and Tap-p15 or Nxf1-Nxt1 in higher organisms. Mex67-Mtr2 heterodimer is essential for cell viability and a thermosensitive mutation in *MEX67* shows a high amount of poly(A) RNA accumulation in the nucleus. It interacts with the poly(A) RNA via its subunit Mex67 and to nuclear pore proteins via Mtr2 (Santos-Rosa, Moreno et al. 1998).

Recruitment of a mature mRNP to Mex67-Mtr2 receptor for its nuclear export is a highly coordinated process and requires a number of factors or protein complexes. The evolutionary conserved transcription and export complex (TREX complex) plays an important role in this process. The involvement and requirement of the TREX complex for the efficient export of mRNPs is discussed in Section 1.2. Additionally, the mRNA proteins Npl3 and Nab2 also act as adaptor proteins for Mex67-Mtr2 receptor. Npl3 is a multifunctional SR-like protein which functions in transcription, mRNA processing, export and translation. It is recruited to the transcriptional machinery at early stages of mRNA formation. It interacts with the CTD and remains associated with the mRNA until its export to the cytoplasm. It is also essential for cell viability and mutations in *NPL3* lead to defects in the mRNA export (Hackmann, Gross et al. 2011, Santos-Pereira, Herrero et al. 2014). Nab2 interacts with Mex67 and Yra1 (a component of TREX complex) and also plays a key role in mRNP biogenesis. After the matured mRNP containing Nab2 and other mRNP biogenesis factors transverses through nuclear pore channels, a DEAD-box helicase Dbp5 is recruited on the cytoplasmic side of NPC. Dbp5 causes remodeling of the mRNP in such a manner that Mex67 and Nab2 are dissociated and released,

further causing its disassembly. After the release of Nab2, it is recycled back to the nucleus (Batisse, Batisse et al. 2009).

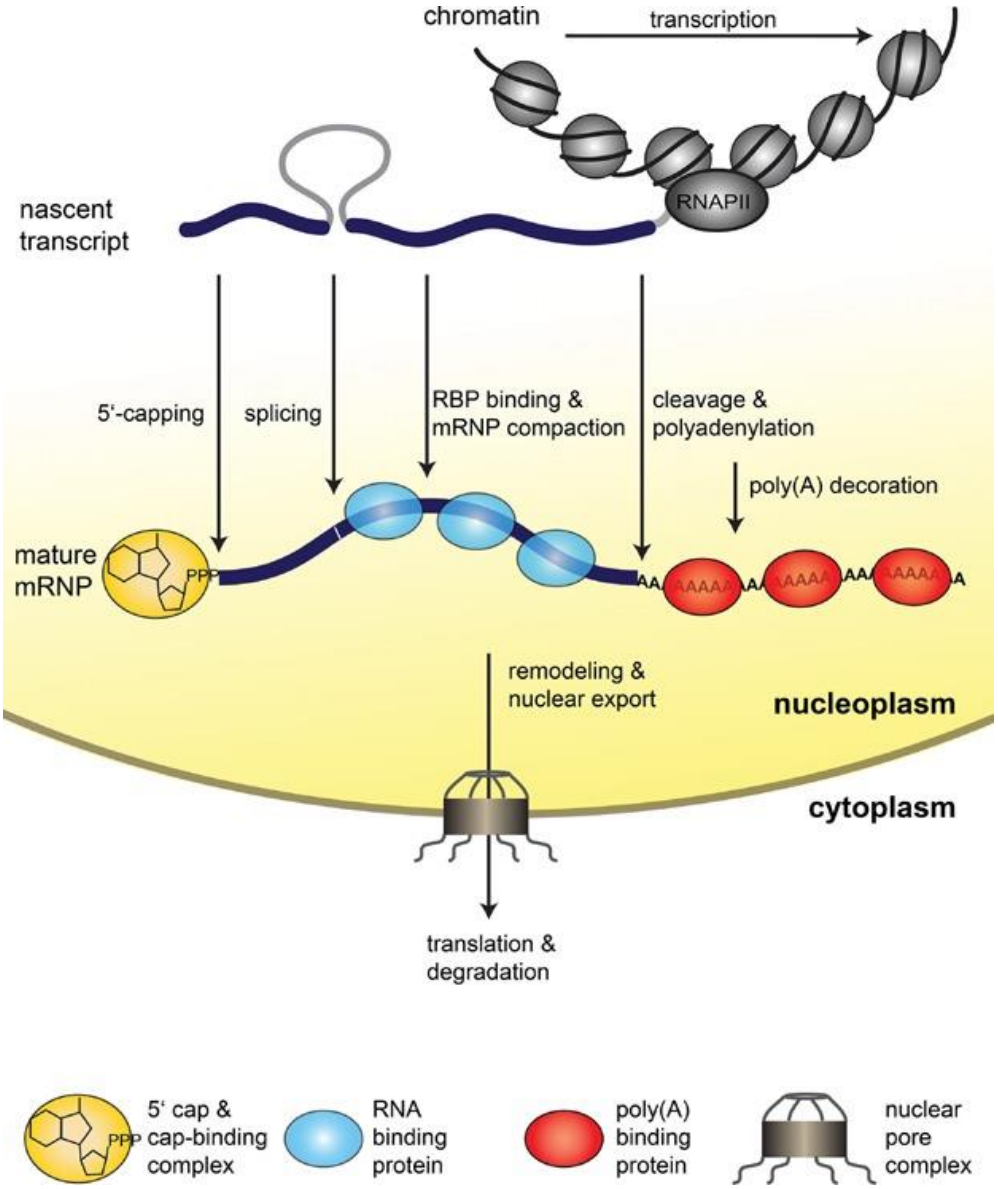


Figure 2. Steps of mRNP biogenesis (Meinel and Strasser 2015).
The nascent mRNA transcribed by RNA Polymerase II (RNAPII) undergoes various processing events such as capping at the 5' end, removal of introns (splicing), cleavage and polyadenylation at the 3'end. It is also bound by many RNA binding proteins, which followed by various remodeling events, leads to the formation of a matured messenger ribonucleoprotein particle (mRNP). This mRNP is finally exported from the nucleus to the cytoplasm for translation.

1.1.4 Coupling of transcription with mRNP formation and nuclear export

All the steps involved in gene expression from transcription to mRNA processing and export are highly coupled and interlinked to ensure quality control. As soon as the nascent mRNA emerges out of the transcription bubble, it is capped by capping machinery. Splicing of mRNA also takes place cotranscriptionally. One of the key factors which plays an important role in coupling transcription to mRNA export is the TREX complex. In order to regulate complex steps of gene expression, the cell's transcriptional machinery employs three different recruitment platforms as outlined below.

1.1.4.1 C-terminal domain (CTD) of RNAPII

The CTD of Rpb1, the largest subunit of RNAPII, plays an important role in coordinating different steps of gene expression. It consists of tandem repeats of a highly conserved and consensus heptapeptide: YSPTSPS (Figure 3). This feature of RNAPII is unique and only exists in eukaryotes, although the number of heptapeptide repeats vary from one species to another. There are 26 of such repeats in yeast and 52 in human. (Zhang, Rodriguez-Molina et al. 2012). Minimum number of heptapeptide repeats required in yeast to maintain cell viability is eight (West and Corden 1995).

The CTD provides a recruitment platform for various mRNA binding proteins, transcription and processing factors involved in mRNP biogenesis. Several residues in the CTD are subjected to modifications. For example, the amino acids tyrosine, threonine, and all the three serines can be phosphorylated and dephosphorylated during the course of transcription. Moreover, glycosylation on threonine and serine residues and isomerization on proline residues has also been observed. Therefore, based on differential modification patterns in these heptads, CTD can attain a noteworthy number of different conformational states (Hsin and Manley 2012, Zhang, Rodriguez-Molina et al. 2012).

Initiation of transcription is marked by the phosphorylation of CTD on its residues Ser5 and Ser7. A component both from TFIIH complex and mediator complex is a cyclin-dependent kinase, known as kinase Kin28 and Srb10 respectively in yeast (named as Cdk7 and Cdk8 respectively in humans). Both Kin28 and Srb10 can phosphorylate the residue Ser5 while Ser7 is phosphorylated by Kin28 kinase only. Ser5 phosphorylation decreases the affinity of RNAPII

from the mediator complex and helps in promoter clearance, so that RNAPII can efficiently proceed for transcription. Additionally, Ser5 phosphorylation also recruits the protein Set1 which trimethylates H3K4 histone, leading to various chromatin remodeling events important for efficient transcription (Zhang, Rodriguez-Molina et al. 2012). Ser5 phosphorylation also recruits as well as directly binds to the capping enzyme subunit Ceg1, which helps in its rapid activation. Furthermore, the yeast spliceosomal complex is known to interact with phosphorylated Ser5 (Ser5P) CTD and several splicing intermediates in higher organisms also copurified Ser5P CTD, supporting the significant role of the CTD in cotranscriptional splicing (Harlen, Trotta et al. 2016, Harlen and Churchman 2017). After the promoter clearance, Ser5 marks are gradually removed from the site of transcription by the CTD phosphatase Rtr1. (Zhang, Rodriguez-Molina et al. 2012, Harlen and Churchman 2017).

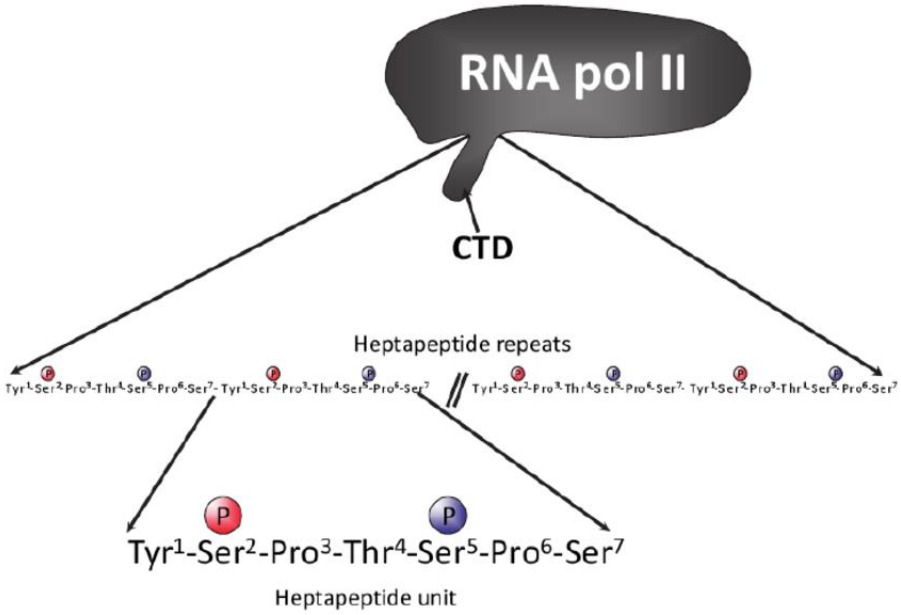


Figure 3. The Rpb1 CTD (Adam Hall and Georgel 2011).
The CTD of Rpb1, largest subunit of RNAPII consists of heptapeptide repeats of the consensus sequence YSPTSPS. Except for proline residues, all other residues in these repeats can be phosphorylated, and these phosphorylation-dephosphorylation patterns among these residues provide a fine platform for the recruitment of various mRNA binding proteins and mRNP biogenesis factors at the site of transcription.

After the transition from initiation to elongation phase, CTD attains a different phosphorylation pattern in order to allow the recruitment of factors which promote elongation. This is marked by an increase in Ser2 phosphorylation and a gradual decrease in Ser5 phosphorylation. More specifically, from TSS to ~450 nt downstream, irrespective of the length of the gene, both Ser5

and Ser2P phosphorylation coexist, with gradual increase of Ser2 phosphorylation and gradual decrease or removal of Ser5 phosphorylation along the length of the gene. Around ~600 nt downstream of TSS, Ser2 phosphorylation gets saturated and remains at its peak till 3' end of the gene, until it starts decreasing downstream of poly(A) site. In *S. cerevisiae*, Bur1/Bur2 cyclin-dependent kinase complex regulates transcription through histone mono-ubiquitylation, recruitment of Paf1 complex and CTD phosphorylation. Bur1 (Cdk9 in humans) is recruited as a result of Ser5 phosphorylation at the promoter. It further phosphorylates Ser2, which recruits Ctk1, a Ser2 kinase (Wood and Shilatifard 2006, Hsin and Manley 2012).

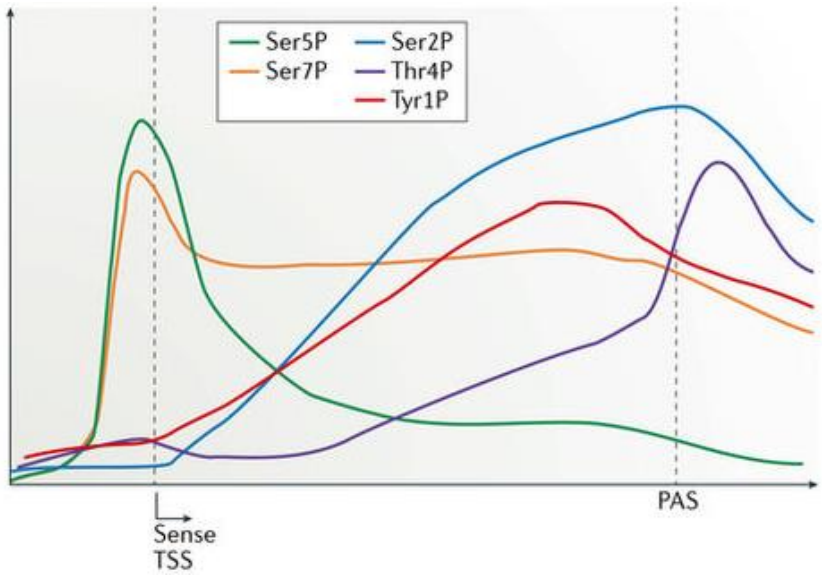


Figure 4. The phosphorylation pattern of amino acid residues in a CTD heptapeptide unit during the course of transcription (Harlen and Churchman 2017).
Initiation of transcription is marked by the phosphorylation of Ser5 and Ser7. The transition from initiation to elongation phase is marked by an increase in Ser2 phosphorylation. Tyr1 phosphorylation is also present downstream of TSS but sharply declines upstream of poly(A) site. Thr4 phosphorylation levels increase downstream of poly(A) site.

Ser2 phosphorylation plays a pivotal role in the interaction of RNAPII with mRNA processing machinery. For example, Prp40, a protein involved in splicing binds to Ser2P and residual Ser5P CTD in yeast (Morris and Greenleaf 2000). Mammalian splicing protein U2AF65 binds to the phosphorylated CTD. The binding of U2AF65 to the RNAPII CTD further promotes recruitment of both U2AF65 and Prp19C at the spliceosomal site and hence activates the splicing process (David, Boyne et al. 2011). Npl3, an RNA-binding protein involved in elongation, termination, 3' end processing and export, also binds to Ser2P (Zhang, Rodriguez-

Molina et al. 2012). Ser2 phosphorylation also plays a central role in the recruitment of cleavage and polyadenylation machinery. In yeast, Pcf11, a component of cleavage factor IA, binds to Ser2P. Ser2P also binds directly and probably helps in the recruitment of Rtt103, another protein involved in transcription termination machinery (Harlen and Churchman 2017).

Other than Ser2P, Tyr1P is also involved in the regulation of transcription and termination. Although in lower amounts, Tyr1 phosphorylation is detected downstream of TSS. Around 100-200 nt downstream from poly(A) site, there is a sharp decrease in Ser2P, while the Tyr1P already falls down around 200 nt upstream of poly(A) site. The role of Tyr1P is to block the termination factors Pcf11 and Rtt103, so that they can't terminate transcription before the poly(A) site. This also explains why levels of Tyr1P fall before poly(A) site when the transcript is ready to terminate (Mayer, Heidemann et al. 2012). Downstream of the poly(A) site, Thr4P levels increase. Thr4P has been shown to interact with the termination factor Rtt103. It has been further suggested that both Ser2P and Thr4P together recruit Rtt103 and thereby ensure efficiently controlled transition from elongation to termination (Harlen, Trotta et al. 2016). The phosphorylation patterns of amino acid residues in a CTD heptapeptide unit during the course of transcription are depicted in Figure 4.

1.1.4.2 The nascent RNA

The nascent RNA itself can act as a platform for binding of various factors mRNP biogenesis. For example, during splicing, RNA components of snRNPs do base pairing with the nascent RNA and also promote interaction of the spliceosome with nascent RNA (Meinel and Strasser 2015). The proteins involved in the 3' end processing machinery recognize various sequence elements (called *cis* elements) in the pre-mRNA. The three most important sequence elements identified in mammalian 3' end processing machinery are the polyadenylation signals (PAS), the cleavage site and the downstream element (DSE). Lack or disruption of any of these elements can lead to defects in 3' end processing (Mandel, Bai et al. 2008). Many other proteins involved in mRNP biogenesis also have preferences for specific kind of sequences. For example, TREX components Gbp2 and Hrb1 recognize highly degenerate sequence motifs in the RNA. Also, the RNA binding protein Nab2 binds only to A-rich motifs (Meinel and Strasser 2015).

1.1.4.3 Elongation factor Spt5

Spt5 is a part of the conserved Spt4-Spt5 complex that is known to play a role in RNAPII elongation. It helps to increase the processivity of RNAPII transcription by binding directly to RNAPII and ensuring engagement of DNA inside the elongation complex (Schneider, French et al. 2006, Blythe, Yazar-Klosinski et al. 2016). Spt5 has also been identified as a recruitment platform for various transcription and processing factors. A mass spectrometric analysis has shown that Spt5 in yeast copurifies several proteins involved in mRNP biogenesis (Lindstrom, Squazzo et al. 2003). Like the CTD of RNAPII, the C-terminal region of Spt5 (CTR) also contains hexa-repeats and Ser1 residue on these repeats can be phosphorylated by kinase Bur1 during transcription elongation. This phosphorylation pattern affects the recruitment of many other factors. For example, the Paf1 complex, which associates with RNAPII and plays several roles throughout the transcription cycle, interacts with Spt5 and its recruitment to chromatin is modulated by the phosphorylation of Spt5 (Squazzo, Costa et al. 2002, Jaehning 2010). Spt5 also stabilizes the recruitment of various enzymes involved in mRNA capping (Lidschreiber, Leike et al. 2013). Also, the CTR of Spt5 functions in recruitment of the pre-mRNA cleavage factor I (CFI) to genes. The deletion of CTR of Spt5 confers 6-azauracil (6AU) sensitivity to the cells (Mayer, Schreieck et al. 2012). Thus, Spt5 provides an important platform for the recruitment and association of various mRNP biogenesis factors.

1.2 TREX complex

Different steps in the gene expression from transcription to mRNA processing and export are performed by highly complex and multi-component cellular machineries. All these distinct machineries are physically and functionally coupled to each other. A key player involved in such coupling is the highly conserved TREX complex, which couples transcription elongation to mRNA export. In *S. cerevisiae*, TREX exists as a heterononameric complex, consisting of a pentameric subcomplex called THO complex, along with two SR-like proteins Gbp2 and Hrb1, the DEAD-box helicase Sub2 and Mex67-Mtr2 adaptor protein Yra1 (Figure 5). The THO complex is sub composed of Tho2, Hpr1, Mft1, Thp2 and Tex1 (Strasser, Masuda et al. 2002, Katahira 2012, Heath, Viphakone et al. 2016).

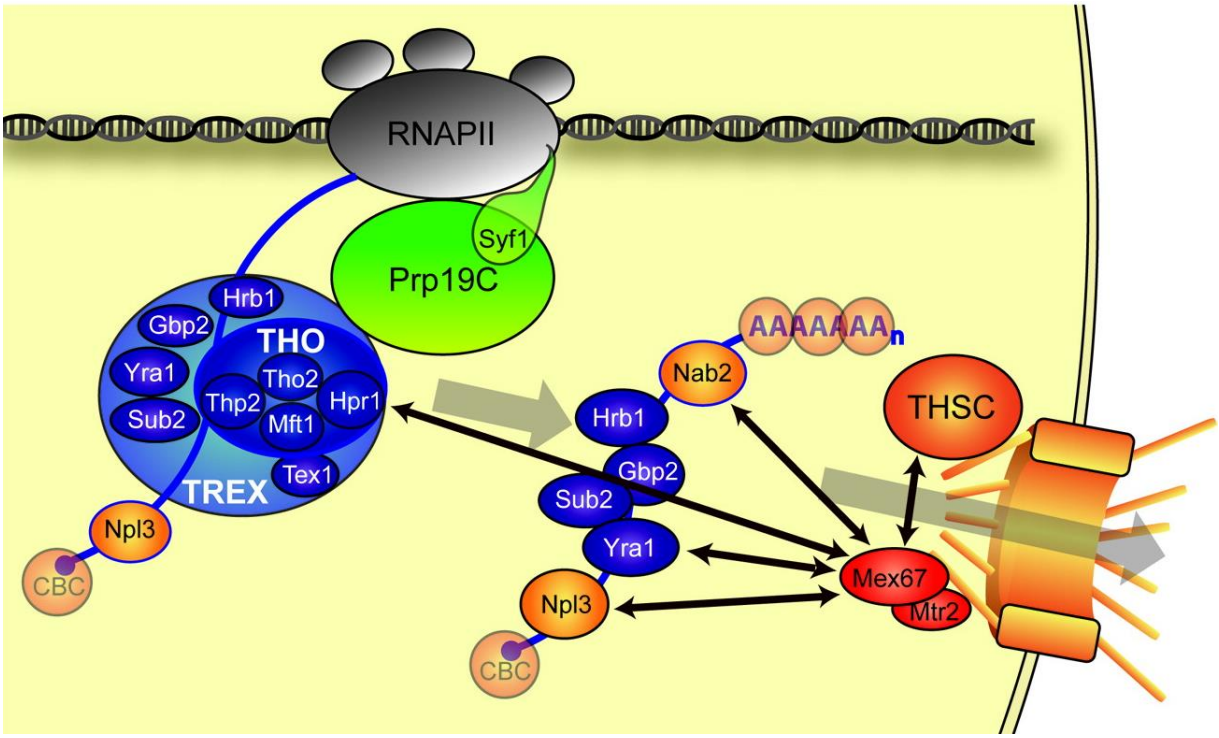


Figure 5. Role of TREX, Prp19C, the export receptor Mex67-Mtr2 and other mRNA binding proteins in mRNP biogenesis (Chanarat, Burkert-Kautzsch et al. 2012).
The highly conserved TREX complex is recruited to the transcriptional machinery and couples transcription to mRNA export. The C-terminus of Syf1, a component of Prp19C, stabilizes the recruitment of TREX complex at transcribed genes. The nascent RNA is bound by several mRNA binding proteins, which process and remodel the RNA to form an mRNP. The Mex67-Mtr2 export receptor is recruited to the processed mRNP by the direct interaction with TREX components Yra1 and Hpr1, Nab2, Npl3 and the THSC complex. The mRNP is finally exported through nuclear pore to the cytoplasm.

During the active transcription by RNAPII, TREX is recruited to transcriptional machinery and travels along with RNAPII and several other factors throughout the length of the gene. The composition of TREX and its function in mRNA export is conserved from yeast to higher organisms although its recruitment mechanism varies. In *D. melanogaster*, THO consists of yeast orthologs Hpr1, Tex1 and Tho2 as well as three other components: Thoc5, Thoc6 and Thoc7, which represent the functional analogs of Mft1 and Thp2 (Rehwinkel, Herold et al. 2004). In mammalian cells, THO also consists of hHpr1, hTex1 and hTho2 as well as counterparts of Thoc5, Thoc6 and Thoc7 (fSAP79, fSAP35, and fSAP24 respectively). Like in yeast, TREX in higher organisms also contain UAP56 and REF1/Aly (homologs of Sub2 and Yra1) associated with the THO complex (Masuda, Das et al. 2005).

The THO complex was originally discovered to be involved in transcription elongation (Jimeno, Rondón et al. 2002). Although null mutations of the components of THO complex are viable, they exhibit phenotypes similar to that of the defects in transcription elongation. Additionally,

the null mutants of components of THO complex display phenotypes of transcription-dependent hyper-recombination (Chavez, Beilharz et al. 2000). The hyper-recombination phenotype is linked to the formation of DNA/RNA hybrids, also known as R-loops. Formation of R-loops occurs during transcription when nascent transcript folds back and hybrids to the template DNA, thereby causing a pause in transcription elongation. The formation of these loops can mediate transcription elongation defects and leads to transcription-associated genetic instability (Huertas and Aguilera 2003). Thus, the THO complex plays an important role in ensuring the formation of correct mRNP and also by providing an mRNA quality control mechanism. Apart from this, the THO complex has also been shown to be required for cell development and differentiation (Wang, Chang et al. 2006).

The TREX subunits Sub2 and Yra1 interact physically and genetically with the THO complex. The THO complex loads Sub2 and Yra1 to transcriptional site for the subsequent export of mRNPs. It has been shown that Yra1 interacts with Mex67 and recruits the Mex67-Mtr2 receptor to the mRNP. Also, mutations in *YRA1* leads to bulk poly(A) mRNA accumulation at the non-permissive temperature (37°C) (Strasser and Hurt 2000). Sub2, another essential subunit of the TREX complex, is a DEAD-box RNA helicase which interacts with Yra1 *in vivo* and *in vitro.* The mutations in *SUB2* as well as the overexpression of Sub2 leads to mRNA export defects. Interestingly, Yra1 uses the same domain to interact both with Sub2 and Mex67-Mtr2 export receptor and it was suggested that Sub2 recruits Yra1 to the transcriptional machinery (Strasser and Hurt 2001). It has also been reported that Pcf11 (a component of the yeast cleavage factor IA) interacts with Yra1, and this interaction is important for recruitment of Yra1 to the transcriptional machinery. Interestingly, interaction between Pcf11 and Yra1 also requires the same domain of Yra1 as required for Sub2-Yra1 interaction (Johnson, Cubberley et al. 2009). Thus, the function of TREX in transcription elongation and export is also coupled to the cleavage and polyadenylation machinery.

The other two subunits of TREX complex, Gbp2 and Hrb1, belong to the serine-arginine-rich (SR) family of proteins. They are recruited to actively transcribed genes and remain associated throughout the length of the gene. It has been shown that both of them interact with Ctk1 kinase, a kinase required for the phosphorylation of Serine 2 in RNAPII CTD. Thus, Gbp2 and Hrb1 are also important for efficient transcription elongation (Hurt, Luo et al. 2004).

The TREX complex interacts with various factors of the transcriptional machinery in order to ensure efficient nuclear export of properly packaged mRNPs. For example, Tho1, an mRNA binding protein, interacts with the transcriptional machinery in a THO dependent manner.

Overexpression of Tho1 can suppress mRNA export defects and hyper-recombination phenotypes in THO mutants (Jimeno, Luna et al. 2006). Moreover, the non-snRNP Prp19C also interacts with TREX in a transcription dependent manner. Prp19C helps to stabilize TREX at transcribed genes and is also required for efficient transcription elongation (Chanarat, Seizl et al. 2011).

In *S. cerevisiae*, TREX is recruited cotranscriptionally to the genes in a transcription-dependent manner (Strasser, Masuda et al. 2002, Zenklusen, Vinciguerra et al. 2002). TREX recruitment to transcriptional machinery occurs via interaction of the THO subcomplex with the S2 phosphorylated CTD (Meinel, Burkert-Kautzsch et al. 2013). On the other hand, in higher organisms, TREX is known to be recruited during splicing rather than transcription. Aly, the human homolog of Yra1, interacts with cap-binding protein CBP80 and it might also be possible that TREX is recruited in a cap-dependent manner in higher organisms. (Rehwinkel, Herold et al. 2004, Masuda, Das et al. 2005, Cheng, Dufu et al. 2006).

1.3 The splicing protein Mud2

Splicing of pre-mRNA requires a series of events which assemble the macromolecular spliceosomal complex in an activated form. This activated enzymatic complex then carries out the splicing reaction. The first event in this process is marked by the interaction of U1 snRNP with pre-mRNA. U1 snRNP can interact with 5' as well as 3' of the splice site, leading to formation of two early complexes called commitment complex 1 and 2 (CC1 and CC2) in yeast and complex E in mammals. While CC1 is formed by the interaction of U1 snRNP with 5' splice site only, CC2 requires the branchpoint sequence and a number of other factors for its assembly (Rutz and Seraphin 1999).

The protein Mud2 was first identified in a genetic screen for identifying factors involved in branchpoint recognition and commitment complex formation. It was shown that Mud2 interacts with U1 snRNP to form CC2 and also interacts with the pre-mRNA in a branchpoint sequence dependent manner. Also, Mud2 affects the addition of U2 snRNP. The name 'Mud2' comes from 'mutant U1 die', which means the otherwise viable mutants of U1snRNP are synthetically lethal with *MUD2* deletion. Prp11, a component of U2 snRNP, interacts directly with Mud2. Mud2 is similar to the well-known mammalian splicing factor U2AF65 (Abovich, Liao et al. 1994).

Mud2 interacts with the branchpoint binding protein (yBBP/Msl5) in both yeast and mammalian cells. While Mud2 is a non-essential protein, Msl5 is required for the cell viability. Mud2-Msl5 heterodimer is present in CC2 but is absent in the pre-spliceosomal complex as well as in the mature spliceosome. Thus, Mud2 and Msl5 are released from the spliceosomal machinery during the transition from CC2 phase to pre-spliceosomal complex formation phase. Also, *in-vitro* studies showed that Mud2 and Msl5 do not affect the formation of pre-spliceosomal complex. Moreover, no splicing defects have been observed if *MUD2* is disrupted or *MSL5* is depleted, though CC1 complex accumulation has been observed in *MUD2* deleted and *MSL5* depleted cells (Rutz and Seraphin 1999, Wang, Zhang et al. 2008). This suggests that both of these proteins do not play a direct role in splicing but are required for splice site recognition and recruitment of spliceosomal machinery in the initial stages of splicing. Both of them are recycled after the release from spliceosomal machinery, either during or before the formation of pre-spliceosomal complex (Rutz and Seraphin 1999, Wang, Zhang et al. 2008).

Sub2, a component of TREX complex, also plays a role in pre-mRNA splicing. Sub2 is required for the formation of CC2 complex and also helps in the stable association of U2 snRNP to the branchpoint region. *SUB2* depletion leads to an accumulation of CC2 complex as well as decrease in the pre-spliceosomal complex (Libri, Graziani et al. 2001, Zhang and Green 2001). Previous studies reported that Sub2 interacts with Mud2 *in vitro* (Fan, Merker et al. 2001). Furthermore, this interaction is antagonistic. The otherwise lethal deletion of *SUB2* at 25°C is viable when *MUD2* is deleted. Thus, in the absence of Mud2, the essential function of Sub2 becomes dispensable. This further suggests that Sub2 helps in the removal of Mud2 from spliceosomal machinery, which is required for the effective formation of pre-spliceosomal complex (Kistler and Guthrie 2001).

Besides its interaction with Sub2, Mud2 is also shown to interact with Syf3, an essential component of well-known splicing complex Prp19C (Chung, McLean et al. 1999, Vincent, Wang et al. 2003). The interaction of Mud2 with Prp19C has also been shown in mammalian cells by Manley's lab (David, Boyne et al. 2011). They showed that U2AF65, the mammalian Mud2, interacts with Prp19C *in vivo* as well as *in vitro* in an RNA-independent manner. This interaction is CTD-dependent and is required for the activation of splicing (David, Boyne et al. 2011).

Human U2AF65 and Drosophila U2AF50 are the presumptive orthologues of Mud2. They all contain three RNA recognition motifs (RRMs) at the C-terminus (Figure 6). The three RRMs

of U2AF65 (marked as RRM1, RRM2 and RRM3) are quite closely related and share remarkable similarities to the three RRMs of U2AF50. On the contrary, in yeast, only RRM3 is well recognized and closely related to the mammalian Mud2 (with around 30% sequence identity; (Soucek, Zeng et al. 2016). RRM1 and RRM2 are required for binding of Mud2 to the polypyrimidine tract (Sickmier, Frato et al. 2006), while RRM3 is required for interacting with Msl5 (Selenko, Gregorovic et al. 2003). As polypyrimidine tract region is not a prominent feature of yeast introns, it explains why RRM1 and RRM2 are not so well characterized in yeast Mud2.

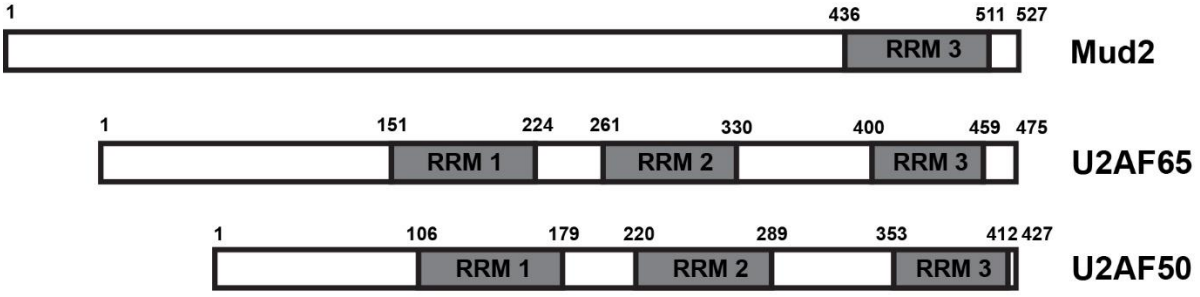


Figure 6. Schematic representation of domain structure of Mud2, U2AF65 and U2AF50.
The boxes represent the primary structure of the labeled protein and the numbers represent the amino acid positions in the primary sequence. The shaded regions inside the boxes indicate the RRM motifs. Mud2 in yeast, human (U2AF65) and *Drosophila* (U2AF50) consist of three RNA Recognition Motifs (RRMs) at the C-terminus. While the three motifs are quite well characterized in higher organisms (U2AF65 and U2AF50), only RRM3 is more conserved and well recognized in yeast Mud2.

The RNA binding protein Nab2 also associates to the spliceosomal machinery and thus, couples splicing with 3' end mRNA processing. The *nab2* mutant cells accumulate splicing defects. The genetic interaction analysis revealed that a *nab2* mutant combined with the deletion of *MUD2* results in severe growth defects. The RRM3 domain of Mud2 is critical for its interaction with Nab2. Mud2 also interacts physically with Nab2 and this interaction is conserved in ZC3H14, the mammalian orthologue of Nab2. Thus, Nab2 participates in the functional coupling of splicing with mRNA processing by interacting with the spliceosomal machinery (Soucek, Zeng et al. 2016).

PAR-CLIP analysis of mRNP biogenesis factors showed that Mud2 is not only bound to the intron-containing genes but also to the intronless genes. Also, Mud2 tends to co-localize with factors like Hpr1, Hrb1, Nab2, and Npl3 (Baejen, Torkler et al. 2014). This strongly indicates that besides its defined role in splicing, Mud2 might also play a role in the other steps of mRNP biogenesis.

1.4 Prp19 complex (Prp19C)

Prp19C, also known as the NineTeen Complex or NTC complex, is an evolutionary conserved, non-snRNA containing and multisubunit protein complex, which was first shown to be involved in the pre-mRNA splicing. It was identified in *S. cerevisiae* as a complex containing Prp19 subunit and at least 10 other core proteins along with some loosely associated proteins (Ohi, Vander Kooi et al. 2005). The composition of Prp19C and growth requirement of its components in yeast is summarized in Table 1.

Table 1. Prp19C composition and growth requirement of its components.

Components of Prp19C	MW (kDa)	Essential for viability in yeast?
PRP19	56.6	Yes
SYF1/NTC90	100	Yes
CEF1/NTC85	68	Yes
CLF1/SYF3/NTC77	82.4	Yes
PRP46/NTC50	50.7	Yes
CWC2/NTC40	38.4	Yes
SYF2/NTC31	25	No
ISY1/NTC30	28	No
SNT309/NTC25	21	No
NTC20	16	No
CWC15	20	No

The Prp19C is conserved from yeast to the higher organisms, although its composition varies. In mammalian cells, it is called as Prp19/CDC5 complex and consists of hPrp19, CDC5L (yCef1p), PRL1 (yPrp46p), SPF27 (ySnt309p) and other proteins not having homologous partners in yeast. Besides Prp19/CDC5 complex, mammalian cells also contain two other complexes containing Prp19 protein and few other human homologues of yeast Prp19C subunits: PRP19-associated complex and the XAB2 complex. In *Drosophila*, two protein complexes with Prp19 have been identified, which are similar to human Prp19/CDC5 complex and PRP19-associated complex (Chanarat and Strasser 2013).

Prp19C is best known for its role in splicing. Prp19C along with some other Prp19-associated proteins interacts with the spliceosome during or after the dissociation of U4 snRNP. It helps in the stabilization of U5 and U6 snRNA by promoting RNA-RNA interactions between pre-mRNA and the U5 and U6 snRNA. Following these interactions, spliceosome undergoes a couple of rearrangements and structural remodeling which leads to the formation of an activated form of spliceosome, ready for the splicing reaction. Prp19C remains associated with the spliceosome until the second catalytic step of splicing (Tarn, Hsu et al. 1994, Chan, Kao et al. 2003, Chan and Cheng 2005, Hogg, McGrail et al. 2010). The depletion or deletion of the components of Prp19C leads to accumulation of free U4 snRNP, which provides an evidence that Prp19C might also play a role in the spliceosome biogenesis and recycling (Hogg, McGrail et al. 2010).

Besides its widely characterized role in splicing, Prp19C is also known to function in genome maintenance, cell cycle progression and ageing. Defects in the cell cycle and DNA repair have been observed in *prp19* mutant cells in yeast. In humans, mutations in *PRP19* leads to double-strand breaks in DNA (Chanarat and Strasser 2013).

Prp19 (precursor RNA processing protein 19) is one of the most integral and essential components of Prp19C. It interacts with a number of other components of the complex and is also required to maintain the structural integrity of the whole complex (Chen, Yu et al. 2002). The temperature sensitive mutants of *PRP19* are defective in the splicing (Ruby and Abelson 1991). The Prp19 protein contains a U-box at the N-terminus, a coiled-coil in the center and WD40 repeats at the C-terminus (Figure 7). The basic domain organization of Prp19 is conserved from yeast to the higher organisms, although the number of WD40 repeats vary from one species to another. An X-ray crystal structure of Prp19 has been determined according to which the U-box domain of Prp19C is structurally similar and shares the same interaction surface as of the RING finger domains in the RING-E2 complex. Disruption of this U-box disrupts the function of Prp19C *in vivo.* Additionally, Prp19 oligomerizes to form a stable tetrameric structure. This tetramer consists of a central stalk of four coiled-coils surrounded by four U-box domains and four WD-40 domains, two on each side of the central stalk. This property of Prp19 to tetramerize is essential for its function *in vivo* (Ohi, Vander Kooi et al. 2003, Vander Kooi, Ren et al. 2010).

Prp19 interacts directly with three other components of Prp19C: Cef1, Snt309 and Cwc2. Cef1 and Snt309 interact via the N-terminus of Prp19 whereas Cwc2 interacts via the WD40 domains of Prp19 (Chen, Jan et al. 1998, Tsai, Chow et al. 1999, Ohi and Gould 2002, Ohi, Vander Kooi

et al. 2005). Syf1 is another core component of the Prp19C, which interacts with most of other components of the complex and is important for the structural integrity of whole complex (Chen, Yu et al. 2002).

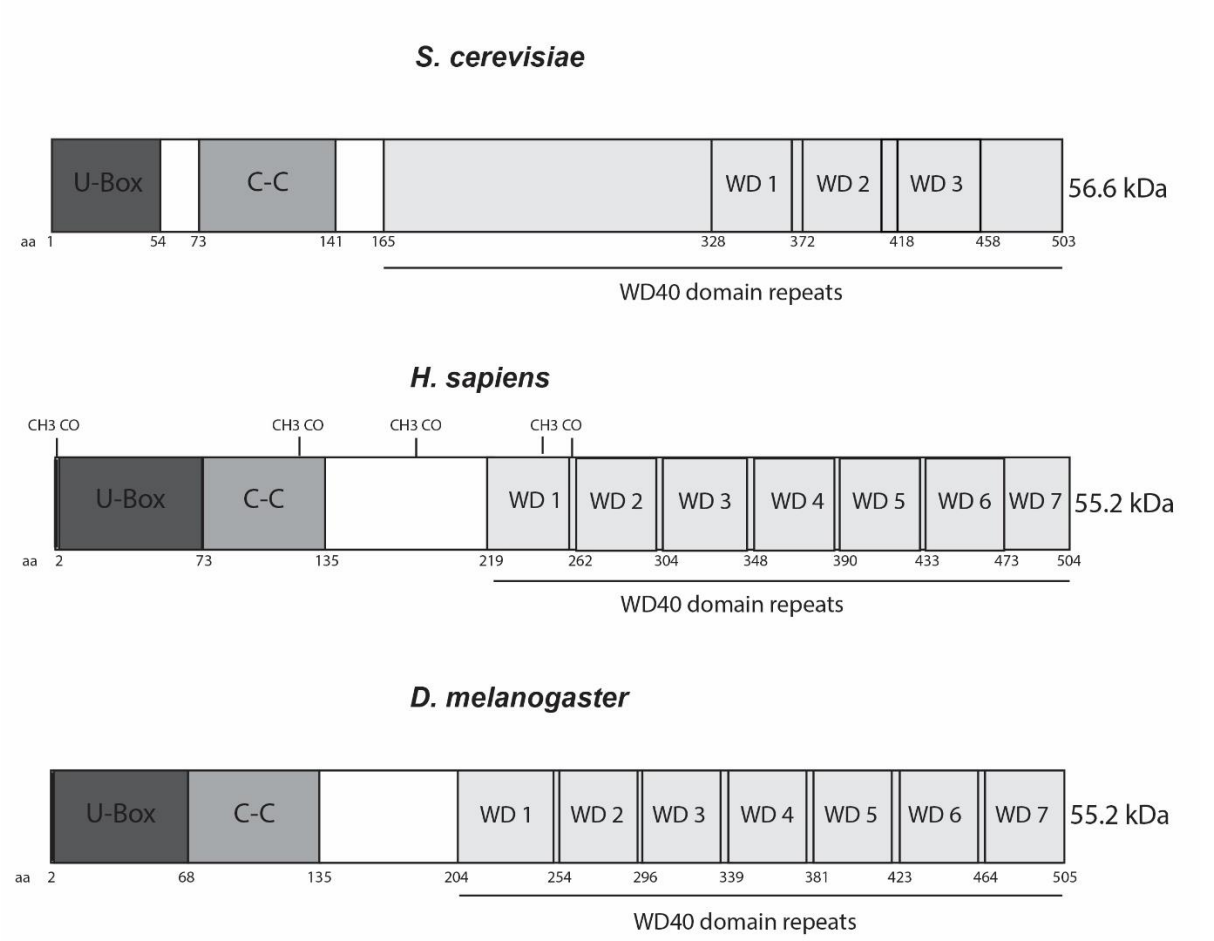


Figure 7. Schematic organization of structural domains of Prp19 in *S. cerevisiae*, *H. sapiens* and *D. melanogaster*.
The conserved Prp19 protein contains a U-box domain at the N-terminus, a coiled-coil domain (C-C) in the center and WD40 repeats at the C-terminus. U-box is required for the E3 ligase activity of the protein. C-C domain is important for the interaction with Cef1 protein and also for the tetramerization of Prp19. Cwc2 subunit of Prp19C interacts with the WD40 domains of Prp19. aa: amino acid position.

In this study, three subunits of Prp19C: Isy1, Ntc20 and Syf2, and a NTC related protein (NTR) Cwc15 (complexed with Cef1) have been focused. These components of Prp19C are recruited to spliceosome in the same manner as Prp19. All these proteins are not essential for yeast growth and their deletion mutants do not show any obvious defects in the cell growth. However, when deletion of *NTC20* is combined with the deletion of *ISY1/NTC30*, the double mutants show severe growth defects and a high accumulation of pre-mRNA. Double deletion of *ISY1/NTC30* and *SYF2/NTC31* is temperature sensitive while the double deletion of *NTC20* and

SYF2/NTC31 does not show any obvious growth defects. Strikingly, deletion of all the three factors is lethal (Chen, Tsai et al. 2001, Chen, Kao et al. 2006). Deletion of *SNT309/NTC25* gives a temperature sensitive phenotype and Δ*snt309* cells also accumulate pre-mRNA at non-permissive temperatures. Snt309 only interacts with Prp19 and not with any other component of the complex. In Δ*snt309* cells, the other components dissociate from the complex leading to the destabilization of whole complex (Chen, Jan et al. 1998). It also leads to the failure in spliceosome recycling. *SYF1* and *SYF3* have common interaction partners, as both of them interact with *NTC85, NTC31, NTC30,* and *NTC20.* Also, they both do not interact either with *SNT309* or *PRP19*. Furthermore, *NTC31, NTC30,* and *NTC20* also show similar interaction profiles, as all of them interact with *CEF1, SYF1* and *SYF3* but not with *SNT309* or *PRP19*. According to the existing model of how all these components are associated in the complex, Snt309 first interacts with Prp19, which then binds to Cef1 followed by the recruitment of Ntc20 and Ntc30. Isy1, Ntc20 and Syf2, which perform overlapping functions, exist together along with Syf1 to form a NTC sub-complex (Figure 8) (Tarn, Hsu et al. 1994, Tsai, Chow et al. 1999, Chen, Yu et al. 2002, Ohi and Gould 2002).

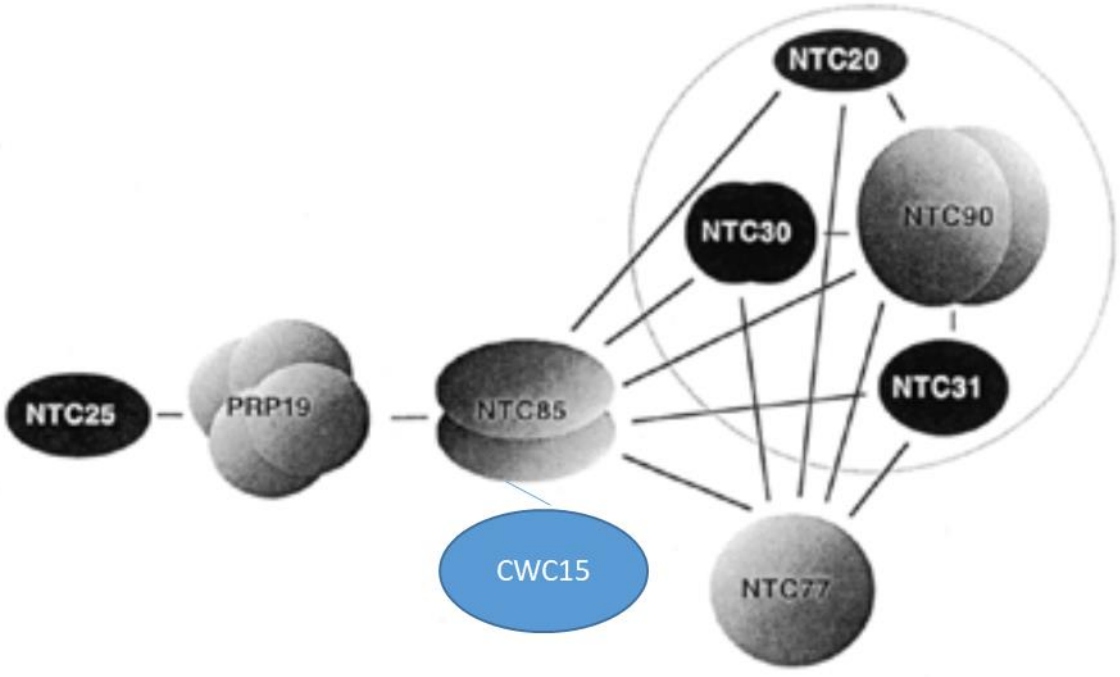


Figure 8. A model representing association of different components of the Prp19C. Figure adapted from (Chen, Yu et al. 2002).
The black boxes represent non-essential proteins while the grey boxes represent essential proteins of the Prp19C. The lines represent the interactions between the two proteins. Cwc15 (blue box) is associated with Ntc85 component of the Prp19C. Ntc20, Ntc30, Ntc31 and Ntc90 exist together in a sub-complex known as NTR sub-complex.

Interestingly, Prp19C also plays a role in transcription elongation by maintaining TREX occupancy at transcribed genes. It interacts with TREX genetically and biochemically. It has been shown previously that a truncation in the C-terminus of Syf1 results in decrease of TREX occupancy (Chanarat, Seizl et al. 2011). This function of Prp19C in stabilizing TREX at genes is independent of splicing, as splicing was is not affected in these mutants. Thus, Prp19C links RNAPII transcription to mRNP biogenesis by stabilizing TREX on the genes (Chanarat, Seizl et al. 2011, Chanarat, Burkert-Kautzsch et al. 2012).

1.5 Aim of this study

Mud2 is required for splice site recognition in the initial stages of splicing. An interaction of Mud2 with the Prp19C subunit Syf3 and the TREX subunit Sub2 has been published before (Chung, McLean et al. 1999, Kistler and Guthrie 2001). Furthermore, previous findings report an interaction of U2AF65 (mammalian Mud2) with the S2 phosphorylated CTD as well as with Prp19C (David, Boyne et al. 2011). All these results strongly suggest that Mud2 might play a role in mRNP biogenesis besides its defined role in splicing. Additionally, the molecular mechanism of how Prp19C functions in maintaining TREX at transcribed genes and in transcription is still unknown. Especially, the roles of the non-essential subunits of Prp19C in coordinating Prp19C-TREX and Prp19C-Mud2 interaction remain elusive.

Therefore, the overall aim of this project was to gain deeper insights into the roles of Mud2 and Prp19C in the highly complex process of mRNP biogenesis. To this end, the following two specific objectives were approached (Figure 9):

1) The first objective of this study was to determine whether Mud2 is recruited to transcriptional machinery for a process distinct from splicing. If it does, what role it plays or what is the possible mechanism involved? In particular, the following questions were asked:

- Is Mud2 recruited to all the transcribed genes including intronless genes? Is the recruitment of Mud2 to the transcribed genes dependent on active transcription? Does Rpb1 CTD play a role in the recruitment of Mud2 (Figure 9A)?
- Does Mud2 interact physically with the Prp19C in *S. cerevisiae* as it does in the mammalian cells (Figure 9A)? Or is the function of Mud2 conserved?

- If Mud2 is present at genes during the course of transcription, does it also affect the occupancy and stability of Prp19C and TREX (Figure 9B)?
- Besides its established role in splicing, does Mud2 function in any other process of mRNP biogenesis, such as transcription, mRNA packaging and nuclear export of mRNPs?

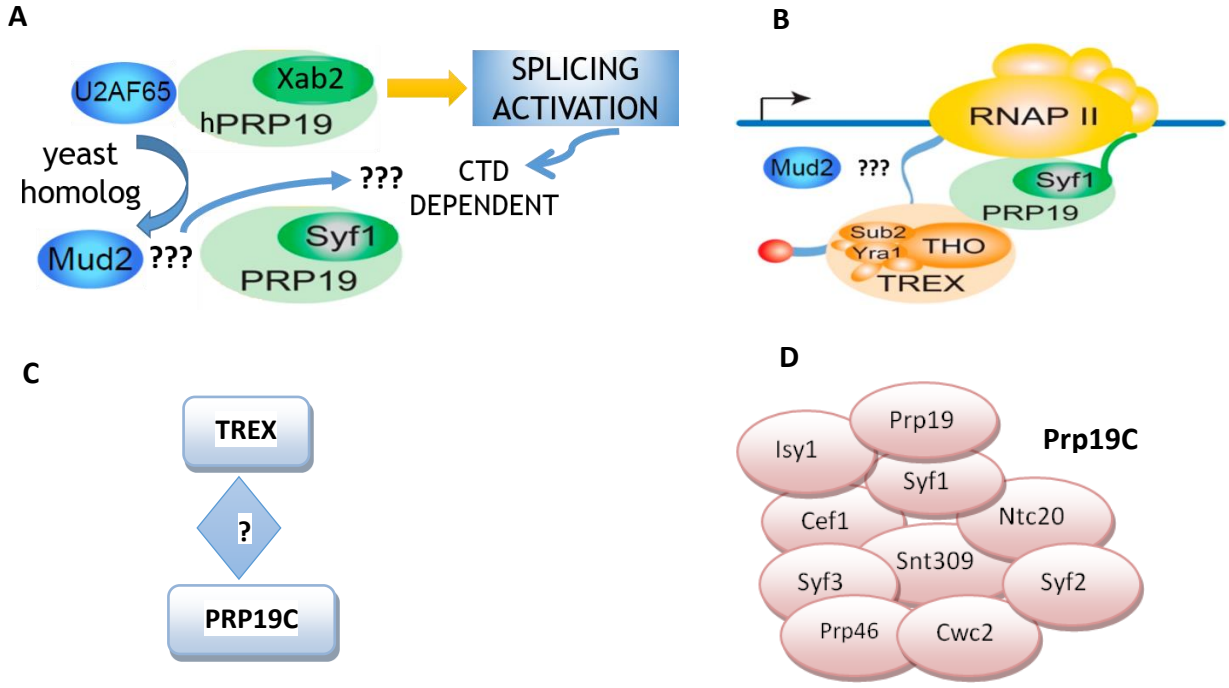


Figure 9. Schematic model to outline the key objectives of this study. Figure 9B adapted from (Chanarat, Seizl et al. 2011).

(A) Interaction of the mammalian Mud2 (U2AF65) with Prp19C and the CTD activates the splicing process. Investigating the role of CTD in recruitment of yeast Mud2 (blue) was one of the objectives of this study. **(B)** Prp19C interacts with TREX and also helps to maintain its occupancy at transcribed genes. Investigating the role of Mud2 in transcription and in maintaining Prp19C and TREX occupancy at transcribed genes was another major aim of this study. **(C)** Whether Prp19C binds directly to THO was also an open question addressed in this study. **(D)** The various subunits of Prp19C forming a stable complex. Only a pictorial representation of all the components of the complex is depicted here and it does not represent their actual interactions within the complex. The second objective of this study was to determine functional significance of the non-essential components of Prp19C in Prp19C-TREX interaction and transcription.

2) The second aim of this project was to unravel the functional significance of the components of Prp19C, especially in relevance to their interaction with TREX. To address this objective, the following questions were asked:

- Does Prp19C interact with the THO directly (Figure 9C)?

- Do non-essential subunits of Prp19C (Figure 9D) play a role in maintaining Prp19C and TREX occupancy at transcribed genes?
- What are the functions of the non-essential Prp19C subunits in Prp19C stability, Prp19C-TREX, Prp19C-Mud2 and Prp19C-RNAPII interaction?
- Do Prp19C non-essential components also play a role in transcription?

2 Materials and Methods

2.1 Materials

2.1.1 Consumables and chemicals

Consumables and chemicals were purchased from the following companies: Applichem GmbH (Darmstadt, Germany), Applied Biosciences (Darmstadt, Germany), Fermentas (St. Leon- Rot, Germany), Bio-Rad (Hercules, USA), Eppendorf (Hamburg, Germany), Fermentas (St. Leon-Rot, Germany), GE Healthcare Europe (Freiburg, Germany), Invitrogen (Karlsruhe, Germany), Life Technologies (Carlsbad, USA), Macherey&Nagel (Düren, Germany), Millipore (Molsheim, France), Mobitec (Göttingen, Germany), Beckman Coulter (Krefeld, Germany), Open Biosystems (Huntsville, USA), Promega (Mannheim, Germany), Qiagen (Hilden, Germany), Roche (Mannheim, Germany), Sarstedt (Nümbrecht, Germany), Sigma (Taufkirchen, Germany), Thermo Scientific (Munich, Germany), Biozym (Hess. Oldendorf, Germany), Carl Roth (Karlsruhe, Germany), Diagenode (Liege, Belgium), Formedium (Norwich, UK), Fujifilm Corporation (Tokyo, Japan), Gilson (Bad Camberg, Germany), Hartmann Analytic GmbH (Braunschweig, Germany), Jena Bioscience GmbH (Jena, Germany), MembraPure (Bodenheim, Germany), Merck Biosciences (Darmstadt, Germany), NEB (Frankfurt, Germany), Neolab (Heidelberg, Germany), Stratagene (Amsterdam, Netherlands), VWR (Ismaning, Germany), Axon (Kaiserslautern, Germany), Biomol (Hamburg, Germany), Biorad (Munich, Germany), Chemicon (Temecula, Canada), MP Biomedical (Illkirch, France), Santa Cruz (Santa Cruz, USA), PSL (Heidelberg, Germany) and Serva (Heidelberg, Germany).

2.1.2 Equipments

Table 2. List of equipments

Name	Supplier
Avanti JXN-26, JLA-8.1 Rotor	Beckman Coulter (Krefeld, Germany)
Beckman DU650 spectrophotometer	Beckman Coulter (Krefeld, Germany)
Beckman J2-HS, JA-20 Rotor	Beckman Coulter (Krefeld, Germany)
Beckman J6-HC, JS-4.2 Rotor	Beckman Coulter (Krefeld, Germany)
Bioruptor UCD-200 Diagenode	Liege (Belgium)
ChemoCam Imager ECL HR 16-3200	Intas, (Göttingen, Germany)
CME microscope	Leica (Buffalo, USA)
CO8000 Cell Density Meter	WPA (Cambridge, UK)
Dissection microscope MSM 400	Singer (Somerset, UK)
Electrophoresis Power Supply Consort E835	Neolab (Heidelberg, Germany)
Eppendorf centrifuge FA-45-24-11 5424/5424R	Eppendorf (Hamburg, Germany)
Gel iX20	Intas (Göttingen, Germany)
Heating oven	BINDER GmbH (Tuttlingen, Germany)
Heidolph Duomax 1030	Heidolph Instruments GmbH & Co. KG (Schwabach, Germany)
Incubator shaker ISF-1-V	Adolf Kühner AG (Basel, Switzerland)
Innova 44 shaking incubator	New Brunswick Scientific (Nürtingen, Germany)
L80 Ultracentrifuge	Beckman Coulter (Krefeld, Germany)
Megafuge 40R, 75003180 Rotor	Thermo Scientific (Munich, Germany)

Mini-Protean II system, Mini Trans-Blot Cell	Bio-Rad Laboratories (Herucles, USA)
Multitron Pro / Labotron shaker	Infors HT (Einsbach, Germany)
Optima TM L-90 K and L80 ultracentrifuge	Beckman Coulter (Krefeld, Germany)
Optima XPN-80 Ultracentrifuge	Beckman Coulter (Krefeld, Germany)
pH 211 Microprocessor pH meter	HANNA instruments (Woonsocket, USA)
Pipetboy acu	INTEGRA Biosciences AG (Zizers, Switzerland)
Planetary Mono Mill Pulverisette 6	Fritsch (Idar-Oberstein, Germany)
QuantStudio 3 Real-Time PCR System	Applied Biosystems (Darmstadt, Germany)
Research Pipettes P2, P20, P200 P1000	Gilson (Bad Camberg, Germany)
SLC 6000, GS3, SW34 Rotors	Thermo Fisher Scientific (Munich, Germany)
Sonifier 250	Branson (Danbury, USA)
Sorvall Evolution RC, RC 5B Plus	Thermo Fisher Scientific (Munich, Germany)
Spectrophotometer ND-1000	Thermo Fisher Scientific (Munich, Germany)
StepOnePlus™ Real Time PCR System	Applied Biosystems (Darmstadt, Germany)
SW32, SW40, Type 70 Ti Rotors	Beckman Coulter (Krefeld, Germany)
T3 Thermocycler	Biometra (Göttingen, Germany)
Tecan Sunrise - The 96-well Microplate Reader	Tecan Group Ltd. (Männedorf Switzerland)
Thermomixer compact	Eppendorf (Hamburg, Germany)
Typhoon FLA-9500 / 9400	GE Healthcare (Freiburg, Germany)
Universal Analytical Balance	Satorius (Göttingen, Germany)
Vortex Genie	Neolab (Heidelberg, Germany)

2.1.3 Commercially available kits

Table 3. List of commercially available kits

Name	Supplier
ECL Kit	Applichem GmbH (Darmstadt, Germany)
Nucleobond AX PC100	Macherey&Nagel (Düren, Germany)
NucleoSpin® Gel and PCR Clean-up	Macherey&Nagel (Düren, Germany)
PureYield™ Plasmid Miniprep System	Promega (Mannheim, Germany)
West Dura ECL kit	Thermo Fisher Scientific (Munich, Germany)

2.1.4 Growth Media and Buffers

2.1.4.1 Growth Media

Table 4. Media and their composition

Name	Composition
Yeast Full Medium (YPD)	2% (w/v) peptone; 2% (w/v) glucose; 1% (w/v) yeast extract; (2% (w/v) agar added for petri plates)
Yeast Full Medium (YPG)	2% (w/v) peptone; 2% (w/v) galactose; 1% (w/v) yeast extract; (2% (w/v) agar added for petri plates)
LB (Luria-Bertani Broth)	1% (w/v) tryptone; 0.5% (w/v) yeast extract; 0.5% (w/v) NaCl; (2% (w/v) agar added for petri plates)
Synthetic complete dropout medium (SDC)	0.67% (w/v) yeast nitrogen base; 0.06% (w/v) complete synthetic mix of amino acids; drop out as required; 2% (w/v) glucose; when required 0.1% (w/v) 5-FOA was added; (2% (w/v) agar added for petri plates)
Sporulation medium (YPA)	2% (w/v) peptone;1%(w/v) yeast extract; 1% (w/v) potassium acetate; (2% (w/v) agar added for petri plates)

2.1.4.2 Buffers and Other Solutions

All the standard buffers and their compositions are listed below. Buffers related to some specific experiments are listed along with their protocols (in 'Methods' section).

Table 5. Buffers and their compositions

Name	Composition
100X Protease inhibitors cocktail	8 ng/ml Leupeptin; 137 ng/ml Pepstatin A; 17 ng/ml PMSF; 0.33 mg/ml Benzamidine; dissolved in 100% EtOH (p.a.)
10X KNOP buffer	500 mM Tris-HCl (pH 9.2); 160 mM $(NH_4)_2 SO_4$; 22.5 mM $MgCl_2$
10X Phosphate-buffered saline (PBS) buffer	1.37M NaCl; 27mM KCl; 20mM $KH_2 PO_4$; 10mM $Na_2 HPO_4 \cdot 2 H_2O$
10X TBE electrophoresis buffer	1 M Tris base; 1 M boric acid; 0.02 M EDTA (disodium salt)
10X TE buffer	100 mM Tris-HCl; 10 mM EDTA, pH 7.5
10X Tris-buffered saline (TBS) buffer	1.37M NaCl; 27mM KCl; 125mM Tris-HCl, pH 7.4
1X High Salt FA Lysis Buffer (for ChIP)	50mM HEPES-KOH pH to 7.5; 0.5M NaCl; 1mM EDTA; 1% Triton X-100; 0.1% sodium deoxycholate
1X Low Salt FA Lysis Buffer (for ChIP)	50mM HEPES-KOH pH to 7.5; 150Mm NaCl; 1mM EDTA; 1% Triton X-100; 0.1% sodium deoxycholate; 0.1% SDS
1X SDS-PAGE Running Buffer	25 mM Tris, 0.1% (w/v) SDS, 0.19 mM glycine
1X TAP Buffer	50 mM Tris-HCl pH 7.5; 100 mM NaCl; 1.5 mM $MgCl_2$; 0.15% NP40
1x TLEND Buffer (for ChIP)	10mM Tris-HCl pH 8.0; 0.25M LiCl; 1mM EDTA; 0.5% Nonidet P-40; 0.5% SDS
1X Wet blotting buffer	25mM Tris; 192mM glycine; 10% methanol

20x SSC (pH 7.0)	3 M NaCl; 0.3 M sodium citrate
20X SSC buffer	300 mM sodium citrate (pH 7), 3M NaCl
2X Formamide Gel-Loading Dye	90% (v/v) formamide; 1x TBE buffer; 0.05% bromphenol blue
4X SDS sample loading buffer	0.2 M Tris pH 6.8 at 25°C; 40% (v/v) glycerol; 8% (w/v) SDS; 0.2% (w/v) bromophenol blue; 0.1M DTT
4X separating SDS-gel buffer	3 M Tris, 0.4% (w/v) SDS, pH 8.8 (HCl)
4X stacking SDS-gel buffer	0.5 M Tris, 0.4% (w/v) SDS, pH 6.8 (HCl)
50X TAE buffer	2 M Tris; 1 M acetic acid; 100 mM EDTA, pH 8.0
5X Bradford reagent solution	0.05% (w/v) Coomassie Brilliant Blue G-250, 25% ethanol, 42.5% phosphoric acid
6X DNA loading dye	40% (w/v) sucrose; 0.25% bromphenol blue; 0.25% xylene cyanole FF
AE buffer	50mM NaAc; 10mM EDTA; 10mM Tris/HCl pH 7.4
Coomassie destaining solution	30% (v/v) ethanol; 10% (v/v) acetic acid
Coomassie staining solution	0.25% (w/v) Coomassie Brilliant Blue R-250; 30% (v/v) ethanol; 10% (v/v) acetic acid
Elution buffer (for ChIP)	50 mM Tris/HCl pH 7.5; 10 mM EDTA
Peptide-Assay High Salt (PA-HS) buffer	1 M NaCl; 25 mM Tris/HCl pH 8.0; 5% Glycerol; 2.5 mM DTT; 0.025% NP-40; 1x PI; 1xPhosI
Peptide-Assay Low Salt (PA-LS) buffer	100 mM NaCl; 25 mM Tris/HCl pH 8.0; 5% Glycerol; 2.5 mM DTT; 0.05% NP-40; 1x PI; 1xPhosI
Phosphatase Inhibitors (PhosI) 100x	10 mM NaF; 1 mM Na_3VO_4; 5 mM ß-glycerophosphate
Ponceau staining solution	0.1% (w/v) PonceauS; 5% acetic acid
Prehybridisation buffer	50% formamide; 10% dextran sulphate; 125 µg/ml of *E.coli* tRNA; 500 µg/ml H.S. DNA

Spheroblasting buffer	0.1 M K-phosphate; 1.2 M sorbitol
Yeast Transformation Sol I	0.5 ml 10x TE; 0.5 ml 10x LiAc; 4 ml H_2O
Yeast Transformation Sol II	0.5 ml 10x TE; 0.5 ml 10x LiAc; 4 ml 50% PEG

2.1.5 Organisms

2.1.5.1 *Escherichia coli* strains

Table 6. *Escherichia coli* strains

Strain	Genotype	Reference
Escherichia coli DH5α	F^- endA1 glnV44 thi-1 recA1 relA1 gyrA96 deoR nupG Φ80dlacZΔM15 Δ(lacZYAargF) U169 hsdR17(rK^- mK^+) λ–	Woodcock et al. (1989)
Escherichia coli Rosetta™ (DE3)pLysS	F^- *ompT hsdS*$_B$(r_B^- m_B^-) *gal dcm* (DE3) pLysSRARE2 (Cam^R)	Novagen®

2.1.5.2 Yeast strains

Table 7. Yeast strains

Strain name	Strain Background	No.	Genotype	Reference
	RS453	Y1	*MAT a; ade2-1; his3-11,15; ura3-52; leu2-3,112; trp1-1; can1-100; GAL+*	(Strasser and Hurt 2000)
	W303	Y8	*MAT a; ade2-1; his3-11, 15; ura3-1; leu2-3, 112; trp1-1; can1-100; rad5-535*	(Thomas and Rothstein 1989)
	GHY498	Y202	*Mat a; rpb1Δ187::HIS3; his3Δ200; his4-912δ; lys28δ; leu2Δ1; ura3-52*	(Lindstrom and Hartzog 2001)
	BY4741	Y2657	*MAT a; his3Δ1; leu2Δ0; met15Δ0; ura3Δ0*	Euroscarf

MUD2 shuffle	RS453	Y2868	*MAT a; mud2::kanMX4; ade2-1; his3-11,15; ura3-52; leu2-3,112; trp1-1; can1-100; GAL+, pRS316-MUD2*	this study
MUD2-TAP	RS453	Y2869	*MAT a; MUD2-TAP::TRP1; ade2-1; his3-11,15; ura3-52; leu2-3,112; trp1-1; can1-100; GAL+*	this study
RPB1 shuffle	Cross of GHY498 and RS453	Y298	*MAT a; rpb1Δ187::HIS3; his3; ura3-52; leu2; trp1-1; YCp50-RPB1*	this study
RPB1 shuffle MUD2-TAP	Cross of GHY498 and RS453	Y2870	*MAT a; rpb1Δ187::HIS3; MUD2-TAP::TRP1; his3; ura3-52; leu2; trp1-1; YCp50-RPB1*	this study
RPB1 shuffle SYF1-TAP	Cross of GHY498 and RS453	Y2871	*MAT a; rpb1Δ187::HIS3; SYF1-TAP::TRP1; his3; ura3-52; leu2;trp1-1; YCp50-RPB1*	this study
MUD2-TAP PRP19-HA	RS453	Y2872	*MAT a; MUD2-TAP::TRP1; PRP19-HA::HIS3MS6; ade2-1; his3-11,15; ura3-52; leu2-3,112; trp1-1; can1-100; GAL+*	this study
PRP19-HA	RS453	Y2873	*MAT a; PRP19-HA::HIS3MX6; ade2-1; his3-11,15; ura3-52; leu2-3,112; trp1-1; can1-100; GAL+*	this study
Δmud2	RS453	Y2874	*MAT a; mud2::kanMX4; ade2-1; his3-11,15; ura3-52; leu2-3,112; trp1-1; can1-100; GAL+*	this study
HPR1-TAP	RS453	Y46	*MATα; HPR1-TAP::TRP1-KL; ade2-1; his3-11,15; ura3-52; leu2-3,112; trp1-1; can1-100; GAL+*	(Strasser, Masuda et al. 2002)
Δmud2 HPR1-TAP	RS453	Y2875	*MAT a; mud2::kanMX4; HPR1-TAP::TRP1; ade2-1; his3-11,15; ura3-52; leu2-3,112; trp1-1; can1-100; GAL+*	this study
SYF1-TAP	RS453	Y641	*MAT a; SYF1-TAP::TRP1-KL; ade2-1; his3-11,15; ura3-52; leu2-3,112; trp1-1; can1-100; GAL+*	(Chanarat, Seizl et al. 2011)
Δmud2 SYF1-TAP	RS453	Y2876	*MAT a; mud2::kanMX4; SYF1-TAP::TRP1; ade2-1; his3-11,15; ura3-52; leu2-3,112; trp1-1; can1-100; GAL+*	this study

PCF11-TAP	RS453	Y713	*MAT a; PCF11-TAP::TRP1; ade2-1; his3-11,15; ura3-52; leu2-3,112; trp1-1; can1-100; GAL+*	(Meinel, Burkert-Kautzsch et al. 2013)
RIX1-TAP	RS453	Y616	*MAT α; RIX1-TAP::TRP1; ade2-1; his3-11,15; ura3-52; leu2-3,112; trp1-1; can1-100; GAL+*	(Meinel, Burkert-Kautzsch et al. 2013)
Δdst1	RS453	Y3033	*MAT a; dst1:: HIS3; ade2-1; his3-11,15; ura3-52; leu2-3,112; trp1-1; can1-100; GAL+*	this study
Δmud2 Δdst1	RS453	Y3034	*MAT a; mud2::kanMX4; dst1:: HIS3; ade2-1; his3-11,15; ura3-52; leu2-3,112; trp1-1; can1-100; GAL+*	this study
MUD2-FTpA	RS453	Y3025	*MAT a; MUD2-FLAG-TEV-protA::HIS3; ade2-1, his3-11,15, ura3-52, leu2-3,112, trp1-1, can1-100, GAL+*	this study
PAF1-TAP	RS453	Y2751	*MAT a; PAF1-TAP::TRP1; ade2-1; his3-11,15; ura3-52; leu2-3,112; trp1-1; can1-100; GAL+*	this study
Δmud2 PAF1-TAP	RS453	Y3026	*MAT a; mud2::kanMX4; PAF1-TAP::TRP1; ade2-1; his3-11,15; ura3-52; leu2-3,112; trp1-1; can1-100; GAL+*	this study
SPT5 - TAP	RS453	Y533	*MAT a; SPT5-TAP::TRP1; ade2-1; his3-11,15; ura3-52; leu2-3,112; trp1-1; can1-100; GAL+*	this study
Δmud2 SPT5-TAP	RS453	Y2997	*MAT a; mud2::kanMX4; SPT5-TAP::TRP1; ade2-1; his3-11,15; ura3-52; leu2-3,112; trp1-1; can1-100; GAL+*	this study
Δmud2 SYF1 shuffle	RS453	Y2980	*MAT a; mud2::kanMX4; syf1:: HIS3; ade2-1; his3-11,15; ura3-52; leu2-3,112; trp1-1; can1-100; GAL+, pRS316-SYF1*	this study
SYF1 shuffle	W303	Y2392	*MAT a; syf1::kanMX4; ade2-1; his3-11, 15; ura3-1; leu2-3, 112; trp1-1; can1-100; rad5-535, pRS316-SYF1*	this study
MUD2-TAP SYF1 shuffle	W303	Y2993	*MAT a; syf1::kanMX4; MUD2-TAP::TRP1; ade2-1; his3-11, 15;*	this study

			ura3-1; leu2-3, 112; trp1-1; can1-100; rad5-535, pRS316-SYF1	
MUD2-TAP	W303	Y2990	*MAT a; MUD2-TAP::TRP1; ade2-1; his3-11, 15; ura3-1; leu2-3, 112; trp1-1; can1-100; rad5-535*	this study
MUD2-TAP Δhpr1	W303	Y2998	*MAT a; hpr1::HIS3; MUD2-TAP::TRP1; ade2-1; his3-11, 15; ura3-1; leu2-3, 112; trp1-1; can1-100; rad5-535*	this study
Δmud2	BY4741	Y2906	*MAT a; mud2::kanMX4; ura3Δ0; leu2Δ0; his3Δ1; met15Δ0*	Euroscarf
Δsnu66	BY4741	Y2907	*MAT a; snu66::kanMX4; ura3Δ0; leu2Δ0; his3Δ1; met15Δ0*	Euroscarf
Δlea1	BY4741	Y2908	*MAT a; lea1::kanMX4; ura3Δ0; leu2Δ0; his3Δ1; met15Δ0*	Euroscarf
Δlin1	BY4741	Y2909	*MAT a; lin1::kanMX4; ura3Δ0; leu2Δ0; his3Δ1; met15Δ0*	Euroscarf
Δnam8	BY4741	Y2910	*MAT a; nam8::kanMX4; ura3Δ0; leu2Δ0; his3Δ1; met15Δ0*	Euroscarf
Δmud1	BY4741	Y2911	*MAT a; mud1::kanMX4; ura3Δ0; leu2Δ0; his3Δ1; met15Δ0*	Euroscarf
Δ*cwc15*	BY4741	Y3048	*MAT a; YDR163w::kanMX4; ura3Δ0; leu2Δ0; his3Δ1; met15Δ0*	Euroscarf
Δ*isy1*	BY4741	Y3049	*MAT a; YJR050w::kanMX4; ura3Δ0; leu2Δ0; his3Δ1; met15Δ0*	Euroscarf
Δ*ntc20*	BY4741	Y3050	*MAT α; YBR188c::kanMX4; ura3Δ0; leu2Δ0; his3Δ1; met15Δ0*	Euroscarf
Δ*syf2*	BY4741	Y3051	*MAT α; YGR129w::kanMX4; ura3Δ0; leu2Δ0; his3Δ1; met15Δ0*	Euroscarf
Δ*snt309*	BY4741	Y3052	*MAT α; YPR101w::kanMX4; ura3Δ0; leu2Δ0; his3Δ1; met15Δ0*	Euroscarf
HPR1-TAP	BY4741	Y2659	*MAT a; HPR1-TAP::URA3; his3Δ1; leu2Δ0; met15Δ0; ura3Δ0*	(Meinel, Burkert-Kautzsch et al. 2013)
Δ*cwc15 HPR1-TAP*	BY4741	Y3053	*MAT a; YDR163w::kanMX4; HPR1-TAP::HIS3; ura3Δ0; leu2Δ0; his3Δ1; met15Δ0*	this study
Δ*isy1 HPR1-TAP*	BY4741	Y3054	*MAT a; YJR050w::kanMX4; HPR1-TAP::HIS3; ura3Δ0; leu2Δ0; his3Δ1; met15Δ0*	this study

Δ*ntc20 Hpr1-TAP*	BY4741	Y3055	*MAT α; YBR188c::kanMX4; HPR1-TAP::HIS3; ura3Δ0; leu2Δ0; his3Δ1; met15Δ0*	this study
Δ*syf2 Hpr1-TAP*	BY4741	Y3056	*MAT α; YGR129w::kanMX4; HPR1-TAP::LEU2; ura3Δ0; leu2Δ0; his3Δ1; met15Δ0*	this study
Syf1-TAP	BY4741	Y2503	*MAT a; SYF1-TAP::HIS3; his3Δ1; leu2Δ0; met15Δ0; ura3Δ0;*	Chanarat, Sittinan (unpublished)
Δ*cwc15 Syf1-TAP*	BY4741	Y3057	*MAT a; YDR163w::kanMX4; SYF1-TAP::HIS3; ura3Δ0; leu2Δ0; his3Δ1; met15Δ0*	this study
Δ*isy1 Syf1-TAP*	BY4741	Y3058	*MAT a; YJR050w::kanMX4; SYF1-TAP::LEU2; ura3Δ0; leu2Δ0; his3Δ1; met15Δ0*	this study
Δ*ntc20 Syf1-TAP*	BY4741	Y3059	*MAT α; YBR188c::kanMX4; SYF1-TAP::HIS3; ura3Δ0; leu2Δ0; his3Δ1; met15Δ0*	this study
Δ*syf2 Syf1-TAP*	BY4741	Y3060	*MAT α; YGR129w::kanMX4; SYF1-TAP::LEU2; ura3Δ0; leu2Δ0; his3Δ1; met15Δ0*	this study
Δ*snt309 Syf1-TAP*	BY4741	Y3061	*MAT α; YPR101w::kanMX4; SYF1-TAP::LEU2; ura3Δ0; leu2Δ0; his3Δ1; met15Δ0*	this study
N-TAP Hpr1	BY4741	Y2984	*MAT a; NTAP-HPR1; his3Δ1; leu2Δ0; met15Δ0; ura3Δ0*	this study
Δ*cwc15 N-TAP Hpr1*	BY4741	Y2986	*MAT a; YDR163w::kanMX4; NTAP-HPR1; ura3Δ0; leu2Δ0; his3Δ1; met15Δ0*	this study
Δ*syf2 N-TAP Hpr1*	BY4741	Y2988	*MAT a; YGR129w::kanMX4; NTAP-HPR1; ura3Δ0; leu2Δ0; his3Δ1; met15Δ0*	this study

2.1.6 Oligonucleotides

2.1.6.1 Oligonucleotides for genomic tagging

Table 8. Oligonucleotides for genomic tagging

Name	Sequence (5′–3′)
Hpr1-TAP fow	ATGCAGCTACTTCGAACATTTCTAATGGTTCATCTACCCAAGATA TGAAAtccatggaaaagagaag
Hpr1-TAP rev	TAAAATCTATCTGAATTGTTTGGGACACTATGCATGAATTTCTTAT CAGTtacgactcactataggg
Mud2-HA fow	CTTATATAGATGAGGACGACTTTGACATGATGGAAGCAACCCAA CTTTCCcgtacgctgcaggtcgac
Mud2-HA rev	ATGAATACTCAATTCTTTACTTAATTTCGCTCTACAAAATAGACC ATTTAatcgatgaattcgagctcg
Mud2-TAP fow	CTTATATAGATGAGGACGACTTTGACATGATGGAAGCAACCCAA CTTTCCtccatggaaaagagaag
Mud2-TAP rev	ATAATGAATACTCAATTCTTTACTTAATTTCGCTCTACAAAATAG ACCATtacgactcactataggg
NTAP-Hpr1-fow	GATGATTTTAACAATTCAAGAGGCATTAAAACTTGGGCAAAGGA GTAATAgaacaaaagctggagctcat
NTAP-Hpr1-rev	TTCTGGAGAAAACCTATTGAGTTTTGGATCAATTCCTCGGTATTA GACATcttatcgtcatcatcaagtg
Prp19-HA fow	ATTCTGAAGACAAATGATAGTTTCAATATTGTTGCATTGACACCCc gtacgctgcaggtcgac
Prp19-HA rev	ATTACACAGGTTTATTTAGAAAGTACAAACGTGTCAGCGTATTTA atcgatgaattcgagctcg
Syf1-HA fow	CAAAAAAGAACTTATGGTTTCGAAAATGATGCATGATTTTACATA GCTTATATCAatcgatgaattcgagctcg
Syf1-HA rev	ACCCAATCAACCTCTTCATATTCGATTAATCCAGATGAAATAGAA CTAGATATTcgtacgctgcaggtcgac

Syf1-TAP fow	CAATCAACCTCTTCATATTCGATTAATCCAGATGAAATAGAACTA GATATTtccatggaaaagagaag
Syf1-TAP rev	CAAAAAAGAACTTATGGTTTCGAAAATGATGCATGATTTTACATA GCTTATAtacgactcactataggg

2.1.6.2 Oligonucleotide Sequences for qPCR

Table 9. Oligonucleotides for qPCR

Name	Sequence (5′–3′)
YER_fow	TGCGTACAAAAAGTGTCAAGAGATT
YER_rev	ATGCGCAAGAAGGTGCCTAT
PMA1-5'-fow	GTTTTTCGTCGGTCCAATTCA
PMA1-5'-rev	AACCGGCAGCCAAAATAGC
PMA1-M-fow	AAATCTTGGGTGTTATGCCATGT
PMA1-M-rev	CCAAGTGTCTAGCTTCGCTAACAG
PMA1-3'-fow	CAGAGCTGCTGGTCCATTCTG
PMA1-3'-rev	GAAGACGGCACCAGCCAAT
DBP2-5'-fow	CCAAAGCCAATCACCACTTTC
DBP2-5'-rev	CAGCCTTCACTTCATTCAAAACG
DBP2-M-fow	CGTGACTGGGTTCTACAAGAGTTTAG
DBP2-M-rev	GGCCACATCAGTAGCAACCAT
DBP2-3'-fow	CTTCACCGAACAAAACAAAGGTT
DBP2-3'-rev	TCGGGAGGAATATTTTGATTAGCT
DBP2-intron-fow	ACGCATACATACGCTTCGTTG
DBP2-intron-rev	AATCTACCCTTGTACAAATGCCA

ACT1-5’-fow	TGGTATGTTCTAGCGCTTGCAC
ACT1-5’-rev	ATCTCTCGAGCAATTGGGACC
ACT1-M-fow	GTATTGTCACCAACTGGGACG
ACT1-M-rev	TCTGGGGCAACTCTCAATTCG
ACT1-3’-fow	TCAGAGCCCCAGAAGCTTTG
ACT1-3’-rev	TTGGTCAATACCGGCAGATTC
ILV5-5’-fow	AAGAGAACCTTTGCTTTGGC
ILV5-5’-rev	TTGGCTTAACGAAACGGGCA
ILV5-M-fow	TGCCGCTCAATCAGAAACCT
ILV5-M-rev	GGGAGAAACCGTGGGAGAAG
ILV5-3’-fow	TGGTACCCAATCTTCAAGAATGC
ILV5-3’-rev	ACCGTTCTTGGTAGATTCGTACA
CCW12-5’-fow	ACTGTCGCTTCTATCGCCGC
CCW12-5’-rev	TTGGCTGACAGTAGCAGTGG
CCW12-M-fow	CTGTCTCCCCAGCTTTGGTT
CCW12-M-rev	GGCACCAGGTGGTGTATTGA
CCW12-3’-fow	TGAAGCTCCAAAGAACACCACC
CCW12-3’-rev	AGCAGCAGCACCAGTGTAAG
RPL9B-5’-fow	TCCCAGAAGGTGTTACTGTCAG
RPL9B-5’-rev	TCAAAGTACCTCTTGGACCGAC
RPL9B-M-fow	ACATTGTTGAAAAGGATGGTGC
RPL9B-M-rev	CGTTTCTGATCTTCTTGTCACC
RPL9B-3’-fow	AGGACGAAATCGTCTTATCTGGT
RPL9B-3’-rev	CAGATTTGTTGCAAGTCAGCGG

RPL28-5’-fow	ACTAGAAAGCACAGAGGTCACG
RPL28-5’-rev	ACCGTCTTGGTTCTTTCATTCC
RPL28-M-fow	TGGTTGTAGAGAGCGCAATTATG
RPL28-M-rev	GGTTTTCAACTGGACATTTTATCG
RPL28-3’-fow	TGGAAGCCAGTCTTGAACTTGG
RPL28-3’-rev	TTGGTCTCTCTTGTCTTCTGGGA
RPS14B-5’-fow	ACGTGAAAGGGGTGATATCCTG
RPS14B-5’-rev	ACCCGATCACAGTCTCCATC
RPS14B-M-fow	GCAGAAGTTCTGTTTACTAACAAC
RPS14B-M-rev	TGAGAGTTATCGCGAGCTTG
RPS14B-3’-fow	AAGACCCCAGGACCAGGTG
RPS14B-3’-rev	GATACGGCCAATCCTCAAACCAG
Gal1-5’-fow	GTGACTTCTCGGTTTTACCTTTAGCT
Gal1-5’-rev	AATGGATGGATTTTTCTCGTTCA
Gal1-M-fow	ATCATGCTCTATACGTTGAGTTC
Gal1-M-rev	TGCGGAAATTTAAACGGAGTAGC
Gal1-3’-fow	GGTGGTTGTACTGTTCACTTGGTT
Gal1-3’-rev	TCATTGGCAAGGGCTTCTTT

2.1.6.3 Oligonucleotide Sequences for Cloning and Colony PCR

Table 10. Oligonucleotide Sequences for Cloning and Colony PCR

Name	Sequence (5′–3′)
NotI-Mud2-prom-fwd	GGGGCGGCCGCTGTTGCTAAACTGCCAACAGCGG

XhoI-Mud2-term-rev	GGGCTCGAGTTCTTTGGCGCGCTATACTAG
Mud2-colony-pcr-fwd	TGTTGCTAAACTGCCAACAGCGG
Mud2-colony-pcr-rev	CGACGTAATCACTTCGGTTGGATATCTG
MUD2_seq_middle	CAGAGGACCATCGTCAATAATTAC
TAP-RT2	CCACGGCTTCATCGTGTTG
Cus2-up	CAAGTGATTCAAACTCCCACACC
Cus2-down	CAGTATTGTCTGAGCTTCTCG
syf1-mid	TATTCGAGCACGTGAAATCC
Syf1-col-fow	TTGTGCTAATCAGCGGTTG
Syf1-col-rev	TGCTTTATGGTCTATGAAGTACC
His3-fow	AGCAGAATTACCCTCCACG
His3-rev	ATCGCAATCTGAATCTTGG
Syf1-del-fow	ATACAAAAGGAAATAGATAGAAAAAAAGGAAATAATC AGGCCTCCATATTGGACTCAACTATCCTTCAGTACCGCA AAACTTATTGTGTCCATATATCCTcagggaagtcataacacagtcc
Syf1-del-rev	TGTTTGAAGGAAAATTGCTACTTAATCAGTAATTGCAGA AGGCCGCACCAAAAAAGAACTTATGGTTTCGAAAATGA TGCATGATTTTACATAGCTTATAtcgttcagaatgacacgtatag
Syf1-HA-rev	GCAGCGTACGAATATCTAGT
pym17-rev	GTCAAGGAGGGTATTCTGG
Hpr1-col-pcr-rev1	TGCTGATCTGTTCCATTCG
Isy1-rev	TCCTAAGACTCGGGATGATATAG
Syf2-rev	ATCATACTGAGGAAGACGTC
Ntc20-rev	TTGCCATGGCTACTGACAG
Cwc15-rev	ATTCCCGAAATTAAAGCTGC

Snt309-rev	ACTTCTCTGCTACCGGTAAC
Kanmx4-fwd1	ATTGTTGATGCGCTGGCAG
Kanmx4-fwd2	TCCTATGGAACTGCCTCGGTG
dst1-pro-fwd	TTCACACCGCTCTGCCTATTC
dst1-ter-rev	TCTGACCGAGCCATATTTAATGC
dst1-col-fwd	TGCTGGAGTAACCGCATTTC
dst1-col-rev	TCGATCATGATATTGCTAACC
mud2-pro-fwd	TCAATCCCATTCGGATCATCC
mud2-ter-rev	ACAGGTACAGTCCTTTAACGATG
mud2-col-fwd	TCTATTGGCAATCTCTTGGAC
mud2-col-rev	ACTTTAACCCACGGGTTTAGC
Mud2-prom.FOR	ATACCGTCGACTGTTGCTAAACTGCCAACAGCG
Mud2-prom.REV	CTCTTCTCGAGCTCCGTGTGCTTGTTTTCTGTG
pRS315-TAP-ADH1-Mud2-pro.FOR	AGCACACGGAGCTCGAGAAGAGAAGATGGAAAAAGAATTTCATAGC
pRS315-TAP-ADH1-Mud2-pro.REV	GTTTAGCAACAGTCGACGGTATCGATAAGCTTGATAT
Mud2-ORF-N1-N2-N3.FOR	CGGAGCTCGAGATGGCTGATGAAAAGAGACTGGAG
Mud2-ORF-C4-C5-C6.REV	CATCTTCTCTTGGAAAGTTGGGTTGCTTCCATCAT
pRS315-Mud2-prom-ORF-T-ADH1-C4-C5-C6.FOR	CCCAACTTTCCAAGAGAAGATGGAAAAAGAATTTCATAGCCG
pRS315-Mud2-prom-ORF-T-ADH1-N1-N2-N3.REV	TCATCAGCCATCTCGAGCTCCGTGTGCT

Mud2-fragment-N1.REV	CATCTTCTCTTGCTTTGAAACGTAATCCATGAAGTAACTTGT
pRS315-Mud2-fragment-N1.FOR	CGTTTCAAAGCAAGAGAAGATGGAAAAAGAATTTCATAGCCG
Mud2-N2.FOR	GGAGCTCGAGATGGCTGATGAAAAGAGACTGGAG
Mud2-N2.REV	ATCTTCTCTTCCTGCAAAAATCGACCAGATGATCA
pRS315-Mud2-N2.FOR	TTTTTGCAGGAAGAGAAGATGGAAAAAGAATTTCATAGCCG
pRS315-Mud2-N2.REV	CATCAGCCATCTCGAGCTCCGTGTGCT
Mud2-N3.REV	CATCTTCTCTTTGAGGCCATGTTGTAGTAAGCCG
pRS315-Mud2-N3.FOR	ACATGGCCTCAAAGAGAAGATGGAAAAAGAATTTCATAGCCG
Mud2-C4.FOR	AGCTCGAGATGCTGCCAAATCTGGTTACCCAATCTG
pRS315-Mud2-C4.REV	AGATTTGGCAGCATCTCGAGCTCCGTGTG
Mud2-C5.FOR	AGCTCGAGATGGGTACGGTAATTGCATTGGAAAATT
pRS315-Mud2-C5.REV	ATTACCGTACCCATCTCGAGCTCCGTGTGC
Mud2-C6.FOR	AGCTCGAGATGTCCAAGGCAAATTCAAGGCTTGTAATTTCG
pRS315-Mud2-C6.REV	TTTGCCTTGGACATCTCGAGCTCCGTGTG

2.1.6.4 Oligonucleotides for transcription assays

Table 11. Oligonucleotides for transcription assays

Name	Sequence (5′–3′)
ACT1-cy5	Cy5- TGGATTGAGCTTCATCACCAACG
GAL10-cy5_96	Cy5- ATTAGCTCTACCACAGTGTGTG
SCR1-cy5	Cy5-TTTACGACGGAGGAAAGACG
Trans_assay_*in_vitro*	Cy5- TTCACCAGTGAGACGGGCAAC
GAL10-5-fwd	AGGTGGTGCTGGATACATTGG
GAL10-5-rev	TCAGCAACAACACAGTCATATCC
GAL10-M-fwd	ATGCATTCTGCAAAGCTTCTGG
GAL10-M-rev	TCAAAACATCACCTGCTCTTCTG
GAL10-3-rev	TACCGGTGATTTCTTGTCTGC
GAL10-3-rev	TGTATCTACCAGGCTCAATTGC
ACT1-5-fwd	TGGTATGTTCTAGCGCTTGCAC
ACT1-5-rev	ATCTCTCGAGCAATTGGGACC
ACT1-M-fwd	GTATTGTCACCAACTGGGACG
ACT1-M-rev	TCTGGGGCAACTCTCAATTCG
SCR1-5-fwd	TCTGGTGGGATGGGATACG
SCR1-5-rev	TGCAATCCGTGTCTAGCC
SCR1-M-fwd	TCCGTCTCTCTGTCTGGTG
SCR1-M-rev	TAACAGCGGTGAAGGTGG
SCR1-3-fwd	TGTGCAAGTTGATCGCTTCG
SCR1-3-rev	AACGGCCACAATGTGCGAG

2.1.7 Plasmids

Table 12. List of plasmids used

Plasmid	Number	Description	Reference
pBS1479	84	for C-terminal TAP-tagging of a protein by genomic integration with the *TRP1* marker	(Puig, Caspary et al. 2001)
pY1WT(14)	432	plasmid coding for Rpb1 with a CTD containing 14 wild-type repeats	(West and Corden 1995)
pY1WT(9)A2(6)	434	plasmid coding for Rpb1 with a CTD containing 9 wild-type and 6 S2A repeats	(West and Corden 1995)
pRS316-GAL10-*ACT1*	1959	*GAL10* promoter sequence in front of *ACT1* ORF in pRS316	(Chanarat, Seizl et al. 2011)
pYM15	683	for C-terminal HA tagging with the *HIS3MX6* marker	Euroscarf
LL279	1960	Native HIS4 promoter ([-428] - [+24] respective to the A in the start codon), used as a template for *in vitro* transcription assay	(Seizl, Lariviere et al. 2011)
pRS314	6	A pBluescript-based yeast centromere vector with *TRP1* marker	(Sikorski and Hieter 1989)
pRS315	7	A pBluescript-based yeast centromere plasmid with *LEU2* marker	(Sikorski and Hieter 1989)
pRS316	8	A pBluescript-based yeast centromere vector with *URA3* marker	(Sikorski and Hieter 1989)
pRS314-*MUD2*	1998	the coding sequence of *MUD2* with around 500 bp downstream and upstream of the ORF was amplified and cloned into pRS314	this study
pRS315-*MUD2*	1999	the coding sequence of MUD2 with around 500 bp downstream and upstream of the ORF was sub-cloned from pRS316-Mud2 into pRS315	this study
pRS316-*MUD2*	1961	the coding sequence of *MUD2* with around 500 bp downstream and upstream of the ORF was amplified and cloned into pRS316	this study
pRS315-*SYF1*	2007	a HincII-SpeI fragment containing *SYF1* ORF was cloned into pRS315	(Chanarat, Seizl et al. 2011)
pRS315-*syf1-37*	2008	Same as pRS315-*SYF1* but with C-terminal deletion mutant *syf1-37*	(Chanarat, Seizl et al. 2011)

pFA6a-FTpA-HisMX6	1962	Plasmid used for C-terminal Flag-TEV-protein A tagging in the genome	(Kornprobst, Turk et al. 2016)
pRS315- TAP-T-ADH1	787	TAP-tag sequence from pBS1479 was amplified and cloned into pRS315-*T-ADH1*	(Rother and Strasser 2007)
pRS315-MUD2-*TAP-T-ADH1*	2000	Full-length *MUD2* (without the stop codon) plus 500bp of *MUD2* promoter was cloned into pRS315-*TAPT-ADH1*	this study
pRS315- mud2-N1-*TAP-T-ADH1*	2001	*mud2* deletion mutant N1 including 500 bp 5' of the *MUD2* ORF was cloned into pRS315-*TAP-T-ADH1*	this study
pRS315- mud2-N2-*TAP-T-ADH1*	2002	*mud2* deletion mutant N2 including 500 bp 5' of the *MUD2* ORF was cloned into pRS315-*TAP-T-ADH1*	this study
pRS315- mud2-N3-*TAP-T-ADH1*	2003	*mud2* deletion mutant N3 including 500 bp 5' of the *MUD2* ORF was cloned into pRS315-*TAP-T-ADH1*	this study
pRS315- mud2-C4-*TAP-T-ADH1*	2004	*mud2* deletion mutant C4 including 500 bp 5' of the *MUD2* ORF was cloned into pRS315-*TAP-T-ADH1*	this study
pRS315- mud2-C5-*TAP-T-ADH1*	2005	*mud2* deletion mutant C5 including 500 bp 5' of the *MUD2* ORF was cloned into pRS315-*TAP-T-ADH1*	this study
pRS315- mud2-C6-*TAP-T-ADH1*	2006	*mud2* deletion mutant C6 including 500 bp 5' of the *MUD2* ORF was cloned into pRS315-*TAP-T-ADH1*	this study
pET28b-Mud2-6xhis	2009	*MUD2* ORF was inserted into pET28b, plasmid was used for the purification of Mud2 from *E.coli.* with a 6xhis tag at the C-terminal	this study
pBS1761-his	2010	plasmid for N-terminal TAP-tagging of a protein by genomic integration with the *HIS3* marker	this study
YDp-H	15	A pUC9 vector containing a yeast gene disruption cassette based on *HIS3* selectable marker	(Berben, Dumont et al. 1991)

2.1.8 Antibodies

2.1.8.1 Primary antibodies

Table 13. List of primary antibodies used

Name	Source	Dilution used for Western Blotting/ ChIP	Supplier
Peroxidase α-Peroxidase complex (PAP)	rabbit, monoclonal	1:5000 for blotting	Sigma
anti-HA	rat, monoclonal	1:1000 for blotting	Roche
anti-PGK1	mouse, monoclonal	1:10,000 for blotting	Molecular probes
8WG16 (α-Rpb1, CTD)	mouse, monoclonal	1:1000 for blotting/ 4 µl per ChIP experiment	Covance
yN-18 (α-Rpb1, N-term)	goat, polyclonal	20µl per ChIP experiment	Santa Cruz Biotechnology
anti-CBP antibody (CAB1001)	rabbit, polyclonal	1:1000 for blotting	Open Biosystems

2.1.8.2 Secondary antibodies

Table 14. List of secondary antibodies used

Name	Source	Dilution used for Western Blotting	Supplier
α-mouse HRPO	goat, monoclonal	1:3000	Bio Rad
α-rabbit HRPO	goat, monoclonal	1:3000	Bio Rad
α-rat- HRPO	goat, monoclonal	1:5000	Sigma

2.1.9 CTD peptides used in pull-down assay

Table 15. Sequences of CTD peptides used in pull down assay

Name	Sequence	Supplier
CTD	Biotin - YSPTSPS YSPTSPS YSPTSPS	PSL (Heidelberg)
S2P	Biotin - YSPTSPS Y**pS**PTSPS Y**pS**PTSPS	PSL (Heidelberg)

2.2 Methods

2.2.1 Standard techniques

2.2.1.1 Cloning

Standard molecular cloning techniques such as restriction digestion, dephosphorylation of DNA, DNA separation using agarose gel electrophoresis, DNA ligation and transformation in *E.coli.* were performed according to (Sambrook and Russell 2001). All commercially available kits were used according to instructions mentioned by the manufacturer. The restriction enzymes used were from Fermentas or New England Biolabs. DNA sequencing was performed by MWG, Eurofins or GATC Biotech.

2.2.1.2 PCR reactions

For the purpose of TAP tagging, KNOP polymerase mix (a mixture of Taq-DNA Polymerase and Vent Polymerase) was used and 300 µl of PCR reaction mixture was purified (Section 2.2.1.3) and transformed into yeast cells. The PCR reaction protocol and the program of amplification cycle used is mentioned in Table 16.

Table 16. Protocol for KNOP PCR Reaction

100 µl PCR reaction mix	PCR program used
0.5 µM forward -primer; 0.5 µM reverse-primer; 0.2 mM of each dNTP; 1x KNOP buffer (50 mM Tris-HCl, pH 9,2; 16mM (NH4)2SO4; 2,25 mM MgCl2); 100-300ng template DNA; 1µl KNOP polymerase (2 U Taq and 0.56 U Vent), ddH2O added till 100 µl final volume	94°C 2 min 94°C 1 min 55°C 30 sec 68°C 1 min/1000bp 35x 68°C 10 min

For cloning purposes, high-fidelity amplification of DNA was performed using Phusion® High-Fidelity PCR Master Mix with HF Buffer (NEB) according to manufacturer's protocol. For confirmation of genomic integration, freshly growing yeast cells were picked up with a yellow tip and transferred to an eppendorf tube containing 30µL of 0.2% SDS. The cells were lysed for 5 min at 95°C and spinned down at 13,000 rpm for 1 min. Colony PCR was performed according to following reaction mixture (Table 17):

Table 17. Protocol for Colony PCR Reaction

1X Reaction mixture (25 µl)	PCR Program Used
1 µM forward-primer; 1 µM reverse-primer; 62.5 µM of each dNTP; 1x KNOP buffer; 0.3µl KNOP polymerase; 1µl lysed cells	95°C 60s 95°C 30s 45°C 45s 72°C x s 25 cycles 72°C 600s

For the colony PCR from *E.coli.*, the same program as mentioned in Table 17 was used except that the individual colonies were directly picked up with a toothpick and dipped into the PCR reaction mixture tube and initial denaturation was carried for 2 min (to lyse the cells within the reaction tube).

2.2.1.3 Phenol-Chloroform extraction of DNA

PCR products for yeast transformation were purified by standard phenol-chloroform extraction procedure. One volume of phenol:chloroform:isoamyl alcohol (25:24:1, pH 7.5-8.0) was added to the DNA solution, which was then vortexed and centrifuged at RT for 5 min at top speed. The upper aqueous phase was transferred to a fresh tube and the above step was repeated again. For total RNA extraction from yeast cells (for *in vivo* transcription assays), the cell pellet was resuspended in 500ul AE buffer and 55ul of 10% SDS was added. An equal volume of aqua-phenol:chloroform:isoamyl alcohol (25:24:1, pH 4.5-5) was then added. The cell mixture was vortexed vigorously and incubated at 65°C for 30 min (vortexed 2-3-times in between). The sample was then centrifuged at top speed for 5 min and the upper phase containing RNA was transferred to a new tube. To increase the amount and quality of RNA, this step was repeated twice, first with aqua-phenol:chloroform:isoamyl alcohol and then with only chloroform. The aqueous phase thus obtained was collected in the fresh tube.

To precipitate the purified DNA or RNA, 1/10 of the volume of 3M sodium acetate (pH 5.2) and 2.5 volumes of 100% ethanol were added to the extracted aqueous phase. After incubating the tubes at -20°C for 20 min (or longer), DNA was precipitated for 20 min at 4°C at top speed and the pellet was washed with 80% ethanol. The pellet was dried, resuspended in 20 µl of 1X TE buffer and used subsequently.

2.2.1.4 SDS PAGE

Sodium dodecylsulfate polyacrylamide gel electrophoresis (SDS-PAGE) was performed according to (Laemmli 1970). Mini-PROTEAN II Electrophoresis Apparatus from Biorad was used for casting the gels. The samples for SDS-PAGE were prepared by adding SDS sample loading buffer to the final concentration of 1X and boiling the samples at 95°C for 3 min. unless mentioned, 10 % of polyacrylamide mini gels were used for most of the purposes. The protein samples were loaded and resolved in 1X SDS-PAGE running buffer at 200 V until the bromophenol blue dye runs out of the separating gel (around 1 h). After running the gel, the apparatus was dismantled and the separated proteins on the gel were either stained with Coomassie Blue or transferred to a membrane for specific detection with Western Blotting. For Coomassie staining, the gel was incubated in the Coomassie staining solution for 1 h at RT with

gentle shaking and afterwards destained in the Coomassie destaining solution till desired bands were obtained. PageRuler™ Unstained Protein Ladder (Fermentas/ Thermo Scientific) or Unstained Protein Ladder, Broad Range (NEB) was used as a marker for gels to be analyzed with Coomassie staining. For gels meant to be used for Western Blotting, PageRuler™ Prestained Protein Ladder (Fermentas/ Thermo Scientific) or Color Prestained Protein Standard, Broad Range (NEB) was used as a marker. The protocol for preparing 10% SDS polyacrylamide (PAA) gel is described in Table 18.

Table 18. Protocol for SDS polyacrylamide (PAA) gel preparation

Component	Separating Gel (10%)	Stacking gel (4.5%)
4x separating SDS-gel buffer	2.5 mL	-
4x stacking SDS-gel buffer	-	2 mL
30% Acrylamide/Bis-acrylamide (29:1)	3.33 mL	1.2 mL
ddH_2O	4.17 mL	4.8 mL
TEMED (100%)	10 µL	50 µL
10% APS	50 µL	50 µL

2.2.1.5 Western Blotting

For specific detection of a protein of interest, Western Blotting technique was used. The protein samples were separated on a SDS-PAA gel as described in the previous section and the resolved proteins on the gel were electrophoretically transferred onto a nitrocellulose membrane (Porablot, Macherey&Nagel) using a standard wet blotting apparatus (Biorad). A sandwich assembly comprising of 2 layers of sponge material, 3 layers of Whatman paper, the gel, the membrane, 3 layers of Whatman paper and 2 layers of sponge material (all of them were presoaked in Wet Blotting buffer) was assembled. During the assembly, care was taken that no bubbles were trapped inside the sandwich. The prepared sandwich cassette was placed in the transfer tank containing an ice block. The transfer was done at 150 V for 1 h and 10 min in pre-cooled 1X Wet Blotting buffer. After transfer, the membrane was stained with Ponceau S solution to confirm the transfer of proteins. Ponceau S stain was rinsed off by washing the

membrane with ddH_2O. The membrane was blocked for 1 h in the blocking buffer - 2% (w/v) milk in 1X PBS buffer - to prevent non-specific signal due to the background binding of antibodies. Afterwards, the blocking buffer was removed and the membrane was incubated in primary antibody solution (prepared in the blocking buffer with recommended dilution) for 2 h at RT with mild shaking. After incubation, the excess antibody was removed by washing the membrane 3 times with the blocking buffer, each time for 10 min. The membrane was then incubated for 1 h with the secondary antibody conjugated to horseradish peroxidase. The membrane was finally washed three times with 1X PBS buffer, each time for 10 min. The signal from the blot was detected with chemiluminescence, using CheLuminate-HRP PicoDetect ECL kit (Applichem) according to manufacturer's instructions. The bands were imaged either by exposing the blot to a light sensitive X-ray film (GE Healthcare) in a dark room followed by film development through Optimax TR developing machine (MS Laborgeräte) or directly imaged using a ChemoCam Imager (Intas).

2.2.1.6 Bradford assay

Bradford assay was performed according to (Bradford 1976) in order to determine the protein concentration in the nucleic extracts for the transcription assay. A standard curve was prepared using Bovine serum albumin (BSA) as a standard. The unknown protein samples were diluted and mixed with 5X Bradford reagent solution to the final concentration of 1X. The absorbance was measured at 595 nm in a microplate reader (Tecan Sunrise).

2.2.2 Yeast specific techniques

2.2.2.1 Culture of *S. cerevisiae*

Yeast strains were cultivated either in full-media (YPD) or synthetic complete (SC) media (composition in Table 4) at 30°C (if not stated otherwise) with shaking. For solid agar plates, 2% (w/v) agar was added to the media. Cell optical densities (OD) were measured at 600 nm using a spectrophotometer (one optical density unit or 1 OD corresponds to approximately 2.5 × 10^7 cells). For 6-azauracil (6AU) sensitivity assay, 50 μg/ml of 6AU was added to the media before pouring the plates.

2.2.2.2 Genomic tagging in *S. cerevisiae*

Genomic tagging was done in *S. cerevisiae* as described previously (Puig, Caspary et al. 2001, Janke, Magiera et al. 2004). For C-terminal Tandem Affinity Purification (TAP) tagging, pRS1479 (with *TRP1* marker), pBS1479LEU2 (with *LEU2* marker) and pBS1479HIS3 (with *HIS3* marker) plasmids were used as templates. For N-terminal TAP tagging, the existing pBS1761 plasmid (with *TRP1* marker) was re-cloned to make pBS1761HIS3 plasmid (with *HIS3* marker). To achieve this, the *TRP1* fragment was cut from the pBS1761 plasmid using the restriction enzymes Bp1I and XbaI. The *HIS3* cassette was isolated from the plasmid YDp-H using BamHI. Both the linearized fragments were blunt ended using Quick Blunting™ Kit (NEB) and ligated together with T4 DNA Ligase (NEB) to create pBS1761HIS3 (with *HIS3* marker). The new plasmid was sequenced for confirmation. For C-terminal HA tagging, pYM15 was used as the template plasmid.

To achieve the genomic integration of C or N terminal TAP tag or HA tag to the desired protein, primers were designed as described in (Puig, Caspary et al. 2001). 3X KNOP PCR reaction (300 µL) was prepared (Section 2.2.1.2) and the PCR product was purified by phenol chloroform extraction (Section 2.2.1.3). The purified DNA was transformed into yeast cells (below) to integrate the tag into the yeast genome by homologous recombination. The correct integration of the tag was confirmed either by Western Blotting (Section 2.2.1.5) or Colony PCR (Section 2.2.1.2).

2.2.2.3 Transformation of yeast cells

For transformation, 50 ml of yeast culture was grown till OD_{600} 0.5-0.8, harvested, washed once with water and once with Sol I. The pellet was resuspended in 250µl of Sol I (Table 5). One transformation reaction mixture contained 300 µl Sol II (Table 5), 50µl of yeast cell suspension, 5 µl of single stranded carrier DNA (salmon or herring testis DNA, 2mg/ml) and 1-5 µg of DNA to be transformed. The transformation mixture was then incubated on a turning wheel for 30 min at RT and then heat shocked at 42°C for 10 min followed by 3 min on ice. Cells were washed with 1 ml of water and spinned down at 3600 rpm for 3 min. Plasmid transformations were plated directly on selective plates while for genomic integration, cells were allowed to

recover in YPD or YPG on a turning wheel (30°C) for 1-4 h prior to plating. The plates were incubated for 2-5 days in the incubator till single yeast colonies appeared.

2.2.2.4 Dot spots

Dot spot assay was used to compare the growth phenotypes of different yeast strains. Briefly, a loop of freshly-growing yeast cells was resuspended in 1 ml of ddH_2O and its 4 ten-fold serial dilutions were made. 5µl of each dilution was spotted on a desired media plate and incubated for 3-5 days at required temperature.

2.2.2.5 Whole cell extracts (WCE)

To confirm the correct integration of a tag by western blotting, the whole cell extracts were prepared. To achieve this, the colonies to be tested were freshly grown on the selective plates, and a loopful of cells was resuspended in 50 µl of 1X SDS sample loading buffer. For lysis, 50 µl of glass beads were added to the cells in the loading buffer. The samples were then boiled for 90 sec and vortexed for 90 sec. This step was repeated 3 to 4 times after which the samples were spinned down for 5 min at top speed. 10 µl of the prepared lysate was loaded on a SDS-PAA gel and further subjected to Western Blotting (Section 2.2.1.5).

To prepare Whole Cell Extracts for Quantitative Western Blotting, the cells were grown till mid-log phase and 5 OD_{600} of each culture was harvested. The pellet was resuspended in 500 µl of cold H_2O and lysed by adding 150 µl of pretreatment solution (1.85 M NaOH and 7.5 % β-mercapthoethanol) followed by 20 min incubation on ice. To precipitate the proteins, ice-cold 100% Trichloroacetic acid (TCA, 1% final concentration) was added, followed by 20 min incubation on ice and then 20 min centrifugation at maximum speed (4°C). The supernatant was discarded, and the pellet was dried and dissolved in 85 µl of 1X SDS sample loading buffer, and neutralized with 20 µl of 1M Tris base. Samples were analyzed by SDS-PAGE and Western blotting.

2.2.2.6 6-Azauracil sensitivity assay

In order to test for sensitivity to the drug 6-azauracil (6AU), the yeast cells were transformed with the plasmid pRS316 (containing a *URA3* marker), so that the uracil in the medium does not interfere with the uptake of the drug. SDC-ura agar plates supplemented either with DMSO (as a control) or 50 μg/ml of 6AU (dissolved in DMSO) were prepared. To test for drug sensitivity, dot spot assays were performed on the plates without or with 6AU. The cells were incubated at 30°C for 2-3 days.

2.2.2.7 Yeast gene deletion

To delete a gene, the ORF of the gene was replaced with an auxotrophic marker or an antibiotic resistance gene via homologous recombination. To achieve this, either the disruption cassette was amplified from a commercially available deletion strain (Euroscarf), or primers were made containing flanking regions (50-70 nucleotides) homologous to the promoter and terminator of the targeted gene and sequence complementary to the both sides of a marker cassette (for example, *HIS3* cassette from plasmid YDp-H). In both the cases, the amplified PCR product was purified and transformed into yeast cells. Cells were plated on appropriate selective media plates. The positive colonies were verified by Colony PCR using one primer binding outside of the amplified region (in 5' or 3' UTR of the targeted gene) and other primer binding within the selection marker.

2.2.2.8 Glycerol stocks

For long term storage of yeast cells, glycerol stocks were prepared. The yeast cells were freshly grown on agar plate for 1-2 days, scrapped off from the plate and resuspended in 1ml of 50% sterile glycerol. The cell suspension was flash frozen in liquid nitrogen and stored at -80°C.

2.2.3 Protein Purifications

2.2.3.1 Tandem affinity purification (TAP)

TAP-tagged proteins were purified from *S. cerevisiae* using tandem affinity purification technique as described previously (Puig, Caspary et al. 2001, Strasser, Masuda et al. 2002). 2 l of yeast cells were inoculated using an overnight saturated preculture of the required strain. The cells were grown overnight at 30°C in YPD medium to an OD_{600} of 3-3.5 and harvested by centrifugation at 5000 rpm for 5 min. The cells were washed once with water and once with 1X TBS buffer, and frozen in liquid N_2.

For lysis, the pellet (usually around 10 ml) was mixed with 2 x volume of glass beads and equal volume of lysis buffer or TAP buffer (Section 2.1.5), with freshly added 1X Protease inhibitors cocktail and 1 mM DTT, and lysed using bead beater (Pulverisette, Fritsch) at 4°C according to the program: 3x (4min, 500rpm, 2min break). The glass beads were washed with 15 ml of TAP buffer and separated with a syringe to obtain 25ml of cell lysate. The crude lysate was cleared first with centrifugation at 4000 rpm for 10 min, 4°C and then with ultra-centrifugation at 35,000 rpm for 1 h, 4°C. The top fatty phase was removed with a vacuum pump and the clear upper phase was incubated for 1 h (or overnight) with 500 µl of pre-equilibrated IgG-coupled Sepharose 6 Fast Flow slurry (GE Healthcare) at 4°C. To test for the RNA dependency, RNase A was added to the lysate during IgG incubation to a final concentration of 100 µg/ml. After IgG binding, the IgG beads slurry was centrifuged down and transferred to a mobicol with a 35µm filter. The beads were then washed by gravity with 10 ml of TAP buffer containing freshly added 0.5 mM DTT. For TAP purification under high salt conditions, buffer containing high salt concentration (250mM to 1000mM of NaCl) was used during the washing step. After washing, 120 µl of TAP buffer (+0.5 mM DTT) was added to the mobicol and the beads were treated with 5 µl of Tobacco-etch-virus protease (TEV protease) (4 mg/ml) for 1 h and 20 min at 16°C. Afterwards, the bound protein of interest was eluted by spinning at 2000 rpm for 2 min.

The protein was further purified with a second affinity step by incubating the TEV eluate with 500 µl of calmodulin-coated beads for 1 h at 4°C in the presence of calcium. Calmodulin Sepharose beads (Agilent Technologies) were prewashed 3 times with 10 ml of TAP buffer containing 1mM DTT and 2mM $CaCl_2$ and two times with 10 ml of TAP buffer containing 1mM DTT and 4mM $CaCl_2$. After incubation, beads were washed in mobicol column with 5 ml

of TAP buffer (+1mM DTT and 2mM $CaCl_2$). The protein of interest was eluted by adding 600 µl of elution buffer (10 mM Tris-HCl pH 8.0, 5 mM EGTA). The elution was carried out at 37°C for 15 min in a thermomixer with shaking so that the beads don't settle down. The eluted protein was precipitated by adding TCA to a final concentration of 10% (v/v) and incubating for 20 min on ice followed by 20 min centrifugation at top speed (4°C). The pellet was dried and resuspended in 50 µl of 1X SDS-sample buffer and 2-3 µl of 1 M TRIS base was added to neutralize the pH. The purified protein was boiled at 95°C for 3 min, centrifuged and subjected to SDS-PAGE.

2.2.3.2 Recombinant Mud2 purification

To purify Mud2 from *E.coli,* Mud2 ORF was cloned into the vector pET28b to introduce a 6 x His tag to the C-terminal of Mud2. Mud2-His_6 was expressed in *E. coli*. Rosetta 2 (DE3) cells in a 500 ml culture, grown at 16°C, 200 rpm till an OD_{600} of 0.6, and induced with the addition of 1 mM (final concentration) of isopropyl-1-thio-β-D-galactopyranoside (IPTG) for 1.5 h. The cells were harvested, washed and resuspended in 10 ml of lysis buffer (50 mM KPi pH 8.0, 200 mM NaCl, 0.1% Triton X-100, 10 mM ß-mercaptoethanol) plus protease inhibitors and were lysed by sonication (5 x 30 sec, 25% output). The lysed cell extract was centrifuged for 15 min at 15,000 rpm (4°C) and the supernatant was incubated for 1 h at 4°C with Ni-NTA agarose beads (Qiagen), pre-equilibrated with the lysis buffer. After incubation, the whole solution was applied to an 'Econo-Column' glass chromatography column (Biorad) equilibrated with the buffer. The column was washed 4 times with 1 ml wash buffer (lysis buffer without Triton X-100; containing 20 mM imidazole). Mud2 fractions were eluted from the column using wash buffer plus 150 mM imidazole. Peak fractions were collected, aliquoted, frozen in the liquid nitrogen and stored at -80 °C.

2.2.4 Chromatin immunoprecipitation (ChIP)

Chromatin immunoprecipitation (ChIP) assays were performed according to (Rother, Burkert et al. 2010) with some modifications. 100 ml of yeast cell culture was grown till the mid-log phase (OD_{600} of 0.7-0.8). The growing cells were crosslinked with 1% formaldehyde with mild shaking for 20 min at RT. The reaction was quenched with 0.25 M glycin for 5 min. The cross-

linked cells were centrifuged and the pellet was washed two times with 1X TBS buffer, frozen in liquid nitrogen and stored at -80°C.

The crosslinked cells were resuspended in 800 µl of low salt FA lysis buffer (for buffer composition, see Table 5) and lysed by vortexing with an equal volume of glass beads for 7 x 3 min and 3 min on ice in between. The lysate was sonicated using Bioruptor UCD-200 (Diagenode) for 3 x 15 min (30 sec. ON/30 sec. OFF) at 'HIGH' power setting with intermittent cooling in between (5 min). The average size of resulting chromatin fragments was between 200-250 bp. The sheared lysate was cleared by two step centrifugation for 5 and 10 min at 14000 rpm (4°C). The concentration of supernatant was measured in Spectrophotometer ND-1000 at 280 nm. 10 µl of lysate was reserved as 'INPUT'. For ChIP of Mud2, Syf1, Hpr1 and Rpb1, 30 A_{280}, 30 $A_{280,}$ 15 A_{280} and 15 A_{280} of lysate was used respectively in a total volume of 1200 µl low salt FA lysis buffer. For immunoprecipitation (IP) of TAP-tagged proteins, the lysate was incubated with IgG-coupled dynabeads tosylactivated M280 (Thermo Scientific) for 2.5 h at 20°C. For RNAPII ChIP, 4 µl of anti-RNA Polymerase II Rpb1 8WG16 monoclonal antibody (Biolegend) was added for 1.5 h at 20°C followed by 1 h incubation with Protein G dynabeads. For RNAPII ChIPs in CTD mutants, 20 µl of N-terminal polyclonal antibody y-80 (Santa Cruz Biotechnology) was added for 1.5 h at 20°C followed by 1 h incubation with Protein A dynabeads. The immunoprecipitated (IP) samples were washed two times with 800 µl of low salt FA lysis buffer, two times with high salt FA lysis buffer, one time with 1X TLEND buffer and one time with 1X TE buffer and afterwards eluted in ChIP elution buffer (Table 5). Both INPUT and eluted IP samples were treated with Proteinase K (P4850; Sigma) at 37°C for 2 h and incubated at 65°C overnight for reversal of crosslinks. INPUT and IP DNA was purified using NucleoSpin® Gel and PCR Clean-up Kit (Macherey-Nagel) and quantified using Real Time PCR cycler (Applied Biosystems). At least three biological replications of each ChIP experiment were performed.

2.2.5 Real Time PCR

To analyze the DNA from ChIP experiments, 1:20 dilutions of 'INPUT' and 'IP' DNA were used as a template for a 10 µl PCR reaction (Table 19). SYBR Green based 2X PCR Master Mix (Applied Biosystems) was used as recommended by the manufacturer and PCR reactions were carried out on MicroAmp Fast Optical 96-Well Reaction Plate (Applied Biosystems).

Either StepOnePlus™ Real Time PCR System or QuantStudio 3 Real-Time PCR System (Applied Biosystems) along with its recommended software was used for the qPCR and data analysis.

Table 19. Protocol for qPCR Reaction

1X qPCR Reaction Mix: 2.5 µL of diluted DNA 0.1 µL (100 pmol/µL) of each primer 5 µL 2x Power Sybr Green PCR Master Mix 2.3 µL of ddH2O = 10 µL total volume	qPCR program: 95°C 10 min (Initial Denaturation) 95°C 15 sec (Denaturation) 60°C 60 sec (Annealing/ Elongation) - for 45 Cycles Melting curve: 95°C 15 sec 60°C 60 sec 60°C - 95°C continuously 0.3 °C 95°C 15 sec

A non-transcribed gene region (NTR1, 174131–174200 on chr. V) served as a negative control for the experiment to account for nonspecific DNA binding of the target protein and H_2O instead of DNA served as a NTC (non-template control) for the PCR. Four highly expressed intron-containing genes *DBP2*, *ACT1*, *RPS14B* and *RPL28* and four highly expressed intronless genes *PMA1*, *ILV5*, *CCW12* and *RPL9B* were analyzed for the occupancy of the protein of interest. The occupancy of the target protein at a particular gene was calculated as its enrichment at that gene relative to its presence at the non-transcribed region (NTR). The schematic diagram of these genes and positions of the primers used is shown in Figure 10. Standard curves were plotted to calculate PCR efficiencies and to ensure that the detected signal lies in the linear range. Melting curves analysis of the amplicons was done to verify the specificity of primers. Three technical replications of each PCR reaction were performed and the average values were considered for further analysis.

All qPCR data are presented as average ± Standard Deviation (SD) of at least three biologically independent experiments. Asterisks indicate the statistical significance (Student's t-test; * $P < 0.05$ and ** $P < 0.01$).

2.2.6 Transcription assays

2.2.6.1 *In vivo* transcription assay

Wild-type and mutant cells were transformed with a plasmid *pGAL10-ACT1* (*URA3* marker). The cells were grown in a media containing raffinose until OD_{600} reached 0.8. The transcription of *ACT1* was induced by adding 2% of galactose to the media. The cells were harvested at the time points 0 (before induction), 15 min, 30 min, 45 min and 60 min after galactose induction. Total RNA was extracted using standard hot phenol extraction procedure. Amount of *ACT1* mRNA (from plasmid) and *GAL10* mRNA (from genome) transcribed from *GAL10* promoter was measured by primer extension assay using 5' Cy5 labeled reverse primers specific for *ACT1* and G*AL10* mRNA. SCR1 RNA, an RNAPIII transcript, served as a loading control for the primer extension assay. 5µg, 10µg and 5µg of total RNA was used for *ACT1*, *GAL10* and SCR1 reverse transcription reactions respectively. The generated cDNA was further precipitated and analyzed on a 7M urea and 7% polyacrylamide gel using a Typhoon 9400 scanner. Quantification was done using ImageQuant software (GE Healthcare).

2.2.6.2 *In vitro* transcription assay

2.2.6.2.1 Yeast nuclear extract preparation:

Yeast nuclear extracts from wild-type and mutant cells were prepared as mentioned previously (Seizl, Lariviere et al. 2011) with some modifications. For the preparation of nuclear extracts, 6 l of yeast cell culture was grown in YPD at 30°C until the OD_{600} reached between 4-5. The cells were harvested and resuspended in 35 ml of resuspension buffer (50mM Tris pH 7.5, 20mM EDTA, 30mM DTT) and incubated at 30°C for 15 min. The cells were pelleted and resuspended in 20 ml of YPD/S (YPD with 1M sorbitol). 3 ml of 2M sorbitol was further added to the cell suspension. To start spheroplasting, 3 ml of resuspension buffer containing 18 mg zymolyase (Seikagaku) and protease inhibitor cocktail (PI) was added followed by an incubation at 30°C (water bath) with gentle mixing. The progress of spheroplasting was checked after 30 min, and if less than 80% spheroplasts were obtained, zymolyase treatment was repeated again. After spheroplasting, 100 ml YPD/S was added and the cells were pelleted. The cells were resuspended in 250 ml YPD/S (RT) and recovered at 30°C for 30 min (in a water

bath). After that, the cells were pelleted again and washed twice with 200 ml of cold YPD/S and once with 250 ml 1 M Sorbitol (4°C). The cells were finally resuspended in 100 ml of ice-cold Buffer A (10mM Tris pH 7.5, 18% polysucrose 400, 20mM potassium acetate, 5mM magnesium acetate, 1mM EDTA, 0.5mM spermidine, 0.15mM spermine, 3mM DTT and protease inhibitors). The cell membrane was disrupted by passing the cell suspension six times through pre-chilled Dounce glass homogenizer. The cell suspension was centrifuged 4 times at 5000 rpm for 8 min, each time transferring the supernatant to a new bottle. Crude nuclei were isolated by centrifuging the resulting supernatant at 13000 rpm for 30 min at 4°C. The crude nuclear pellet was washed with 15 ml of buffer B (100mM Tris pH 8.0, 50mM potassium acetate, 10mM magnesium sulfate, 20% glycerol, 2mM EDTA, 3mM DTT and protease inhibitor cocktail) and resuspended again in 15 ml fresh buffer B. To lyse the nuclei, ammonium sulfate (pH 7.5) was added to a final concentration of 0.5M for 30 min at 4°C and the nuclear lysate was ultra-centrifuged at 28000 rpm for 90 min at 4°C. The nuclear proteins were further precipitated by adding excess amount of ammonium sulfate and centrifuging. The pelleted nuclear extracts were resuspended in Buffer C (20mM HEPES pH 7.6, 10mM magnesium sulfate, 1mM EGTA, 20% glycerol, 3mM DTT and protease inhibitor cocktail). The nuclear extracts were further dialyzed against buffer C containing 75mM ammonium sulfate, aliquoted and frozen in liquid nitrogen. The protein concentration of nuclear extracts was measured with Bradford assay and equal concentration was used from the wild-type and mutant cells for the in vitro transcription assay.

2.2.6.2.2 *In vitro* transcription assay:

A plasmid-based in vitro transcription assay was performed with a *S. cerevisiae HIS4* promoter as described previously (Seizl et al. 2011) with few changes. The transcription reaction was performed in a 50µl reaction volume and at two temperatures, 18°C and 23°C, for 1 h. The reaction mixture contained 200 µg yeast nuclear extract, 200 ng of template plasmid, 0.1 M DTT, 64 mg/ml of phosphocreatine, transcription buffer (10mM HEPES pH 7.6, 50mM potassium acetate, 0.5mM EDTA, 2.5mM magnesium acetate), 0.2 µg of creatine phosphokinase, 10U of RiboLock RNase inhibitor (Fermentas), 200 ng of recombinant Gcn4 and 100 µM nucleoside triphosphates (NTPs). RNA was isolated with standard phenol chloroform extraction procedure. In vitro generated transcripts were analyzed by standard primer extension assay using a 5’ Cy5 labeled primer for the reverse transcription. The

generated cDNA was further precipitated and analyzed on a 7M urea and 8% polyacrylamide gel using a Typhoon 9400 scanner. Quantification was done using ImageQuant software (GE Healthcare).

2.2.6.2.3 Add back assay:

Add back assay was performed with both yeast purified Mud2 as well as with the recombinant Mud2. Mud2 was purified from *MUD2-TAP* yeast strain using TAP till TEV elution (Section 2.2.3.1), except that after IgG binding, IgG-coupled sepharose beads were washed with the TAP buffer containing 150mM NaCl. Purification from a non-tagged yeast strain served as a negative control for the experiment. Purification of Mud2 from *E.coli* is explained before in Section 2.2.3.2.

2.2.7 CTD-peptide pull down assay

The CTD-peptide pull down assay was performed as described in (Meinel, Burkert-Kautzsch et al. 2013) with some modifications. For each reaction, 50 µl of Streptavidin coupled magnetic beads M280 (Invitrogen) were washed three times with Peptide-Assay High Salt (PA-HS) buffer. The beads were resuspended again in PA-HS buffer (100 µl volume per reaction) and incubated with 10 µl of the non-phosphorylated or S2P phosphorylated biotinylated CTD peptide (1 µg/µl dissolved in water) for 2 h at turning wheel (4°C). The details of the peptide sequences are provided in Table 15. After incubation, beads bound with peptides were washed again one time with PA-HS buffer and two times with Peptide-Assay Low Salt (PA-LS) buffer. The phosphatase inhibitors (PhosI) were added to all the buffers used in this assay and low protein binding tubes were used for the binding reactions.

The proteins to be tested for CTD peptide binding were purified with a TAP tag until TEV elution as described before (Section 2.2.3.1) except that the TAP-tagged protein were bound to IgG coupled tosylactivated Dynabeads M-280 instead of usual IgG Sepharose beads. After binding, the IgG coupled beads were washed on a Dynabeads MPC (Magnetic Particle Concentrator) in 1ml of TAP buffer (100mM NaCl) + 0.5mM DTT. The beads were resuspended in 70 - 100 µl of TAP buffer and treated with 5 µl of TEV protease (4 mg/ml) for 1 h on a thermoshaker (500 rpm at 16°C, precooled in cold room). For Mud2-TAP purification,

2x beads (from 2x 2l cultures) were combined together for 1x TEV cleavage reaction (in order to concentrate the protein). The reason for changing the TAP protocol was to improve the purification and obtain highly clean and concentrated protein.

To test for the CTD peptide binding, TEV-eluates (TEV-E) of the tagged or the non-tagged (as a control) proteins were added to the washed and peptide bound streptavidin beads for 90 min at 4°C on a turning wheel. As observed after several trials, the minimum amount of Mud2 protein required for CTD binding is much more than required for the positive control (Pcf11-TAP), around 25 µl of TEV-E was added to each CTD peptide, while for the Pcf11-TAP (positive control), Rix1-TAP (negative control) and non-tagged strain (negative control), only 5 µl of TEV-E was used for the incubation with the peptide. After incubation, the unbound fractions were collected and the beads were washed 4 times with 500 µl of PA-LS buffer on a Dynabeads MPC. During each washing step, the beads (resuspended in wash buffer) were put on a turning wheel for 2 min (for effective washing). After washing, the beads were resuspended in 20 µl 1x SDS sample loading buffer and cooked for 2 min at 95°C in order to elute the peptide bound protein. Both the bound and unbound proteins were analysed with a western blot using an anti-CBP antibody.

2.2.8 *In vitro* binding assay for Prp19C-THO interaction

To purify Prp19C, Syf1-TAP strain was TAP purified from 12 l (6x purifications) of yeast culture till EGTA elution (EGTA-E). Basically, two TEV-eluates were combined for incubation with calmodulin beads. The 3x EGTA-eluates thus obtained were combined and the eluted protein was concentrated using a centricon (Millipore). Hpr1-TAP strain was used to purify the THO complex. Ceg1-TAP strain was used as a control for the experiment. The lysates from both Hpr1-TAP and Ceg1-TAP strains were treated with 100 µg/ml of RNase A during incubation with IgG Sepharose beads. The protein bound beads were washed after IgG incubation with 10 ml of high salt TAP buffer (1M NaCl) with 0.5 mM DTT, followed by an additional wash with 2 ml of standard TAP buffer containing 0.5 mM DTT. Washing with high salt removes Yra1, Sub2, Hrb1 and Gbp2 proteins from the TREX complex and yields highly pure THO complex. The same amount of purified and concentrated Syf1 protein (Prp19C) was incubated with Hpr1-TAP and Ceg1-TAP bound to IgG-coupled sepharose beads at 4°C on a turning wheel. After incubation, the beads were washed with 5ml of standard TAP buffer + 0.5

mM DTT, the proteins were eluted by incubation with TEV protease and separated by SDS-PAGE.

2.2.9 oligo(dT)-in situ hybridization

10 ml of yeast cells were grown till OD_{600} 0.5-1.0 (mid-log phase) and 1.25 ml of formaldehyde was added to cross link the cells for 90 min at RT on a turning wheel. For *mex67-5* cells, the cells were grown till mid-log phase at 30°C and shifted for 15 min to 37°C before adding formaldehyde. The cells were spinned down for 5 min at 3600 rpm. The cells were washed with 5 ml of 0.1 M KPO_4 buffer (pH 6.4) and again with 1 ml 0.1 M KPO_4 (pH 6.4) and transferred to the eppendorf tube. Then cells were spinned again and resuspended in 1 ml of 0.1 M KPO_4 plus 1.2 M sorbitol buffer (called spheroblasting buffer). The cells were spinned again, resuspended in 200 µl of spheroblasting buffer with 100 µg of Zymolyase 100T powder and incubated for 30 min at 30°C. To stop the Zymolyase 100T treatment, the cells were spinned for 4 min at 2,000 rpm. The cells were washed with 1 ml of spheroblasting buffer and the pellet was resuspended in >10x volume of the pellet. The spheroblasts were then attached to the glass slide pre-coated with polylysine and allowed to sit undisturbed for 5 min. Afterwards, the non-adherent cells were removed by aspiration. The adherent cells were further fixed by putting the slide in a staining jar with methanol at –80°C for 6 min and in staining jar with acetone at –80°C for 30 s. the staining jars with methanol and acetone were pre cooled at -80°C before starting the experiment. The cells were allowed to dry at RT for around 1 h. The dried cells were incubated first with 100 µl 2x SSC solution for 10 min at RT and then with 12 µl of prehybridisation buffer in a humid chamber for 1 h at 37°C. For hybridization with Cy3-labelled oligo $d(T)_{50}$ probe, 0.75 µl of 1 pmol/µl of probe was added and the slide was incubated at 37°C overnight in a humid chamber. After incubation, cells were washed with 100 ml of 0.5x SSC solution in staining jar (100 ml) at RT for 30 min. For staining of the nuclei, 2 µl of 4',6-diamidino-2-phenylindole (DAPI, 2.5 mg/ml stock solution) was added to the 100 ml of 0.5x SSC solution in the staining jar and the slide was incubated inside it for 3 min. Afterwards, the slide was washed again in fresh 0.5x SSC solution for 5 min at RT. The slide was allowed to dry at RT in the dark. The dried slides were finally covered with 80% glycerol in PBS. Olympus BX60 fluorescence microscope was used for analyzing the results.

3 Results

3.1 Role of Mud2 in transcription

3.1.1 Recruitment of Mud2 to the transcription machinery *in vivo*

To assess the occupancy of Mud2 at transcribed genes, chromatin immunoprecipitation (ChIP) experiments were performed with an endogenously TAP-tagged Mud2 strain (*MUD2*-TAP strain). Mud2 occupancy was determined by Real Time PCR at different positions of four intron-containing genes *DBP2*, *ACT1*, *RPS14B* and *RPL28* and four intronless genes *PMA1*, *ILV5*, *CCW12* and *RPL9B* (Figure 10). All these genes are highly transcribed in yeast.

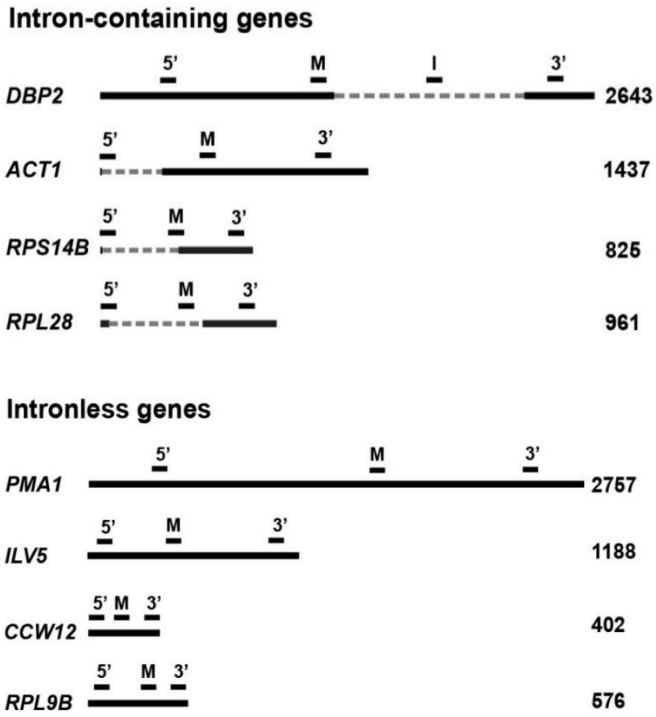


Figure 10. Schematic representation of four intron-containing and four intronless genes used for ChIP analysis.
The solid line represents an open reading frame (ORF) and the hatched line represents an intron. The bars above the solid line represent the position of primer pairs used for ChIP analysis. 5': 5' end of the gene, M: middle of the gene, 3': 3' end of the gene and I: within the intronic region.

A non-transcribed gene region (NTR1, 174131–174200 on chr. V) served as a negative control for the experiment. The occupancy of Mud2 at a specific gene was calculated by its enrichment at that gene relative to its presence at the NTR1. It was observed that Mud2 is recruited not only to the intron-containing genes where splicing is required, but also to the intronless genes (Figure 11). Also, Mud2 is recruited differentially to the different regions of a gene. For most of the analyzed genes, the occupancy was higher at the middle or the 3' region of the gene. Among the intron-containing genes, the occupancy of Mud2 was higher at the intronic site or immediately after the intronic region ends. Particularly, it showed high peaks at the intronic and the 3' region of *DBP2* gene, and also at the middle and the 3' region of *RPL28* gene. Among the intronless genes, the occupancy of Mud2 did not show a specific pattern. Altogether, it can be concluded that Mud2 is recruited to intron-containing and intronless genes and its occupancy is differentially distributed over the length of the gene.

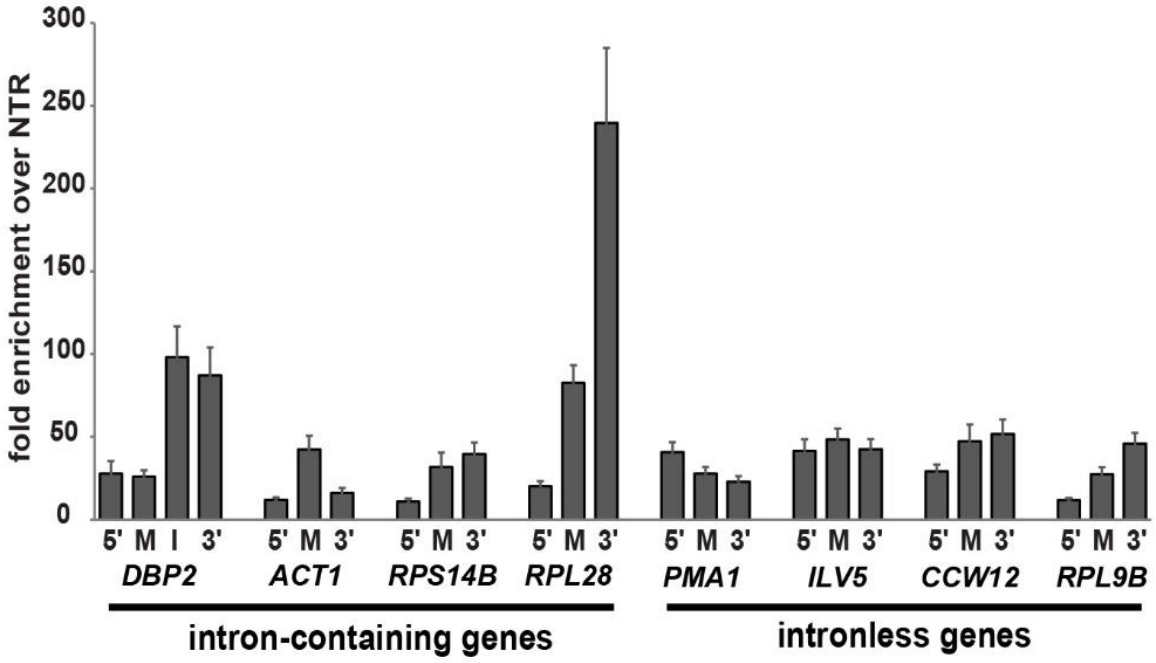


Figure 11. Mud2 is present at intronless genes.
The occupancy of Mud2 analyzed at various positions of four intron-containing genes (*DBP2*, *ACT1*, *RPS14B* and *RPL28*) and four intronless genes (*PMA1*, *ILV5*, *CCW12* and *RPL9B)* by chromatin immunoprecipitation (ChIP). ChIP was quantified by Real Time PCR using the primer pairs shown in Figure 10.

3.1.2 Recruitment of Mud2 is transcription dependent

Next, it was investigated whether recruitment of Mud2 to genes is dependent on active transcription. For this, the occupancy of Mud2 at the galactose inducible gene *GAL1* was assessed. *GAL1* gene is not transcribed when cells are grown under repressed conditions (glucose-containing medium). Its transcription can be induced by shifting the cells to induced conditions (galactose-containing medium). The occupancy of Mud2 was calculated by ChIP both under repressive and inducible conditions. The primers for qPCR amplified the *5'*, middle (M) and the *3'* positions of *GAL1* ORF. It was observed that Mud2 is recruited to *GAL1* only when the *GAL1* gene is transcribed i.e. in the presence of galactose (Figure 12). Therefore, recruitment of Mud2 is dependent on active transcription by RNAPII. Thus, it can be concluded that Mud2 is recruited not only to the intron-containing, but also to the intronless genes and this recruitment of Mud2 is dependent on active transcription.

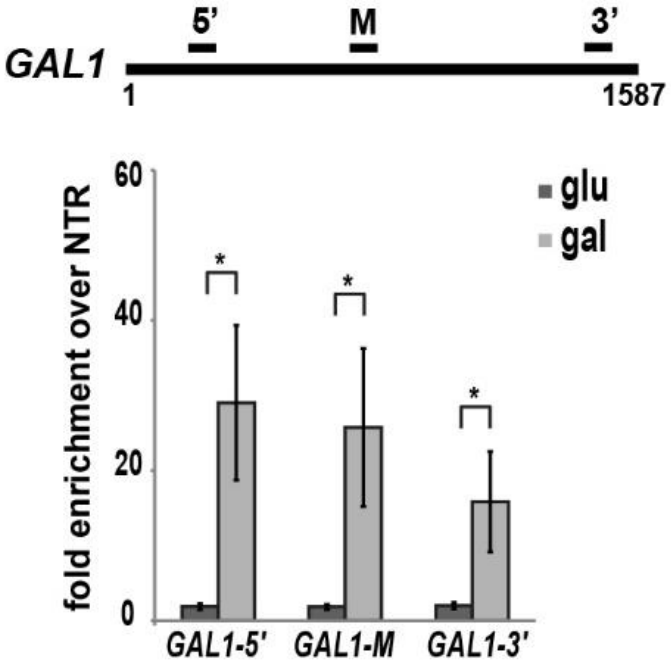


Figure 12. Recruitment of Mud2 to the transcribed genes is transcription dependent.
Occupancy of Mud2 at *GAL1* using ChIP under repressed conditions (glucose-containing medium: glu), and induced conditions (galactose-containing medium: gal). Primer pairs used for Real Time PCR are indicated in the top panel.

3.1.3 S2 CTD phosphorylation is required for the occupancy of Mud2 and Prp19C *in vivo*:

The Rpb1 CTD is the biggest platform for recruitment of various mRNP biogenesis factors to the transcriptional machinery. It has been published that in *S. cerevisiae*, the S2 CTD phosphorylation is required for TREX recruitment *in vivo* (Meinel, Burkert-Kautzsch et al. 2013). Also in mammalian cells, Mud2 binds directly to the S2 phosphorylated CTD (David, Boyne et al. 2011). Thus, the next objective was to investigate whether the recruitment of Mud2 and Prp19C in yeast is also dependent on S2 phosphorylation. To achieve this, the occupancy of TAP-tagged Mud2 in *rpb1 S2A* mutant cells compared to the wild-type cells was assessed by ChIP. The *rpb1 S2A* mutant cells carried nine wild-type and six S2A heptapeptide repeats while the wild-type cells (*WT*) carried 14 wild-type heptapeptide repeats in RNAPII CTD (West and Corden 1995).

As the occupancy of Mud2 was higher at the middle and 3' positions of most of the analyzed genes (except for *DBP2* and *PMA1*) (Figure 11), qPCRs were performed using primer pairs which amplify these regions. For *DBP2* and *PMA1*, primers for qPCRs amplified I and 3', and 5' and 3' regions respectively. Interestingly, the occupancy of Mud2 decreased significantly in *rpb1 S2A* mutants at all the gene positions analyzed in both intron-containing as well as intronless genes (Figure 13). Thus, Mud2 requires phosphorylation of Serine 2 in the CTD for maintaining its full occupancy at transcribed genes.

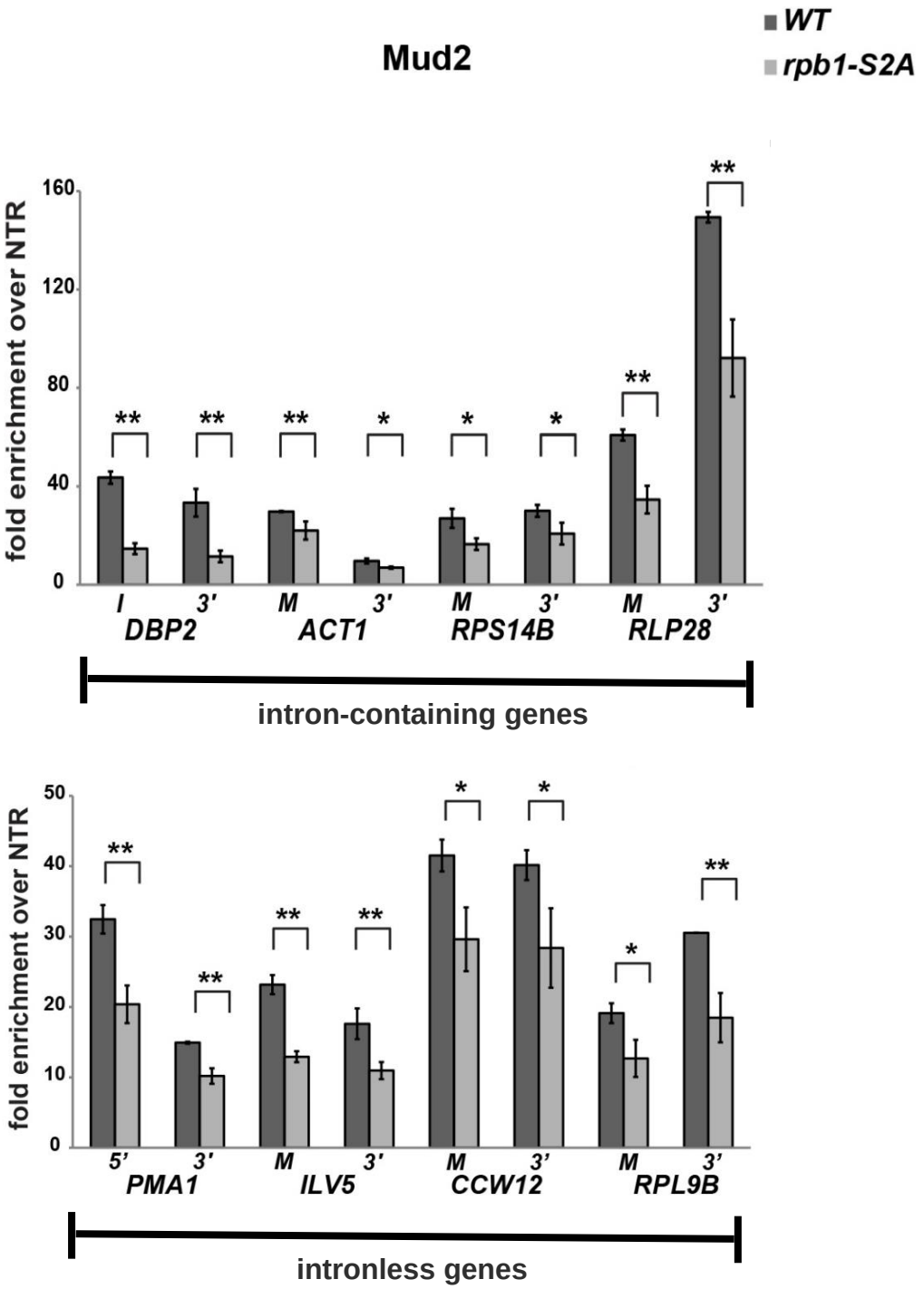


Figure 13. Serine 2 phosphorylation is necessary for Mud2 occupancy at transcribed genes.
Occupancy of Mud2 in wild-type cells (*WT*) and S2A mutant cells (*rpb1- S2A*) at four intron-containing and four intronless genes as assessed by ChIP. ChIPs were quantified by Real Time PCR using the primer pairs shown in Figure 10.

If Mud2 recruits Prp19C to genes, the occupancy of Prp19C should also decrease in the *rpb1 S2A* mutants. This was assessed by the ChIP of TAP-tagged Syf1 in *WT* and *rpb1 S2A* mutant cells. Indeed, the occupancy of Syf1 subunit of Prp19C decreased significantly in these mutants at all the analyzed genes (Figure 14).

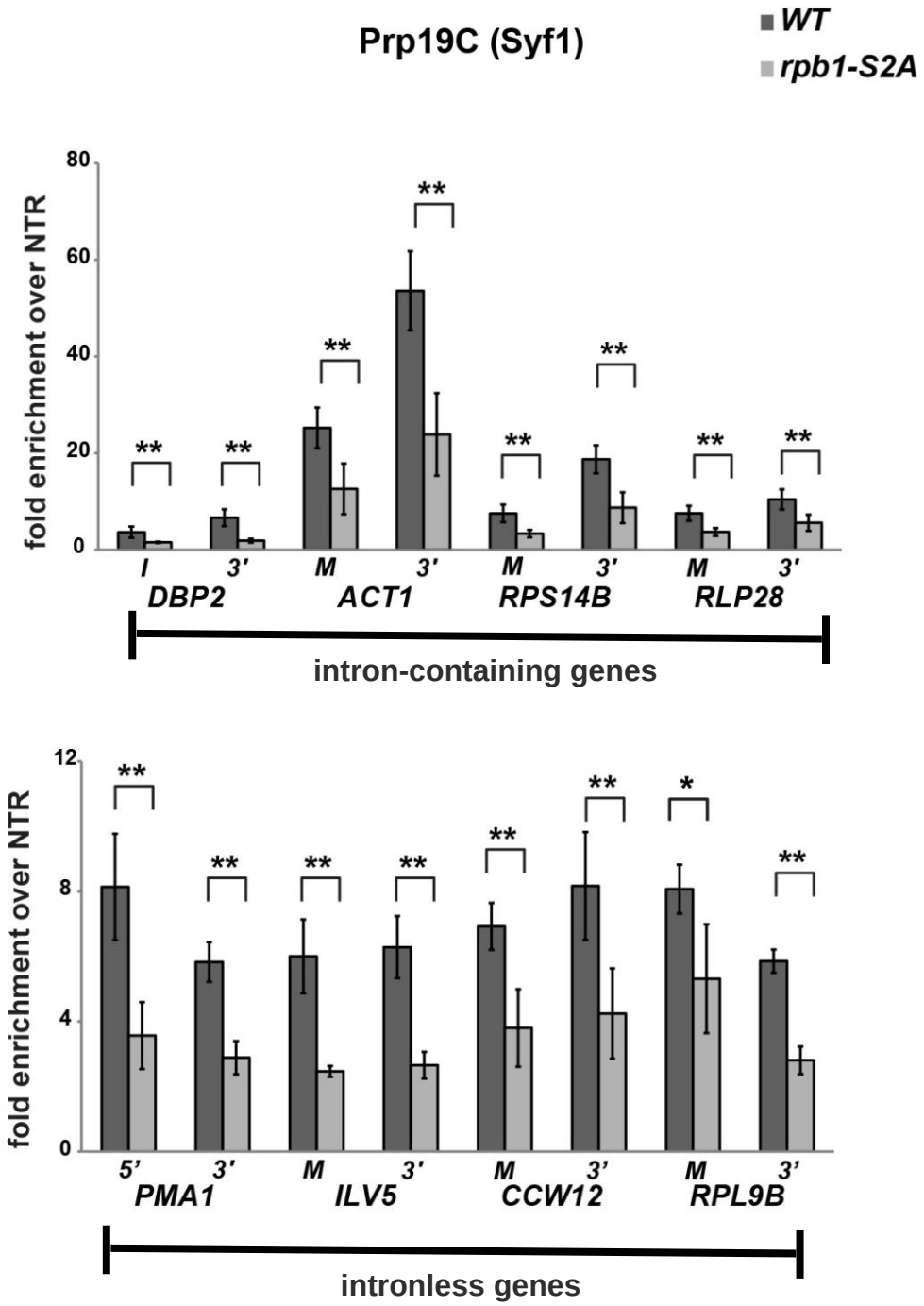


Figure 14. Serine 2 phosphorylation is necessary for Syf1 occupancy at transcribed genes.
Occupancy of Prp19C subunit Syf1 in the wild-type cells (*WT*) and S2A mutant cells (*rpb1- S2A*) at four intron-containing and four intronless genes as assessed by ChIP. ChIPs were quantified by Real Time PCR using the primer pairs as shown in Figure 10.

As Syf1 is an essential and core component of the stoichiometrically stable Prp19C, most likely the occupancy of whole Prp19C is decreased in these mutants. Thus, it can be concluded that S2 phosphorylation in CTD is necessary for the full occupancy of both Mud2 and Prp19C at transcribed genes. Also, as expected, S2 phosphorylated CTD decreases significantly in *rpb1 S2A* mutants as assessed by ChIP using an antibody which binds to the Ser2P (Figure 15).

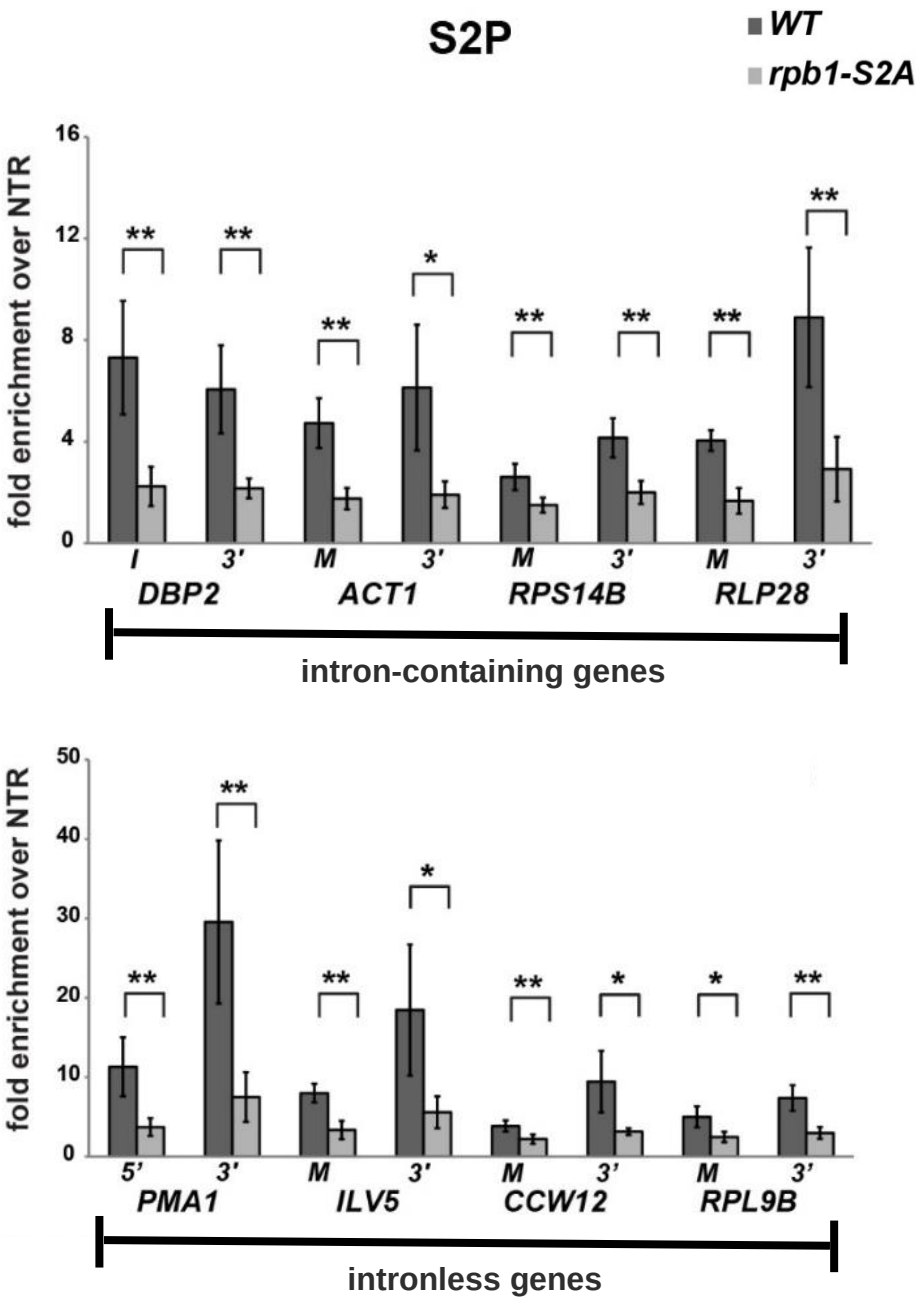


Figure 15. The occupancy of RNAPII with serine 2 phosphorylation decreases in S2A mutant cells. Occupancy of RNAPII with S2 phosphorylated CTD in wild-type cells (*WT*) and S2A mutant cells (*rpb1-S2A*) at four intron-containing and four intronless genes as assessed by ChIP. ChIPs were quantified by Real Time PCR using the primer pairs as shown in Figure 10.

To determine whether the decrease in Mud2 and Prp19C occupancy is because of the decrease in occupancy of RNAPII at genes, ChIP experiments were performed in *WT* and S2A mutant cells using yN-18 antibody directed against the N-terminus of Rpb1. RNAPII occupancy is not affected in the *rpb1 S2A* cells at all the genes and positions analyzed (Figure 16). This suggests that loss of Mud2 and Prp19C from transcribed genes is not due to a general decrease in transcription efficiency in mutant cells. Therefore, S2 phosphorylation is essential for the full occupancy of Mud2 and Prp19C.

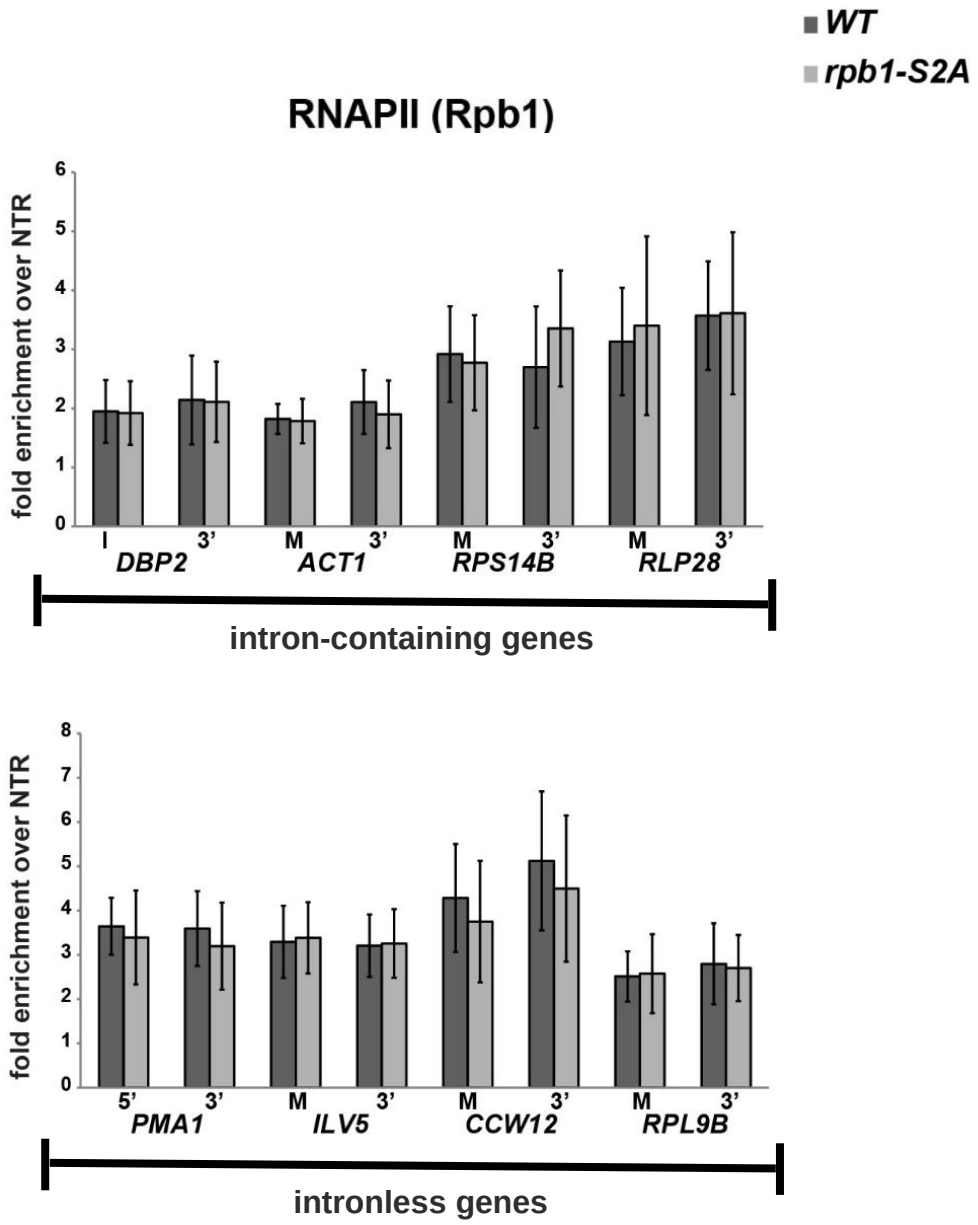


Figure 16. The occupancy of RNAPII is not affected in S2A mutant cells.
Occupancy of RNAPII subunit Rpb1 in wild-type cells (*WT*) and S2A mutant cells (*rpb1- S2A*) at four intron-containing and four intronless genes as assessed by ChIP. ChIPs were quantified by Real Time PCR using the primer pairs as shown in Figure 10.

3.1.4 Mud2 binds to S2 phosphorylated CTD

As shown in the previous section, S2 phosphorylation is required for Mud2 occupancy *in vivo*. Also, the presumptive orthologue of Mud2 in humans, U2AF65, has been shown previously to bind directly to the phosphorylated CTD *in vitro* (David, Boyne et al. 2011). Therefore, the next objective was to elucidate whether Mud2 binds directly to the S2 phosphorylated CTD *in vitro*. This was achieved by performing CTD peptide pull down assay with TAP purified Mud2 protein. The N-terminal biotinylated CTD peptides, either non-phosphorylated or S2 phosphorylated were immobilized on M-280 Streptavidin Dynabeads and incubated with purified Mud2. Each peptide consisted of three heptapeptide repeats of CTD, and two of these repeats were phosphorylated at S2 in the S2P CTD peptide. After incubating with the purified protein, the beads were washed and eluted. Bound and non-bound fractions were analyzed with an anti-CBP (calmodulin binding peptide) Western blot. Highly concentrated Mud2 was obtained as TEV-eluate (TEV-E) from *MUD2-TAP* strain by purifying with IgG coupled tosylactivated Dynabeads M-280 at 100mM NaCl (usual TAP buffer) and eluting in less final volume (to obtain concentrated protein; Figure 17).

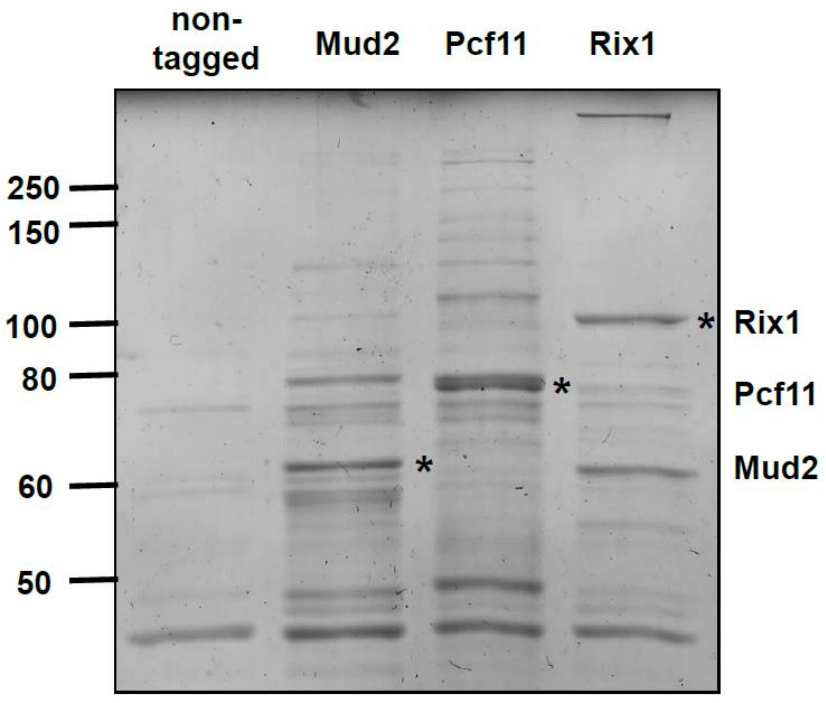


Figure 17. Purification of *MUD2-TAP, PCF11-TAP* and *RIX1-TAP* from *S. cerevisiae*.
The non-tagged (negative control), TAP-tagged Mud2, TAP-tagged Pcf11 and TAP-tagged Rix1 proteins were purified by Tandem Affinity Purification (TAP) till TEV-eluates. The polyacrylamide gel was stained with Coomassie brilliant blue. The purified proteins (marked as asterisks) were used for the CTD peptide pull down assay.

TAP-tagged Pcf11, a component of cleavage factor IA which binds to the Ser2P CTD (Lunde, Reichow et al. 2010), served as a positive control while TAP-tagged Rix1, a protein involved in rRNA processing (Nissan, Galani et al. 2004), served as a negative control for the experiment. TEV-eluates from wild-type (non-tagged), *MUD2-TAP*, *PCF11-TAP* and *RIX1-TAP* strains are shown on a Coomassie stained polyacrylamide gel (Figure 17).

Mud2 binds strongly to the S2P CTD peptide, but very weakly to the non-phosphorylated CTD peptide (Figure 18). A very strong interaction of Pcf11 with the S2P CTD peptide was also observed. Rix1 did not bind to any of the CTD peptides. In conclusion, Mud2 binds strongly to the S2 phosphorylated CTD *in vitro* which is consistent with our previous result that occupancy of Mud2 decreases in the *rpb1 S2A* mutant cells.

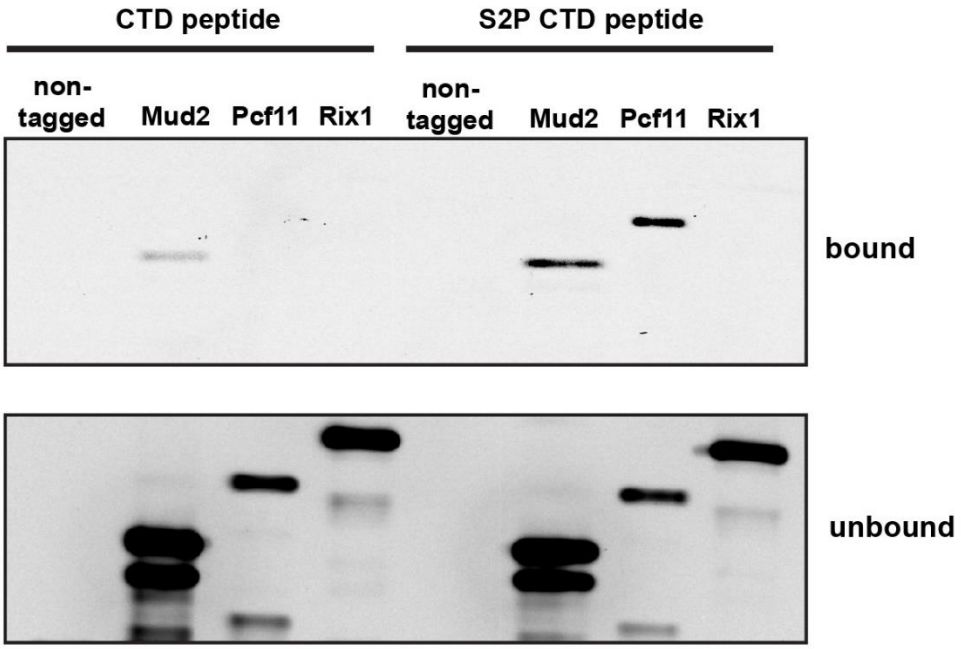


Figure 18. Mud2 binds to S2P phosphorylated CTD *in vitro*.
In vitro pull down assay of purified Mud2 with unphosphorylated and S2P phosphorylated CTD peptide. Pcf11 protein, a member of 3' processing machinery served as a positive control and an unrelated protein Rix1 served as a negative control for the experiment. Bound and unbound fractions were analyzed by Western blotting against CBP.

3.1.5 Mud2 interacts with Prp19C *in vivo* in an RNA-independent manner

U2AF65 interacts with Prp19C *in vitro* as well as *in vivo* in an RNA-independent manner (David, Boyne et al. 2011). In *S. cerevisiae*, Mud2 interacts with Syf3, a component of Prp19C

(Chung, McLean et al. 1999). Also, Mud2 and Prp19C are both recruited to transcribed genes in S2P dependent manner (Figure 13 and 14), there is a possibility that one recruits the other. To address this possibility, physical interaction between Mud2 and Prp19C was tested *in vivo.* The yeast strains expressing *MUD2-TAP* with HA-tagged Prp19 (*PRP19-HA*) and expressing only *PRP19-HA* (to serve as a negative control) were prepared. Mud2 was purified by tandem affinity purification (TAP) under standard (100 mM NaCl) and high salt conditions (250 mM NaCl; Figure 19).

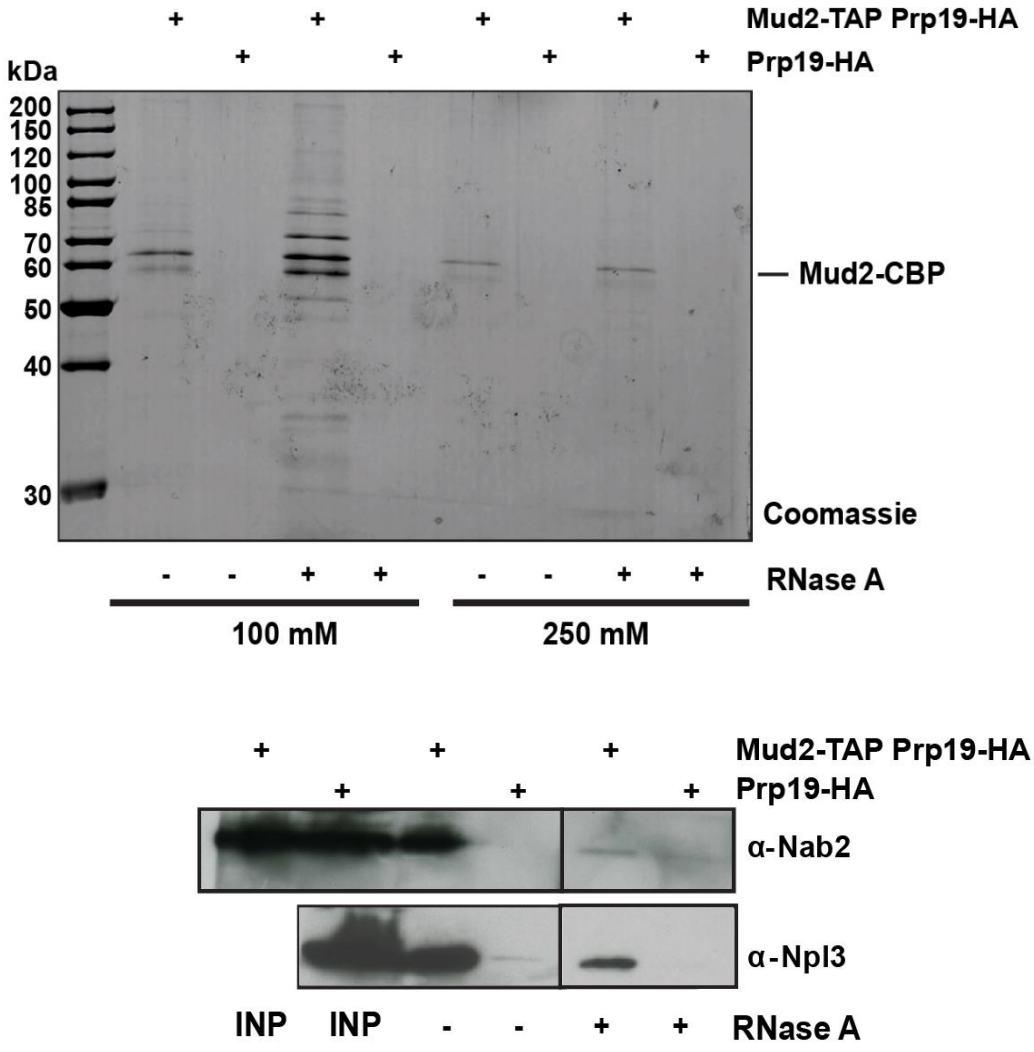


Figure 19. Purification of TAP-tagged Mud2 from *MUD2-TAP PRP19-HA and PRP19-HA* strains on a Coomassie gel.

Upper panel shows EGTA-eluates of Mud2-TAP purifications from *MUD2-TAP PRP19-HA* and *PRP19-HA* strains at 100mM and 250mM salt concentrations in the absence and presence of RNase A. The lower panel shows the Western blot of TEV-eluates of these purifications using Nab2 and Npl3 antibodies. Treatment of lysates with RNase A clearly abrogates and diminishes the co-purification of Nab2 and Npl3 proteins respectively.

The copurification of Prp19C was assessed by Western blotting against endogenous HA-tagged Prp19 (α-HA blot). As consistent with the previous results, Mud2 coimmunoprecipitated Prp19, and most likely the whole Prp19C, as the chemiluminescene signal of HA-tagged Prp19 was detected in the lane corresponding to the TAP tagged Mud2, but not in the lane corresponding to the negative control (*PRP19*-HA) (Figure 20). This interaction is stable even when Mud2 is purified under high salt conditions as Prp19C is still coimmunoprecipitated with Mud2 when the IgG beads are washed with the TAP buffer containing 250mM NaCl.

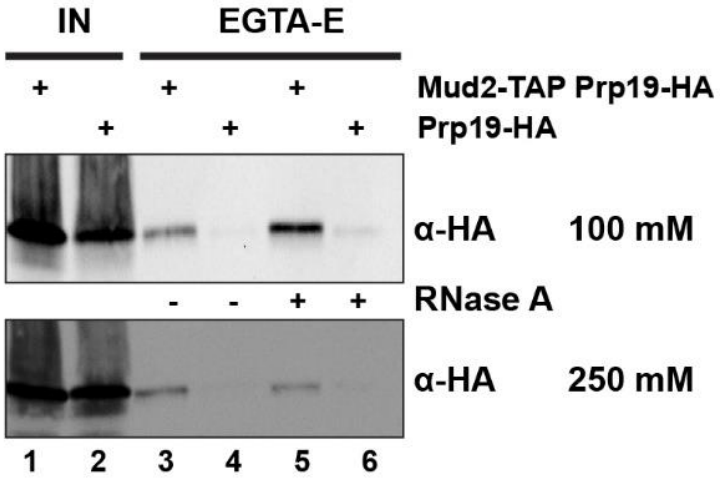


Figure 20. Mud2 interacts with Prp19C *in vivo* in an RNA-independent manner.
Mud2 was purified from the yeast strains expressing *MUD2-TAP PRP19-HA* or *PRP19*-HA. Copurification of Prp19C was assessed by Western blotting using α-HA antibody. To determine whether this interaction is mediated by RNA, cell lysate was treated with RNase A before purification. IN: Input; EGTA-E: EGTA eluate.

To further assess if this interaction between Mud2 and Prp19C is dependent on RNA, the experiment was performed both in the absence and presence of RNase A. Mud2 copurifies Prp19C even in the presence of RNase A. On the other hand, the copurification of mRNA binding proteins Nab2 and Npl3 was abolished and diminished respectively in the RNase treated samples, as assessed by the western blotting of TEV-eluates using Nab2 and Npl3 antibodies (Figure 19, lower panel). The loss of copurified Nab2 and Npl3 proteins in TAP purified Mud2 samples confirms that the RNase treatment worked during the purification. In conclusion, Mud2 interacts with Prp19C *in vivo* in an RNA-independent manner.

3.1.6 Mud2 is needed for the full occupancy of Prp19C and TREX

The next aim was to determine if Mud2 is needed for Prp19C occupancy at transcribed genes. In order to test this possibility, the occupancy of TAP-tagged Syf1 in *MUD2* deleted cells (*Δmud2*) compared to wild-type cells (*WT*) was analyzed using ChIP. As TREX, which is stabilized by Prp19C, has highest occupancy towards the 3' end of the gene (Meinel, Burkert-Kautzsch et al. 2013), the qPCRs were performed at 3' positions of intron-containing and intronless genes used in this study.

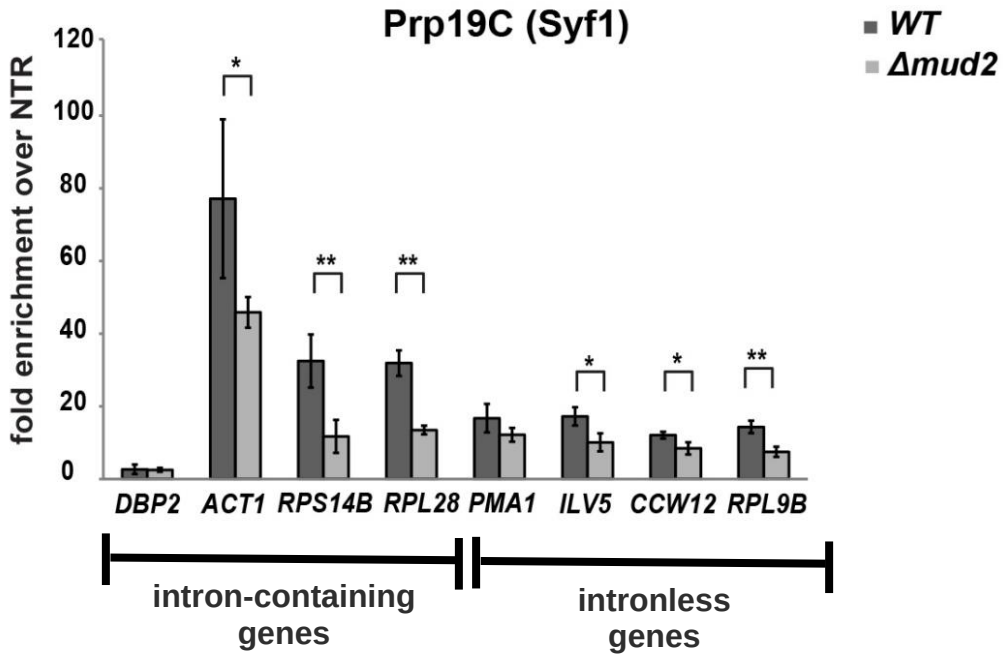


Figure 21. Syf1 occupancy decreases in *Δmud2* cells.
Occupancy of Syf1 (Prp19C) was assessed in *Δmud2* cells compared to the wild-type cells (*WT*) using ChIP. ChIPs were quantified by Real Time PCR using the primer pairs shown in Figure 10.

The occupancy of Syf1 decreased to ~ 50 % in *Δmud2* cells compared to the wild-type cells (Figure 21). As Syf1 is an essential component of the stoichiometrically stable Prp19C (Tarn, Hsu et al. 1994), this implies that the occupancy of the whole Prp19C is decreased in the absence of Mud2. As Prp19C stabilizes the TREX occupancy at transcribed genes, a decrease in Syf1 occupancy should also lead to the loss of TREX from genes. To test this possibility, the occupancy of TAP-tagged Hpr1, a component of the TREX complex, was assessed in wild-type

and *Δmud2* cells by ChIP. Indeed, the occupancy of Hpr1, and most likely the whole TREX complex is decreased in *Δmud2* cells compared to the wild-type cells (Figure 22).

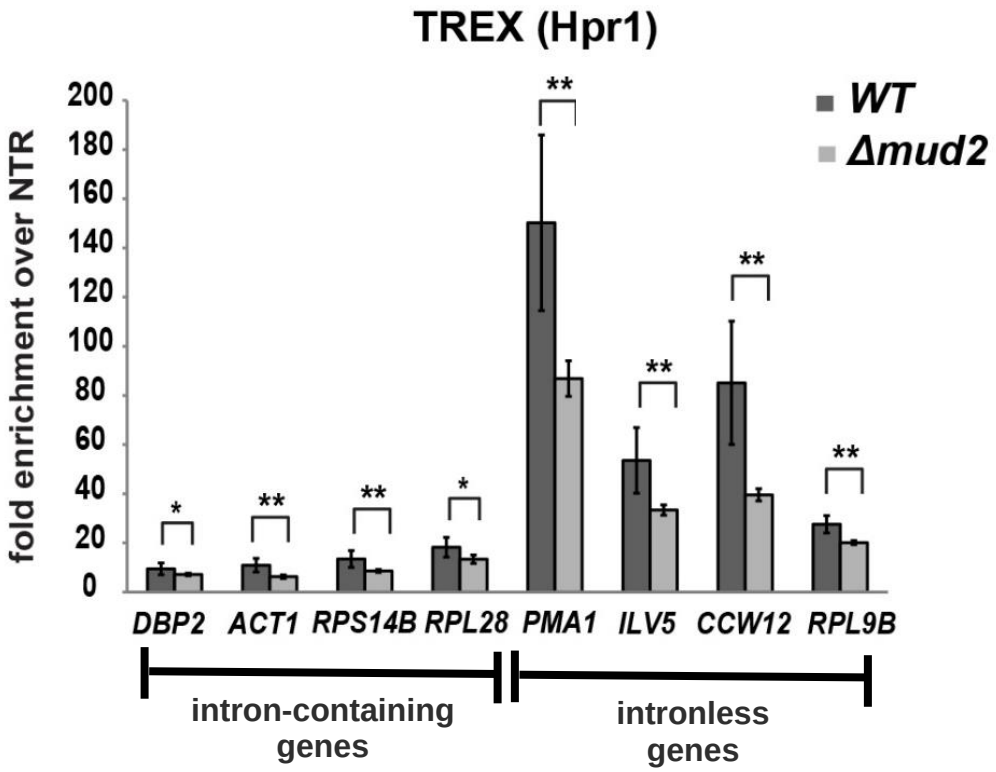


Figure 22. Hpr1 occupancy decreases in *Δmud2* cells.
Occupancy of Hpr1 (TREX) was assessed in *Δmud2* cells compared to the wild-type cells (*WT*) using ChIP. ChIPs were quantified by Real Time PCR using the primer pairs shown in Figure 10.

This decrease in Prp19C and TREX occupancy in not due to a general decrease in the transcription in *Δmud2* cells as the occupancy of Rpb1, as assessed by ChIP, is not affected in the *Δmud2* cells (Figure 23). Also, this decrease should not be due to a decrease in the S2 phosphorylated RNAPII occupancy in the *Δmud2* cells, as assessed by S2P ChIP in *Δmud2* cells (Figure 24). Therefore, the decrease of Prp19C and TREX occupancy in *Δmud2* cells is a specific effect of *MUD2* deletion.

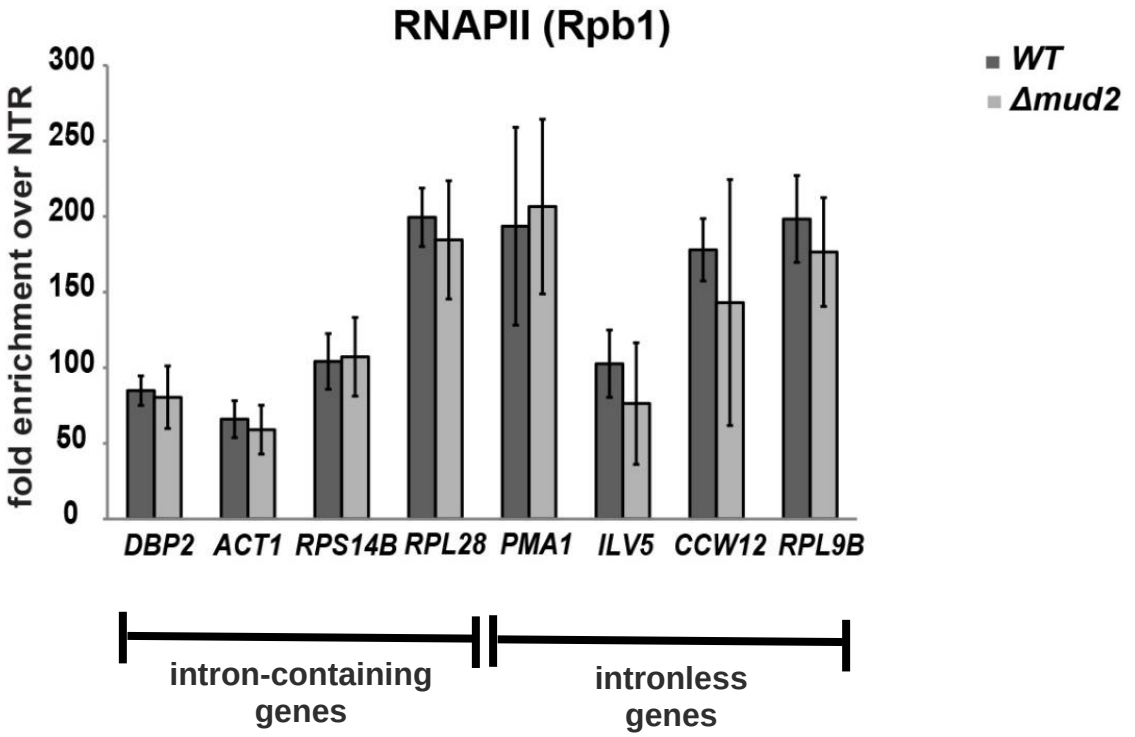


Figure 23. Rpb1 occupancy does not change in *Δmud2* cells.
Occupancy of Rpb1 (RNAPII) was assessed in *Δmud2* cells compared to the wild-type cells (*WT*) using ChIP. ChIPs were quantified by Real Time PCR using the primer pairs shown in Figure 10.

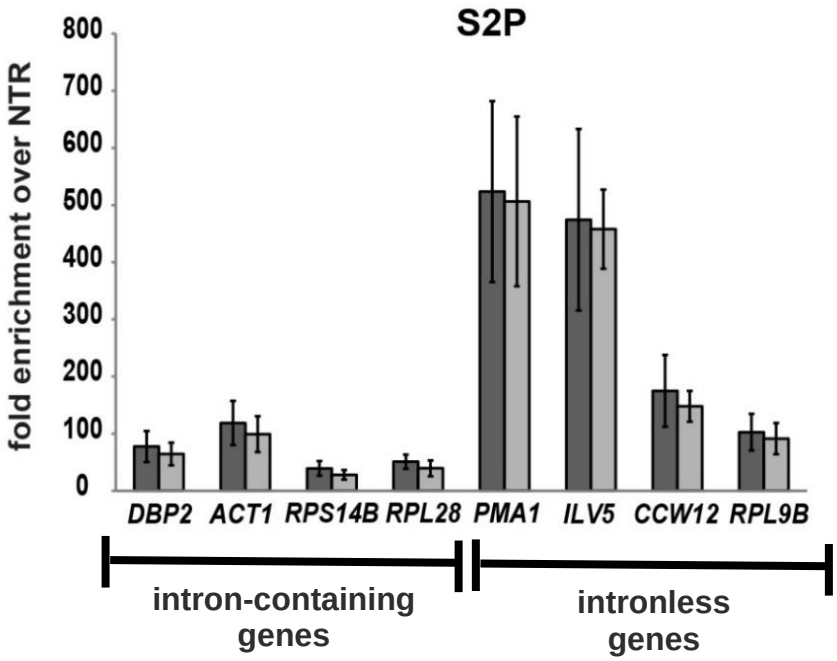


Figure 24. S2P Rpb1 occupancy does not change in *Δmud2* cells.
Occupancy of S2P Rpb1 (S2-phosphorylated RNAPII) was assessed in *Δmud2* cells compared to the wild-type cells (*WT*) using ChIP. ChIPs were quantified by Real Time PCR using the primer pairs shown in Figure 10.

To confirm that equivalent amounts of Prp19C, TREX and RNAPII complexes are expressed in both wild-type and mutant cells, Western blots to detect the total cellular levels of Syf1 (Prp19C), Hpr1 (TREX) and Rpb1 (RNAPII) were performed in wild-type and *Δmud2* cells. As determined from the quantification of these blots, there is a slight increase in the total cellular expression levels of Syf1 (Prp19C), no significant change in the total levels of Hpr1 (TREX) and a slight decrease in the total levels of Rpb1 (RNAPII) (Figure 25). This slight decrease of Rpb1 levels in *Δmud2* cells is still sufficient to provide the full occupancy of Rpb1 at transcribed genes. Nevertheless, ~ 50 % decrease of Prp19C and TREX occupancy in *Δmud2* cells should be a specific effect of *MUD2* deletion. In conclusion, Mud2 ensures full occupancy of Prp19C and TREX complex at transcribed genes.

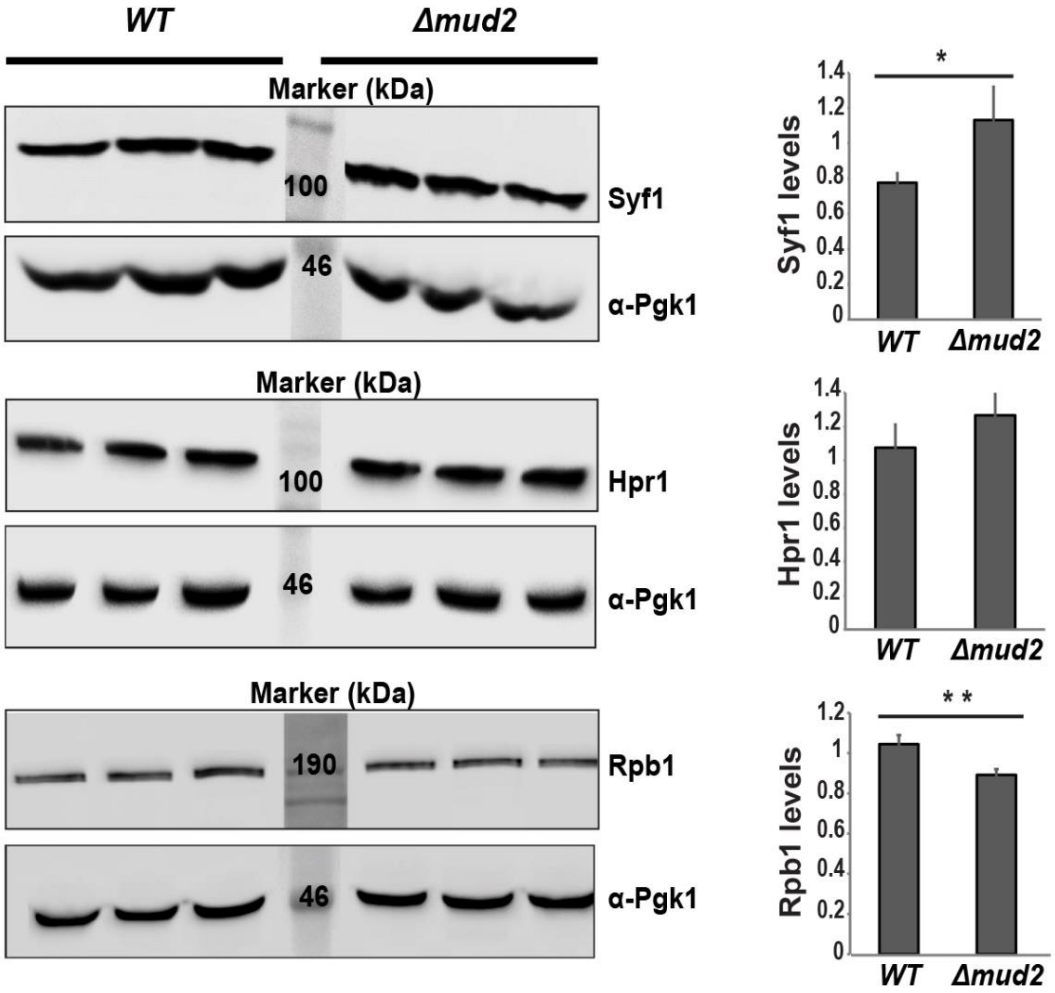


Figure 25. Total cellular levels of Syf1, Hpr1 and Rpb1 in wild-type (*WT*) and *Δmud2* cells.
Western blots were performed to determine the total cellular levels of Prp19C subunit Syf1 (detected via TAP tag), TREX subunit Hpr1 (detected via TAP tag) and RNAPII subunit Rpb1 (detected via 8WG16 antibody against Rpb1) in wild-type (*WT*) and *Δmud2* cells. The Western blot for the Pgk1 levels (detected via antibody against Pgk1) served as the loading control for the analysis. The quantification was performed with LabImage 1D software and is shown in the corresponding right panel of each blot.

3.1.7 Does Mud2 deletion affect the recruitment of other transcription elongation factors?

In order to determine whether the decrease of Prp19C and TREX occupancy in *Δmud2* cells is an indirect effect caused by the loss of other transcription elongation factors, the occupancy of the transcription elongation factors Paf1 (a component of the PAF complex) and Spt5 (which also acts as a recruitment platform for other transcription and processing factors) was assessed in wild-type and *Δmud2* cells by ChIP.

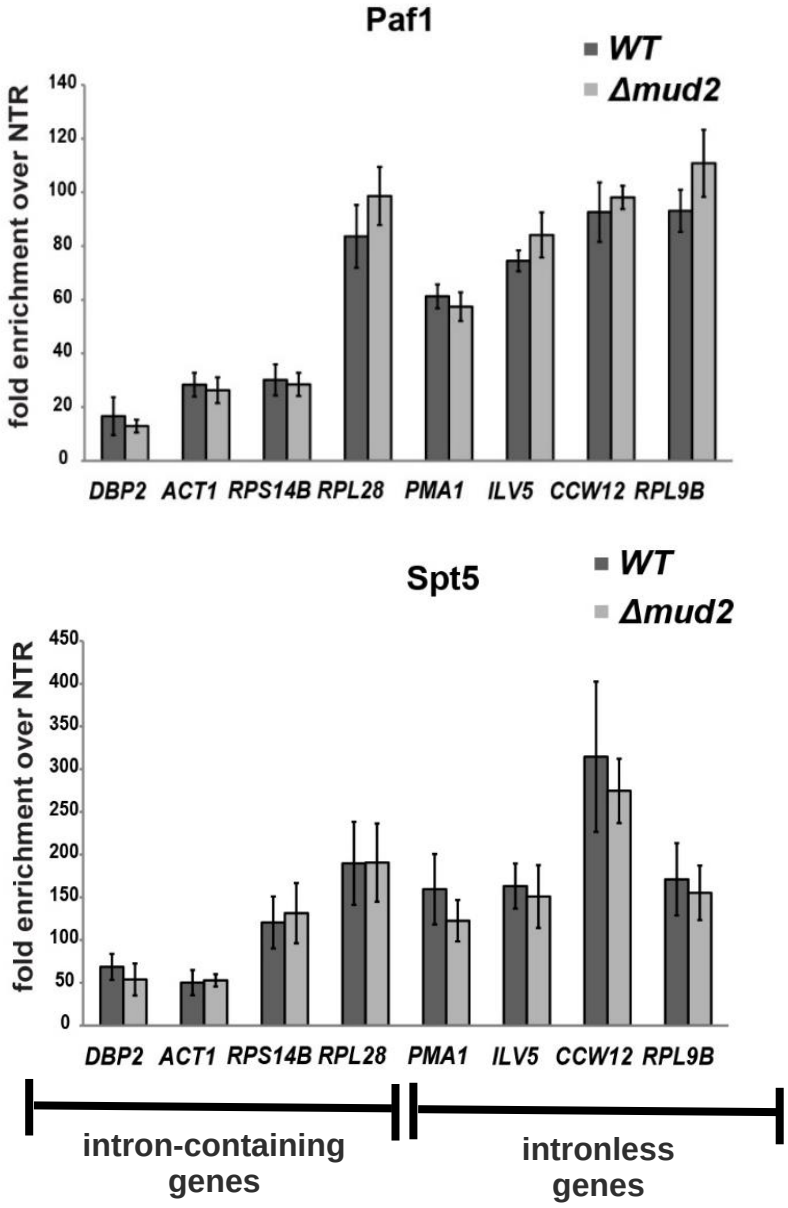


Figure 26. The occupancy of Paf1 and Spt5 is not affected in *MUD2* deleted cells.
Occupancy of transcription elongation factors Paf1 (member of PAF complex) and Spt5 (part of Spt4-Spt5 complex) was assessed in *Δmud2* cells compared to the wild-type cells (*WT*) using ChIP. ChIPs were quantified by Real Time PCR using the primer pairs shown in Figure 10.

It was observed that the occupancy of both elongation factors Paf1 and Spt5 does not change in *Δmud2* cells (Figure 26). Thus, decrease in Prp19C and TREX occupancy is a specific effect of *MUD2* deletion. This strongly suggests that Mud2 might be a recruiting factor for Prp19C and/or the TREX complex.

3.1.8 Does Mud2 function 'upstream' of Prp19C and TREX?

In order to gain more insights into the role of Mud2, the next objective was to elucidate if Mud2 acts 'upstream' of Prp19C and TREX. Thus, it was determined if a mutant of *SYF1* (which leads to the loss of Prp19C from the genes) or the deletion of *HPR1* results in the loss of Mud2 from transcribed genes. To test this possibility, a ChIP of TAP-tagged Mud2 in *syf1-37* mutant cells was performed. The *syf1-37* mutant cells contain a truncation in the C-terminal domain of Syf1. It has been shown previously that the *syf1-37* mutation leads to a loss of TREX from the transcribed genes (Chanarat, Seizl et al. 2011). In contrast, it was observed that Mud2 occupancy is not affected in the *syf1-37* cells (Figure 27). Therefore, Mud2 might recruit Prp19C and not vice versa. Also, deletion of *HPR1* does not lead to a loss of Mud2 from transcribed genes, as assessed by ChIP of TAP-tagged Mud2 in *Δhpr1* cells compared to the wild-type cells (Figure 28). Taken together, Mud2 is needed for the full Prp19C and TREX occupancy and most likely functions 'upstream' of Prp19C and TREX.

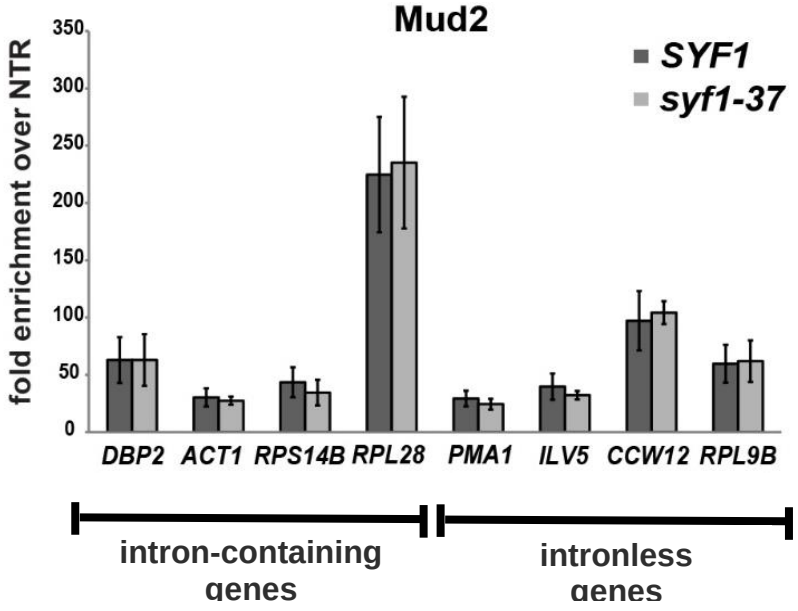


Figure 27. Mud2 occupancy does not change in *syf1-37* mutant cells.
Occupancy of Mud2 was assessed in *syf1-37* mutant cells compared to the wild-type cells (*WT*) using ChIP. ChIPs were quantified by Real Time PCR using the primer pairs shown in Figure 10.

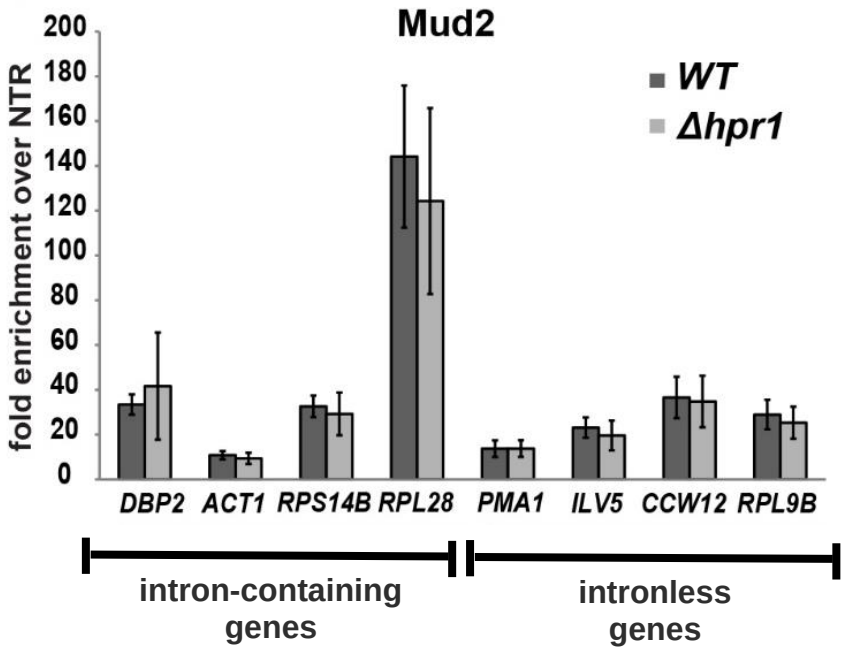


Figure 28. Mud2 occupancy does not change in *Δhpr1* mutant cells.
Occupancy of Mud2 was assessed in *Δhpr1* mutant cells compared to the wild-type cells (*WT*) using ChIP. ChIPs were quantified by Real Time PCR using the primer pairs shown in Figure 10.

3.1.9 Deletion of Mud2 causes 6-azauracil (6AU) sensitivity

Mud2 is recruited not only to the intron-containing but also to the intronless genes and the occupancy of Mud2 is dependent on the active transcription (Figure 11 and 12). Also, Mud2 interacts with Prp19C (Figure 20), a complex which has been previously reported to function in transcription elongation. Furthermore, the deletion of *MUD2* causes a loss of Prp19C and TREX complex from the transcribed genes. All these results are consistent with the possibility that Mud2 might be involved in the transcription.

To get a first indication for a role of Mud2 in transcription, the *Δmud2* cells were tested for sensitivity to a drug called 6-azauracil (6AU). 6AU inhibits transcription elongation by depleting intracellular GTP and UTP levels. This reduction in the cellular nucleotide levels causes a growth defect in cells that are deficient of an important factor involved in transcription elongation (Riles, Shaw et al. 2004, Zhou, Liu et al. 2015). Thus, 6AU sensitivity is an indication that a protein is involved in transcription elongation. Interestingly, *MUD*2 deletion conferred 6AU sensitivity to the cells (Figure 29). To test whether this drug sensitivity effect is also observed in other splicing factors which function in the similar pathways as Mud2, deletion mutants of Snu66 (a component of the U5/U4/U6 tri-snRNP), Lea1 (a component of U2

snRNP), Lin1 (a component of the U5 snRNP), Nam8 (a component of the U1 snRNP) or Mud1 (component of the U1 snRNP) were also tested for 6AU sensitivity. It was observed that all these splicing factors are not sensitive to 6AU (Figure 29). Thus, 6AU sensitivity of *Δmud2* cells is a specific effect of Mud2 deletion.

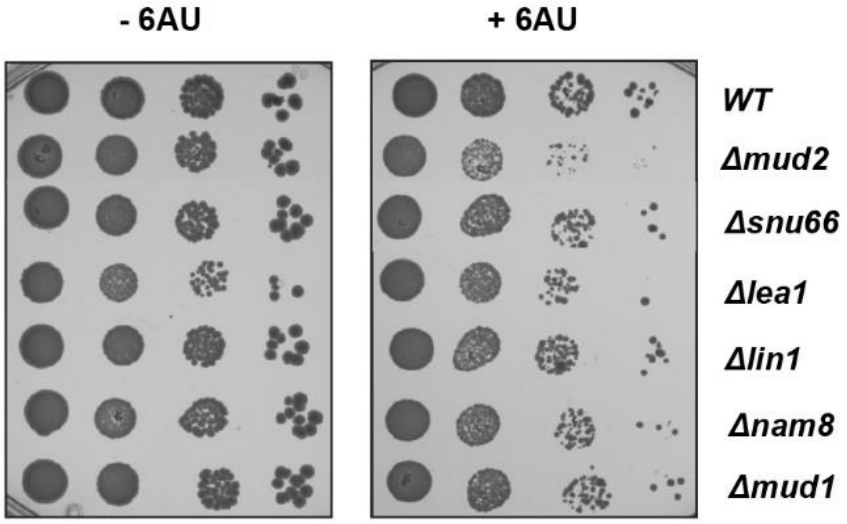


Figure 29. Deletion of *MUD2* confers sensitivity to 6-azauracil (6AU).
Dot spot assays with 10 fold serial dilutions of wild-type (*WT*), *Δmud2, Δsnu66, Δlea1, Δlin1, Δnam8* and *Δmud1* on SDC-ura plates containing DMSO (as solvent control) or 50 µg/ml of 6AU. The plates were incubated at 30°C for 2-3 days.

DST1, the gene which codes for transcription elongation factor TFIIS, enables RNAPII to resume transcription after arrest. TFIIS induces cleavage by RNAPII at the 3' end of the transcript so that the nascent transcript is positioned at the active site of RNAPII and hence the transcription can be resumed again (Wind and Reines 2000, Kettenberger, Armache et al. 2003). The null mutant of *DST1* is viable but sensitive to 6AU, one of the most prominent and well-studied phenotypes of *Δdst1* cells (Mason and Struhl 2005). In order to test the genetic interaction between Mud2 and TFIIS, 6AU sensitivity of *Δmud2* cells was compared to the 6AU sensitivity of *Δdst1* as well as double deletion *Δmud2 Δdst1* mutant cells. As expected, *Δdst1* cells are 6AU sensitive (Figure 30). Interestingly, double deletion of *MUD2* and *DST1* exhibits increased sensitivity to 6AU as compared to each single deletion (Figure 30). This genetic interaction between Mud2 and TFIIS indicates that Mud2 might play a direct role in the transcription elongation.

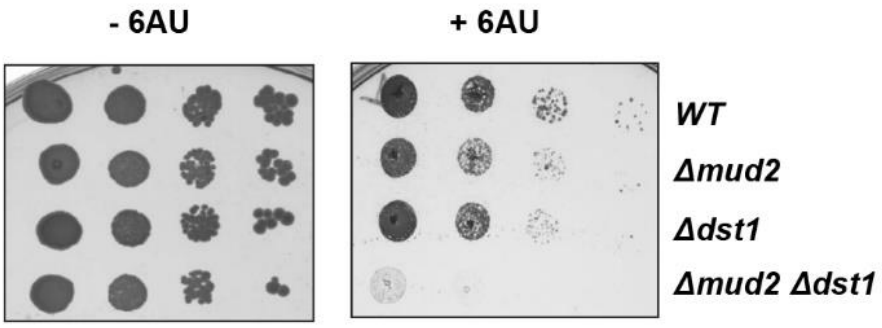
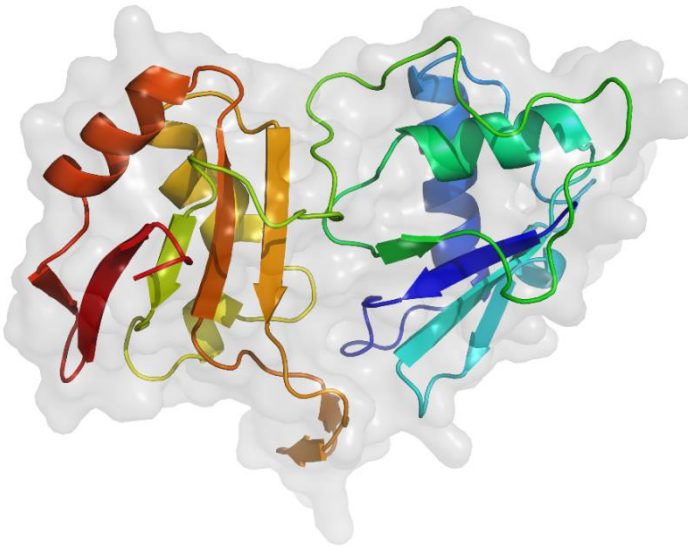


Figure 30. Deletion of *MUD2* is synthetic lethal with the deletion of *DST1* in presence of 6AU.
Dot spot growth assays with 10 fold serial dilutions of wild-type, *Δmud2*, *Δdst1* and *Δmud2 Δdst1* on SDC-ura plates containing DMSO (as solvent control) or 50 µg/ml of 6AU. The plates were incubated at 30°C for 2-3 days.

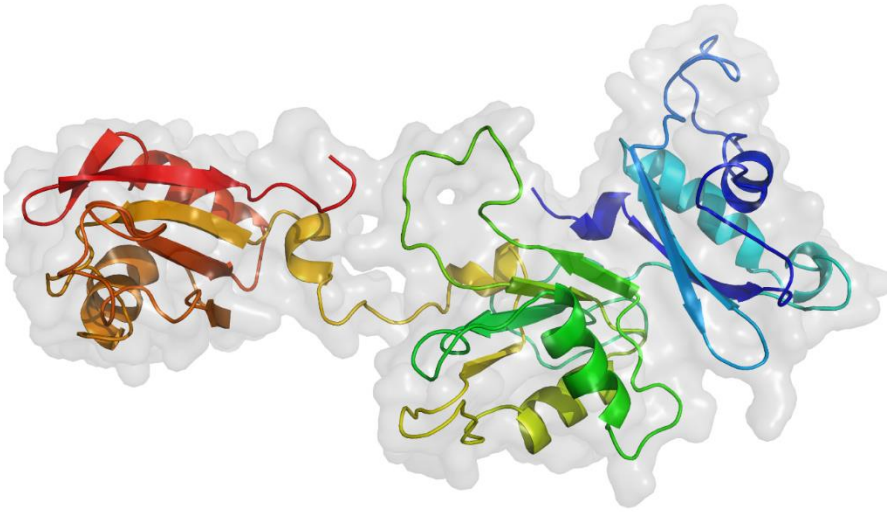
3.1.10 The three RRMs of Mud2 are required to complement the 6AU sensitivity of *Δmud2* cells

Mud2 contains a well characterized RNA Recognition Motif (RRM) and two less characterized RRMs at the C-terminus (Section 1.3). In order to determine which domain of Mud2 is required to rescue its sensitivity to 6AU, structural alignment of Mud2 with U2AF65 and splicing factor Prp24 was performed to predict the structure of less conserved RRMs. The predicted models as a result of alignment of Mud2 with U2AF65 and Prp24 are shown in Figure 31 and 32 respectively.



```
Target    MADEKRLEDLRSKIMESIGKSEKDVVPIENKRFNTDNAVIDTHFKRQKSDGELPKAPKSRNVSHSNNRGPSSIITMSTNR
5ev3.1.A  --------------------------------------------------------------------------------

Target    TTYEQTRAGPHRQSYRDASGRSYNRENRYSSHNTGPQWNNNPYNRQRDERRGRNERFDRRGRNGNGNYDRFNYQRKNEGS
5ev3.1.A  --------------------------------------------------------------------------------

Target    KFNGDRDKRQLQTNKYDMNYNSQNVMYPGSSFDSPAYYNMASSKANSRLVISGLSQSSDPSIVARLKDLLENFISGLQKT
5ev3.1.A  --------------------------------------------------------------------------------

Target    ESNAEDFKISNFYIGEGRPDHIIVEFSSQICSTMVLACRSFFNAKLGTFDLKWRRPNDYVQQLDHLVDFCRGTVIALENL
5ev3.1.A  ------------------------------------------------------------------------ARRLYVGNI

Target    ENIGEGEDYRMKELFSSLNVTNGTAK----PLFYKCSSNTNNTGKESEFTKCILLSFEVVTQDILDKL------KPYKWF
5ev3.1.A  PF--GITEEAMMDFFNAQMRLGGLTQAPGNPVLA------VQINQDKNFAFLEFRSVDETTQAM-AFDGIIFQGQSLKIR

Target    KPNDGKISQVTSWITFQSLPNLVTQSVRVESRVLLLLNCLDPLDLKDETFITEIKETLKYSIAGADTIKICQPGVDYRLN
5ev3.1.A  RPHDYQPLPGMSENPSVYVPGVVSTVVPDSAHKLFIGGL--PNYLN-DD---QVKE-LLTSFGPLKAFNLVKDSATG---

Target    FENLASGAGNIYIKFKTLEAAKHAMEELPGTQFNDRTVLCTYIDEDDFDMMEATQLS
5ev3.1.A  -----LSKGYAFCEYVDINVTDQAIAGLNGMQLGDKKLLVQRASV------------
```

Figure 31. Predicted Model of Mud2's less conserved RRM domains based on the structural alignment with U2AF65-RNA-binding domain (PDB ID: 5EV3).



```
Target    MADEKRLEDLRSKIMESIGKSEKDVVPIENKRFNTDNAVIDTHFKRQKSDGELPKAPKSRNVSHSNNRGPSSIITMSTNR
5tf6.2.A  --------------------------------------------------------------------------------

Target    TTYEQTRAGPHRQSYRDASGRSYNRENRYSSHNTGPQWNNNPYNRQRDERRGRNERFDRRGRNGNGNYDRFNYQRKNEGS
5tf6.2.A  --------------------------------------------------------------------------------

Target    KFNGDRDKRQLQTNKYDMNYNSQNVMYPGSSFDSPAYYNMASSKANSRLVISGLSQSSDPSIVARLKDLLENFISGLQKT
5tf6.2.A  ------------------------------------------RNRELTTVLVKNLPKSYNQN---KVYKYFKHC----G--

Target    ESNAEDFKISNFYIG---EGRPDHIIVEFSSQICSTMVLACRSFFNAKLGTFDLKWRRPNDYVQQLDHLVDFCRGTVIAL
5tf6.2.A  -------PIIHVDVADSLKKNFRFARIEFARYDGALAAITKTH---KVVGQNEIIVSHL--------------TECTLWM

Target    ENLENIGEGEDYRMKELFSSLNVTNGTAKPLFYKCSSNTNNTGKESEFTKCILLSFEVVTQDILDKLKPYKWFKPNDGKI
5tf6.2.A  TNFPP--SYTQRNIRDLLQDINV---VALSIRLPSL---R-FNTSRRFAYIDVTSKEDARYCVEKLNGLKI----EGYTL

Target    SQVTSWITFQSLPNLVTQSVRVESRVLLLLNCLDPLDLKDETFITEIKETLKYSIAGADTIKICQPGVDYRLNFENLASG
5tf6.2.A  VTKVSN--P-LEKSKRTDSATLEGREIMIRNL--STELLDENL---LRE-SFEGFGSIEKINIPAGQKEH-------SFN

Target    AGNIYIKFKTLEAAKHAMEELPGTQFNDRTVLCTYIDEDDFDMMEATQLS
5tf6.2.A  NCCAFMVFENKDSAERAL-QMNRSLLGNREISVSLADKKPF---------
```

Figure 32. Predicted Model of three RRM domains of Mud2 based on the structural alignment with U6 small nuclear ribonucleoprotein core (PDB ID. 5TF6)

Based on these structural alignments, the predicted domains of Mud2 were compared to the well characterized domains of U2AF65 and U2AF50 (Figure 33). To understand which domain of Mud2 is required to complement the 6AU sensitivity of *Δmud2* cells, six deletion mutants of *MUD2* along with full length *MUD2* were cloned on the plasmid *pRS315-TAP-T-ADH1* under the control of natural *MUD2* promoter. Three of them were N-terminal deletions and other three were C-terminal deletions. All of them were expressed with a TAP tag at their C-terminus.

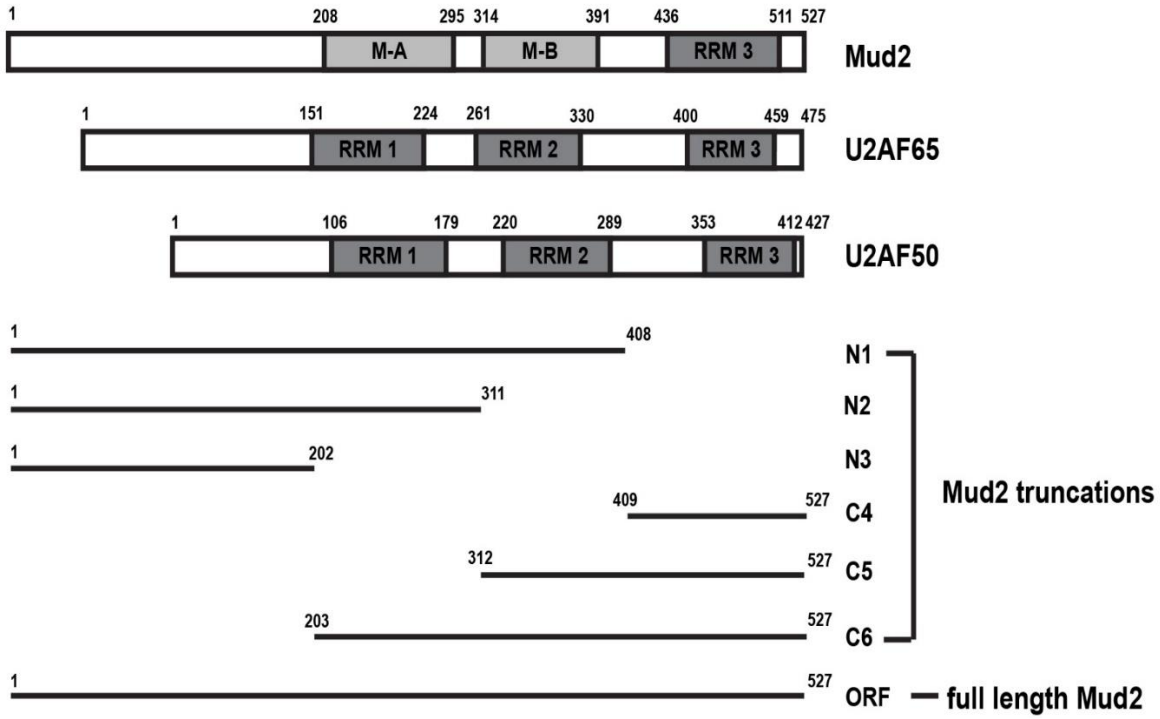


Figure 33. Schematic representation of the three RRM domains in Mud2, U2AF65 (mammalian Mud2) and U2AF50 (Drosophila Mud2).
The full length and six truncated versions of *MUD2* were cloned into the plasmid *pRS315-TAP-T-ADH1*. The boxes represent the primary structure of the protein and the shaded regions inside the boxes represent the RRM motifs. The numbers above the boxes refer to the amino acid position in the primary sequence of the protein (where number 1 refers to the first amino acid of the protein). The N-terminal and C-terminal truncations of Mud2 are shown as solid lines.

To test for 6AU sensitivity, *Δmud2* cells were transformed with the plasmids containing full length *MUD2* as well as all the six truncated mutants of *MUD2*. As controls, the wild-type and *Δmud2* cells were transformed with empty plasmid. As observed from the dot spots of these strains on control plates and plates with 6AU (Figure 34), only the cells expressing all three RRMs of Mud2 can rescue the 6AU sensitivity of *Δmud2* cells. Thus, only the N terminal domain of *MUD2* can be dispensed for its function.

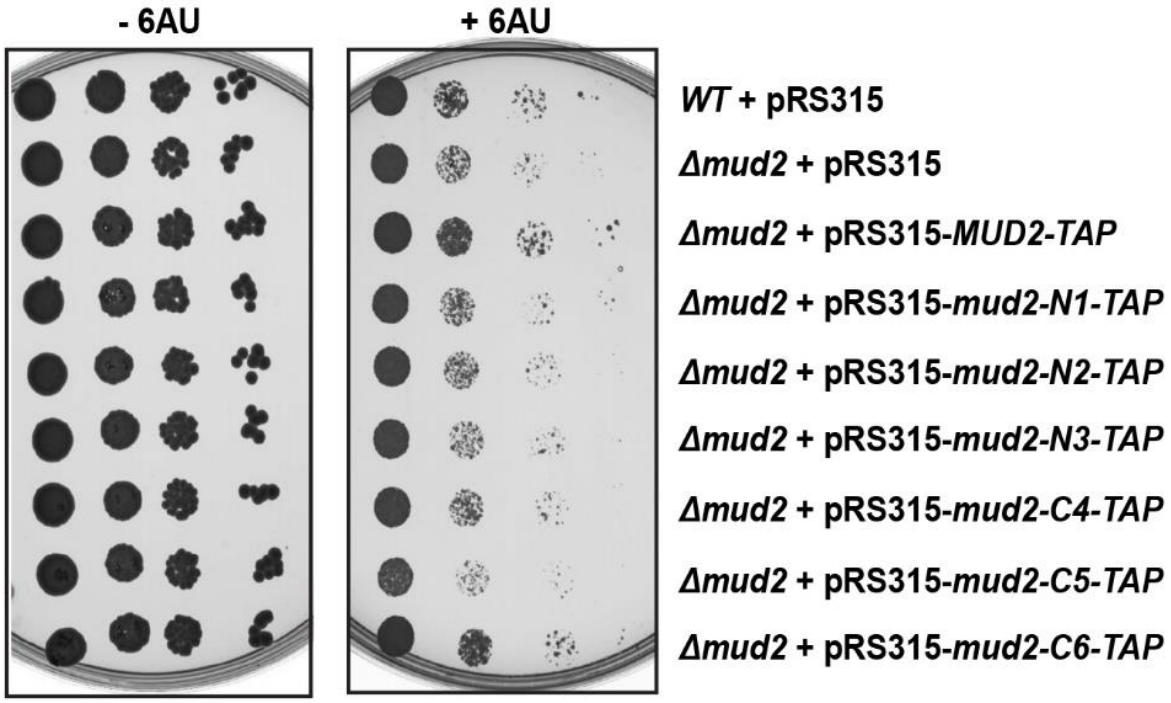


Figure 34. Only the N-terminal domain of *MUD2* can be dispensed for its function.
Dot spot assays of wild-type and *Δmud2* cells carrying the empty plasmid (pRS315) and the *Δmud2* cells carrying the plasmid which expresses either the C-terminally TAP-tagged full length Mud2 or the C-terminally TAP-tagged truncated versions of Mud2 protein (as shown in Figure 30, lower panel) on SDC-ura plates containing DMSO (as solvent control) or 50 µg/ml of 6AU. The plates were incubated at 30°C for 2-3 days.

To test whether all the six deletion mutants of *MUD2* are expressed at the same levels as its full length version, quantitative Western blots were performed with the *Δmud2* cells expressing the full length *MUD2* and *MUD2* truncated versions. All the Mud2 proteins truncated at their C-terminus (*mud2-N1-TAP*, *mud2-N2-TAP and mud2-N3-TAP*) were more likely expressed as the full length protein. On the other hand, among the N-terminally *MUD2* truncated proteins, *mud2-C4-TAP* or *mud2-C6-TAP* were highly expressed. The expression levels of Mud2 were quantified and normalized to the corresponding Pgk1 signal (Figure 35).

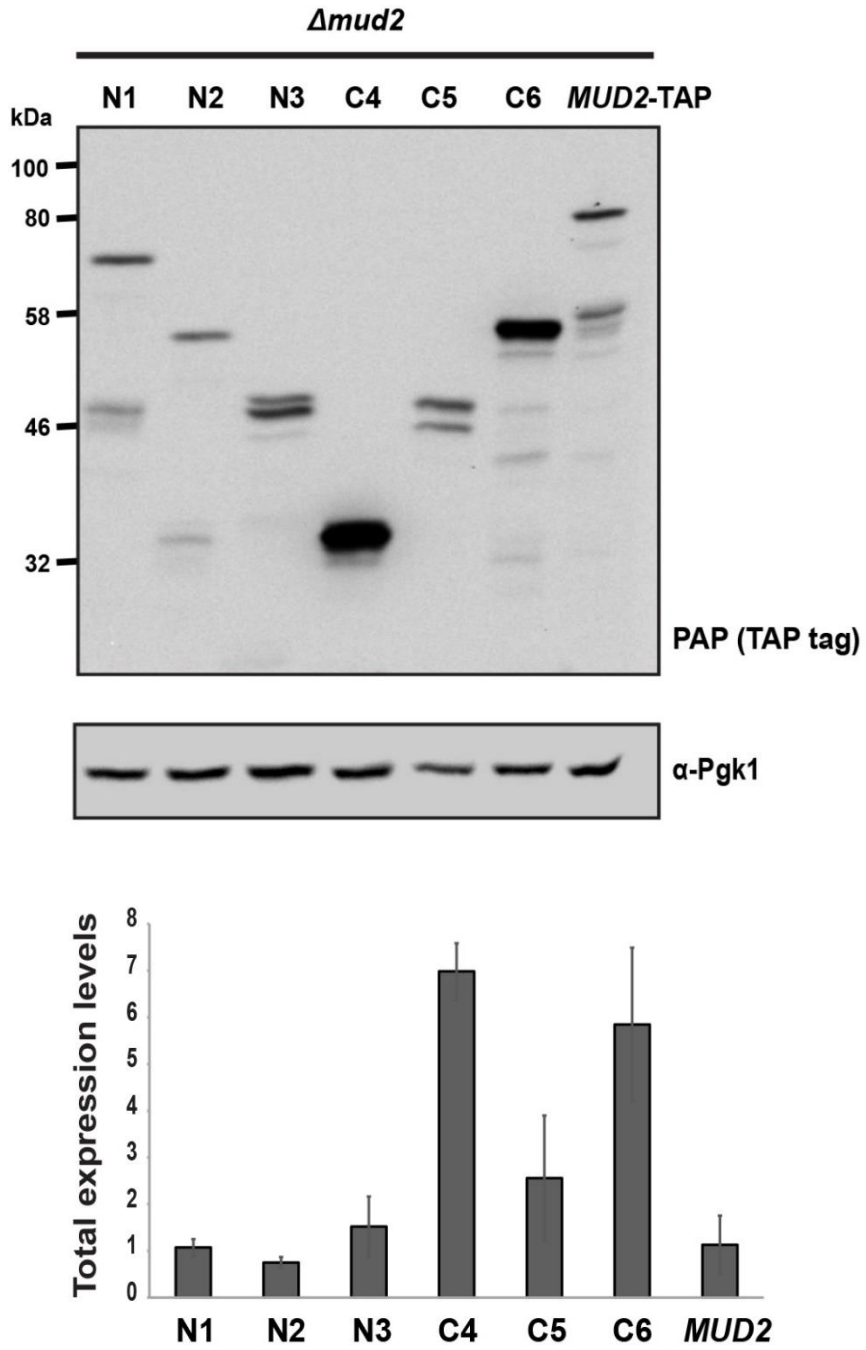


Figure 35. Total cellular levels of C-terminally TAP-tagged full length Mud2 and six truncated versions of Mud2 expressed in *Δmud2* cells.

Total cellular levels of TAP-tagged proteins were determined by Western blotting using a Protein A specific PAP (peroxidase-anti-peroxidase soluble complex) reagent (upper panel). Pgk1 levels served as a loading control. The quantification of three independent experiments was performed with LabImage 1D software. The quantified Mud2 signal normalized to the Pgk1 signal is shown in the lower panel.

3.1.11 Mud2 functions in transcription

To provide a direct evidence for the role of Mud2 in transcription, *in vivo* and *in vitro* transcription assays were performed. For *in vivo* transcription assays, two reporter constructs were used based on galactose inducible *GAL10* promoter. The transcription of an intron-containing *ACT1* gene (on plasmid) and an intronless endogenous *GAL10* gene was assessed by inducing their expression with galactose in the medium. For *ACT1* expression, a plasmid containing *GAL10* promoter in front of *ACT1* gene was transformed into cells. RNA was extracted from wild-type and *Δmud2* cells before induction (as a control) and 5', 10', 20' and 30' after galactose induction. An RNAPIII transcript SCR1 RNA served as a loading control. The amount of *GAL10* and *ACT1* mRNA was quantified using primer extension assay as described in Section 2.2.6.1. The levels of both synthesized *ACT1* and *GAL10* mRNA were decreased significantly in *Δmud2* cells as compared to the wild-type cells (Figure 36 and 37). This shows that Mud2 functions in transcription *in vivo*.

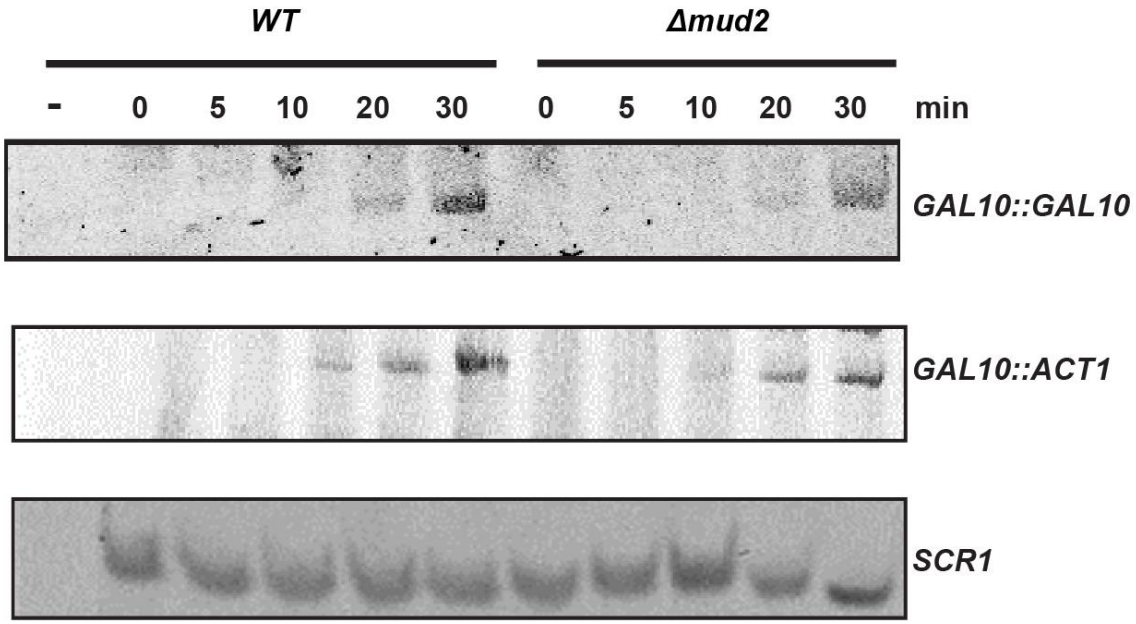


Figure 36. Mud2 is necessary for efficient mRNA synthesis *in vivo.*
The intronless, endogenous *GAL10* gene and plasmid-driven *ACT1* gene under the control of *GAL10* promoter were induced in the -galactose medium and total RNA was extracted from the cells collected 5, 10, 20 and 30 min after induction. The amount of *GAL10* and *ACT1* mRNA synthesized was analyzed by primer extension assay using 5' Cy5 labelled primers specific to *GAL10* and *ACT1* mRNA. The amount of *SCR1* RNA (-), an RNAPIII transcript, served as a loading control for the experiment.

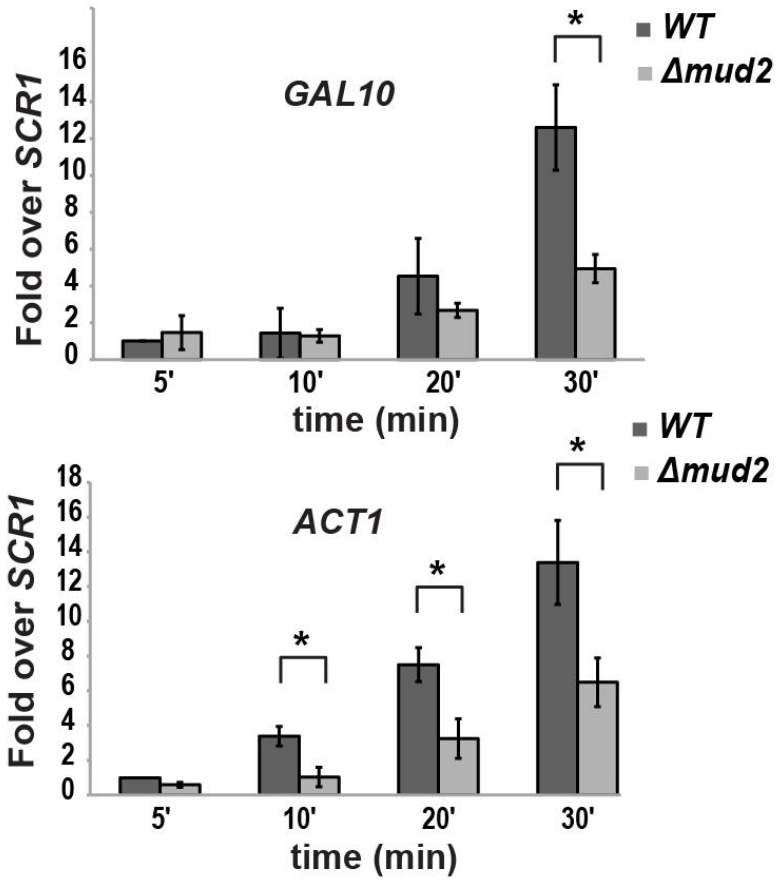


Figure 37. Mud2 is necessary for efficient mRNA synthesis *in vivo* (quantification).
The quantification of three independent *in vivo* transcription assays (a representative of which is shown in Figure 36) was performed using ImageQuant software. The amount of *GAL10* and *ACT1* mRNA synthesized 5, 10, 20 and 30 min after galactose induction in wild-type (*WT*) and *Δmud2* cells was quantified and normalized to the corresponding amount of SCR1 RNA, an RNAPIII transcript. The normalized signal for wild-type cells 5 min after galactose induction was set equal to 1.

In order to confirm that decrease in transcriptional activity in *Δmud2* cells is a direct effect, *in vitro* transcription assays with wild-type and *Δmud2* mutant cells were performed. To achieve this, a plasmid-based *in vitro* transcription assay with a *S. cerevisiae HIS4* promoter was performed. *HIS4* is involved in histidine biosynthesis and the transcription of *HIS4* is regulated by a eukaryotic transcriptional activator protein Gcn4 (Lucchini, Hinnebusch et al. 1984, Hope and Struhl 1987). A pBluescript II KS+ plasmid (LL279) containing a wild-type *HIS4* yeast promoter sequence (-428 to +24 respective to the start codon) served as the template for the assay. (Seizl, Lariviere et al. 2011). Recombinant Gcn4 was used to activate the transcription reaction. *In vitro* transcription reactions using nuclear extracts from wild-type and *Δmud2* cells were prepared. *In vitro* generated transcripts were analyzed by primer extension assay using a Cy5 labeled primer for the reverse transcription. At the optimal temperature for *in vitro* transcription (18°C), transcriptional activity of the nuclear extracts prepared from *Δmud2* cells

was found to be 40-50 % lower compared to the transcriptional activity of the nuclear extracts from the wild-type cells. When temperature of the transcription reaction was increased to 23°C, the decrease in the transcriptional activity in the nuclear extracts from *Δmud2* cells relative to the wild-type cells became more extrusive and conspicuous (Figure 38).

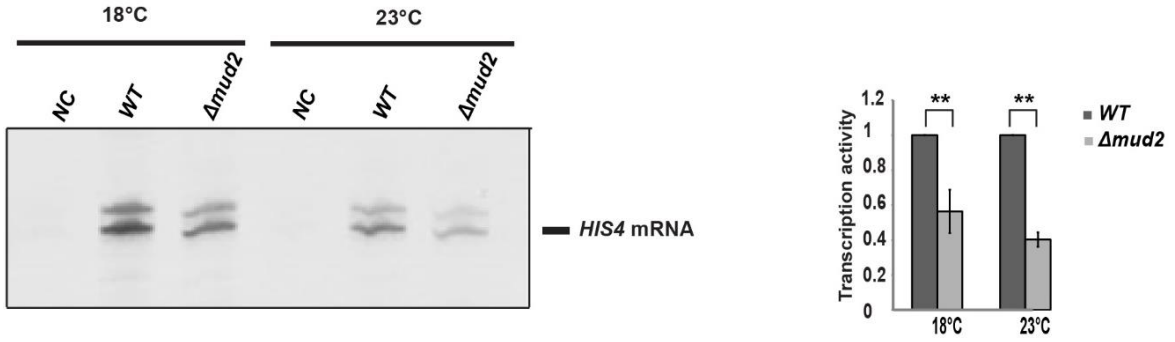


Figure 38. Mud2 is required for efficient *in vitro* synthesis of mRNA.
A plasmid based *in vitro* transcription assay was performed with the nuclear extracts of wild-type (WT) and *Δmud2* cells at 18°C and 23°C (left). The plasmid used as a template for the transcription reaction contained a yeast *HIS4* promoter sequence. The negative control (NC) lacked Gcn4, a transcriptional activator. The quantification of three independent experiments is shown (right).

To corroborate that this decrease in the transcriptional activity is caused by the deletion of *MUD2*, an add-back assay was performed to determine if adding back purified Mud2 protein to the nuclear extracts from *Δmud2* cells can reconstitute the lost transcriptional activity. To achieve this, His tagged Mud2 was purified from *E. coli* and TAP tagged Mud2 was purified from *S. cerevisiae* at high salt concentrations (for details, see Section 2.2.3; Figure 39, A). An increasing concentration of purified Mud2 protein from both the sources was added to the *in vitro* transcription reactions. The addition of Mud2 rescued the decrease in the transcriptional activity of *Δmud2* extracts in a dose dependent manner in both cases (Figure 39, B and C). Furthermore, RNAPII was not associated with Mud2 purified from *S. cerevisiae* (Figure 39D). These results confirm that Mud2 plays a direct role in RNAPII transcription.

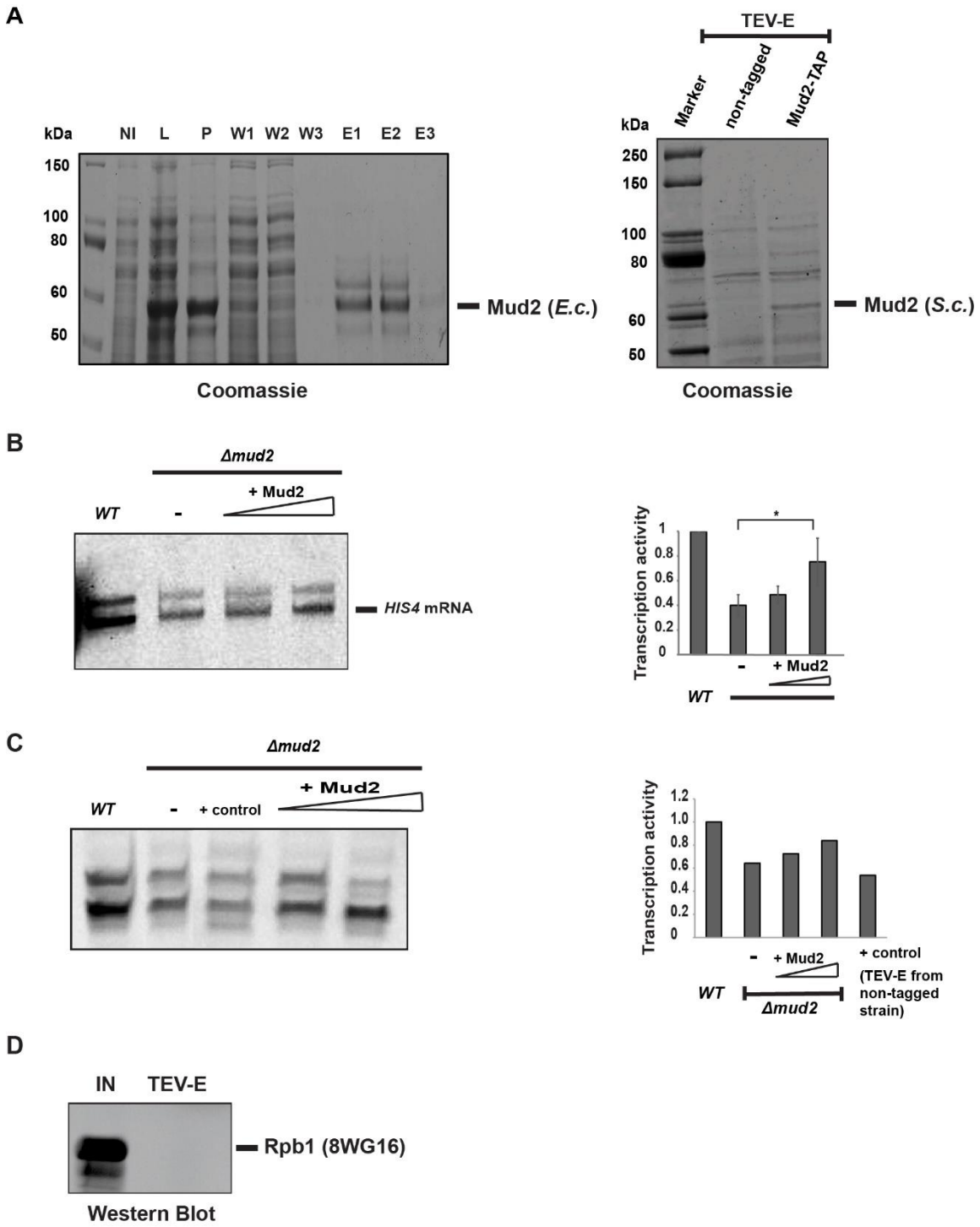


Figure 39. Add back of purified Mud2 stimulates the transcriptional activity of *Δmud2* extracts.
(A, left) Recombinant Mud2 purified for the add-back assay shown in (B). NI: non-induced; L: lysate; P: pellet; W1-W3: flow-through after three column washing steps; E1-E3: eluted Mud2 fractions. **(A, right)** TEV-eluates of a non-tagged strain and a TAP-tagged Mud2 strain purified from *S. cerevisiae* under high salt conditions (250mM NaCl). The purified Mud2 was used for the add-back assay shown in B and C. **(B)** Add back of 50 and 300 ng of recombinant Mud2 increases the activity of *in vitro* transcription reaction from *Δmud2* extracts. Quantification of three independent *in vitro* transcription assays is shown (right). **(C)** Add back of 50 and 150 ng of yeast purified Mud2 increases the activity of *in vitro* transcription reaction from *Δmud2* extracts. Quantification of one representative *in vitro*

transcription assay is shown (right). **(D)** RNAPII does not associate with yeast purified Mud2 (purification shown in A, right panel) as assessed by Western Blotting using Rpb1 antibody 8WG16.

3.1.12 *Δmud2* is synthetically lethal with *syf1-37*

Prp19C and TREX occupancy is reduced in *Δmud2* cells (Figure 21 and 22). Also, Prp19C and TREX occupancy is reduced in *syf1-37* cells (Chanarat, Seizl et al. 2011). This provides a possibility that Mud2 and Prp19C interact with the transcriptional machinery via two separate modes. If that is the case, combination of *MUD2* deletion with *syf1-37* mutation should show more pronounced genetic effects than observed in either of *Δmud2* or *syf1-37* cells. Indeed, it was observed that *Δmud2* is synthetically lethal with *syf1-37* (Figure 40). Thus, Mud2 and Prp19C play important and independent roles in maintaining TREX at transcribed genes.

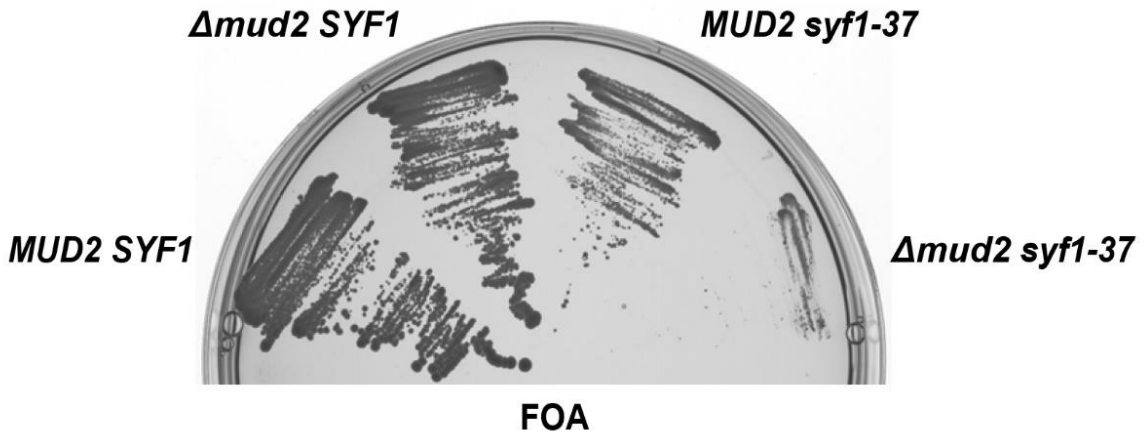


Figure 40. *Δmud2* is synthetically lethal with *syf1-37*.
A *Δmud2Δsyf1* strain complemented by full length Syf1 protein encoded on the plasmid pRS316-*SYF1* was transformed with plasmids encoding full length or C-terminally truncated *SYF1* mutant (pRS315-*SYF1 or* pRS315-*syf1-37*) and plasmids encoding full length *MUD2* or an empty plasmid (pRS314-*MUD2 or* pRS314). The corresponding four strains were streaked on agar plates containing FOA. The plates were incubated at 30°C for 3 days.

3.1.13 Mud2 is not required for mRNA export

It has been reported previously that TREX functions in mRNA export while Prp19C is not required for the same (Strasser, Masuda et al. 2002, Chanarat, Seizl et al. 2011). Thus, it is interesting to investigate whether Mud2 is required for mRNA export. To test this possibility, oligo-dT *in situ* hybridization was performed with the wild-type and *Δmud2* strains. As a positive control, *mex67-5* mutant cells, which show an mRNA export defect at 37°C, were used (Segref, Sharma et al. 1997). While *mex67-5* cells showed nuclear accumulation of mRNA at 37°C, *Δmud2* cells did not show a defect in the mRNA export (Figure 41). Thus, in this respect, Mud2 shows the same behavior as Prp19C.

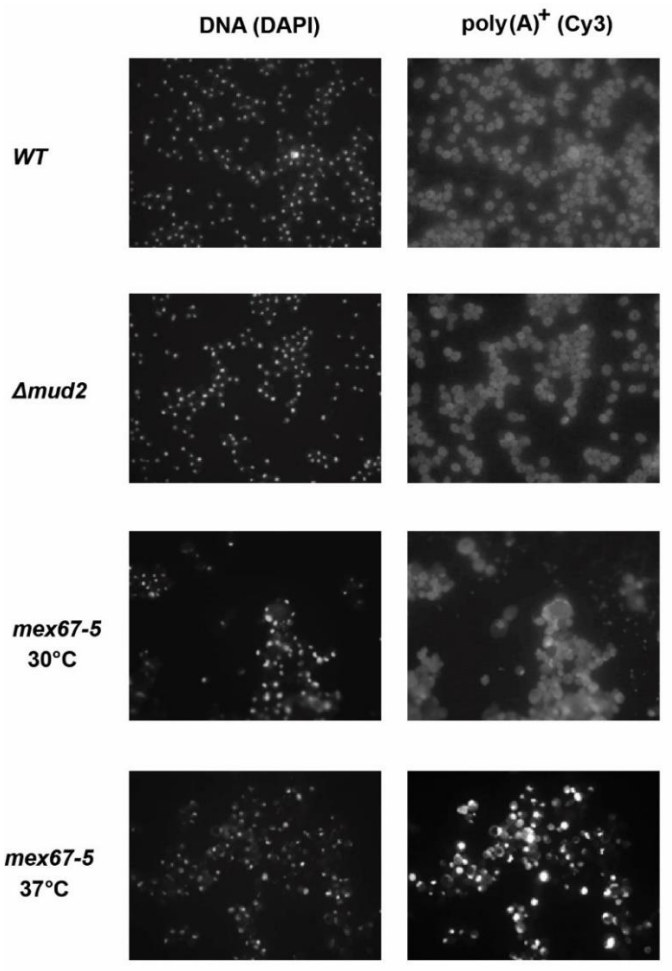


Figure 41. *Δmud2* cells do not show an mRNA export defect.
The wild-type and *Δmud2* cells were grown at 30°C. The *mex67-5* mutant cells, which are known to show a strong export defect at elevated temperatures, were grown at 30°C and shifted to 37°C for 15 min before adding formaldehyde. Oligo-dT *in situ* hybridization was performed as explained in Section 2.2.9. The nuclei were stained with DAPI (4',6-diamidino-2-phenylindole) while the poly(A)-tail of bulk mRNA was detected by hybridization with a Cy3-labelled oligo $d(T)_{50}$ probe.

3.1.14 Purification of Mud2 from *MUD2- FTpA*

During purification of a TAP-tagged Mud2, more copurifying proteins were observed in the Coomassie stained polyacrylamide gel when the lysates were treated with RNase A as compared to the non-RNase A treated lysates (Figure 19). In order to re-confirm this result and to identify the Mud2 copurifying proteins, purification of Mud2 from *S. cerevisiae* was performed with a modified TAP tag (containing a FLAG tag instead of Calmodulin binding peptide). The strain was made using the template plasmid pFA6a-FTpA-HisMX6 (Kornprobst, Turk et al. 2016). The purification was performed as the standard tandem affinity purification, except that the FLAG beads were used instead of Calmodulin Sepharose beads. The purification was performed in the absence as well as presence of RNase A and the Coomassie stained bands in the polyacrylamide gel were analyzed by mass spectrometry. The proteins identified are marked (Figure 42A). It was confirmed that, Mud2 copurified more proteins when the lysate was treated with RNase A compared to the non-RNase A treated lysates. This might be due to the reason that RNA somehow abrogates the interaction of Mud2 with some other proteins. The blot against mRNA binding proteins Nab2 and Npl3 were also performed with the TEV-eluates in order to confirm that RNase treatment worked (Figure 42B).

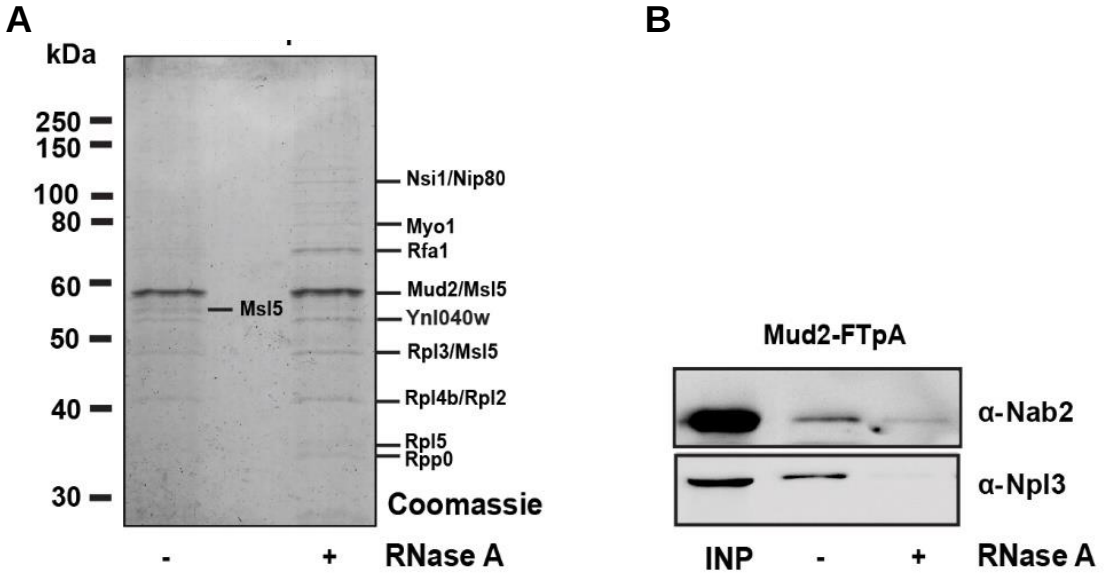


Figure 42. Purification of Mud2 from *MUD2- FTpA* strain.
(A) Mud2 was purified from a strain containing a C-terminal FTpA (FLAG-TEV-protein A) tagged Mud2 in *S. cerevisiae.* The purification was performed in the absence and presence of RNase A and the co-purifying proteins (indicated with a solid black line) were assessed by mass spectrometry. **(B)** The treatment of total cell extracts with RNase A abolishes the copurification of mRNA binding proteins Nab2 and Npl3, as assessed by Western blotting of the TEV-eluates using Nab2 and Npl3 antibody.

3.2 Role of Prp19C subunits in Prp19C-TREX and Prp19C-Mud2 interaction

3.2.1 Prp19C binds directly to the THO complex *in vitro* in an RNA-independent manner

It has been shown previously that Prp19C interacts with the TREX complex *in vivo* and this interaction is independent of RNA (Chanarat, Seizl et al. 2011). In order to determine whether this interaction is direct, *in vitro* binding assays with TAP-purified Syf1 (Prp19C) and the THO complex bound to IgG-coupled Sepharose beads (from *HPR1-TAP* strain) were performed. The details of this assay are explained in Section 2.2.8.

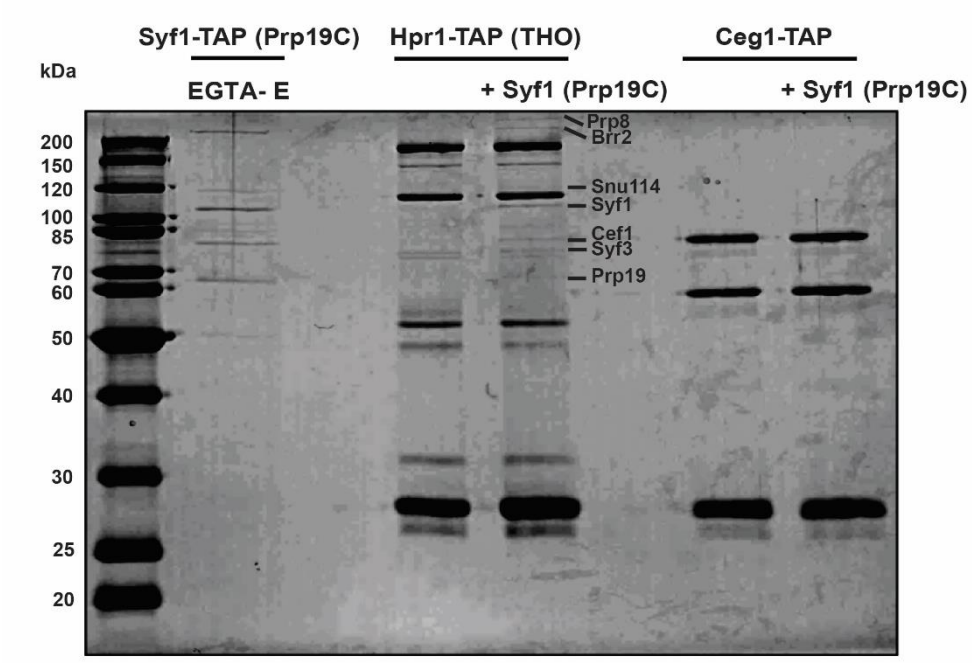


Figure 43. Prp19C binds directly to the THO complex in an RNA-independent manner.
SYF1-TAP strain was purified by TAP till EGTA elution (eluates marked as EGTA-E) and concentrated. Equal amounts were then incubated with Hpr1-TAP and Ceg1-TAP (a negative control) bound to IgG-coupled Sepharose beads washed under high salt conditions (in order to obtain THO complex). The total cell lysates from Hpr1-TAP and Ceg1-TAP strains were treated with RNase A prior to the purification. After incubation, the beads were washed and the bound proteins were eluted by treatment with TEV protease. The eluted proteins were loaded on a 12% SDS-polyacrylamide gel stained with Coomassie brilliant blue. Prp19C subunits that interact with the THO complex are indicated.

This assay revealed that Prp19C binds directly to the THO complex in an RNA-independent manner. In contrast, no interaction was observed between Prp19C and Ceg1 (a subunit of mRNA capping enzyme, which served as a negative control for the experiment). The Prp19C subunits interacting with the THO complex were identified on a Coomassie stained polyacrylamide gel (Figure 43). Thus, it can be concluded that Syf1 and most likely the whole Prp19C binds directly and specifically to the THO complex in an RNA-independent manner.

3.2.2 Role of non-essential Prp19C subunits in Prp19C stability at the genes

Although non-essential components of Prp19C has been studied in splicing (Chan, Kao et al. 2003, Chan and Cheng 2005, Chen, Kao et al. 2006), their functions in the transcription and interaction with the TREX complex remain elusive. In this regard, the first objective was to determine if deletion of any of these components affects the occupancy of the Prp19C at transcribed genes. In order to test this, the occupancy of TAP tagged Syf1 was analyzed by ChIP in the wild-type (*WT*) and Δ*cwc15*, Δ*isy1*, Δ*ntc20* and Δ*syf2* mutant cells. ChIP results were assessed by Real Time PCR using primer pairs which amplify 3' positions of four highly expressed intron-containing genes *DBP2*, *ACT1*, *RPS14B* and *RPL28* and four highly expressed intronless genes *PMA1*, *ILV5*, *CCW12* and *RPL9B*. It is noteworthy to mention here that Δ*snt309* mutants were not included in this analysis, because of the fact that in *SNT309* deleted cells, Prp19C is destabilized and associated proteins are dissociated from the core component Prp19, as published before and reassessed by purifying TAP-tagged Syf1 (Prp19C) from Δ*snt309* cells (where the core components of Prp19C could not be purified; results not shown).

The ChIP results showed that the occupancy of Syf1 and most likely the whole Prp19C decreases at intron-containing genes in Δ*isy1*, Δ*ntc20* and Δ*syf2* mutant cells as compared to the wild-type cells (Figure 44). In contrast to the intron-containing genes, the Prp19C occupancy decreases specifically in Δ*isy1* mutants at the intronless genes. Moreover, no significant decrease of Prp19C occupancy in Δ*cwc15* cells at intron-containing or intronless genes was observed (except at *RPS14B*). Thus, the occupancy of Syf1 and most likely the whole Prp19C decreases specifically in Δ*isy1* mutants at both intron-containing and intronless genes. This could also affect the occupancy of TREX as Prp19C is required to maintain TREX.

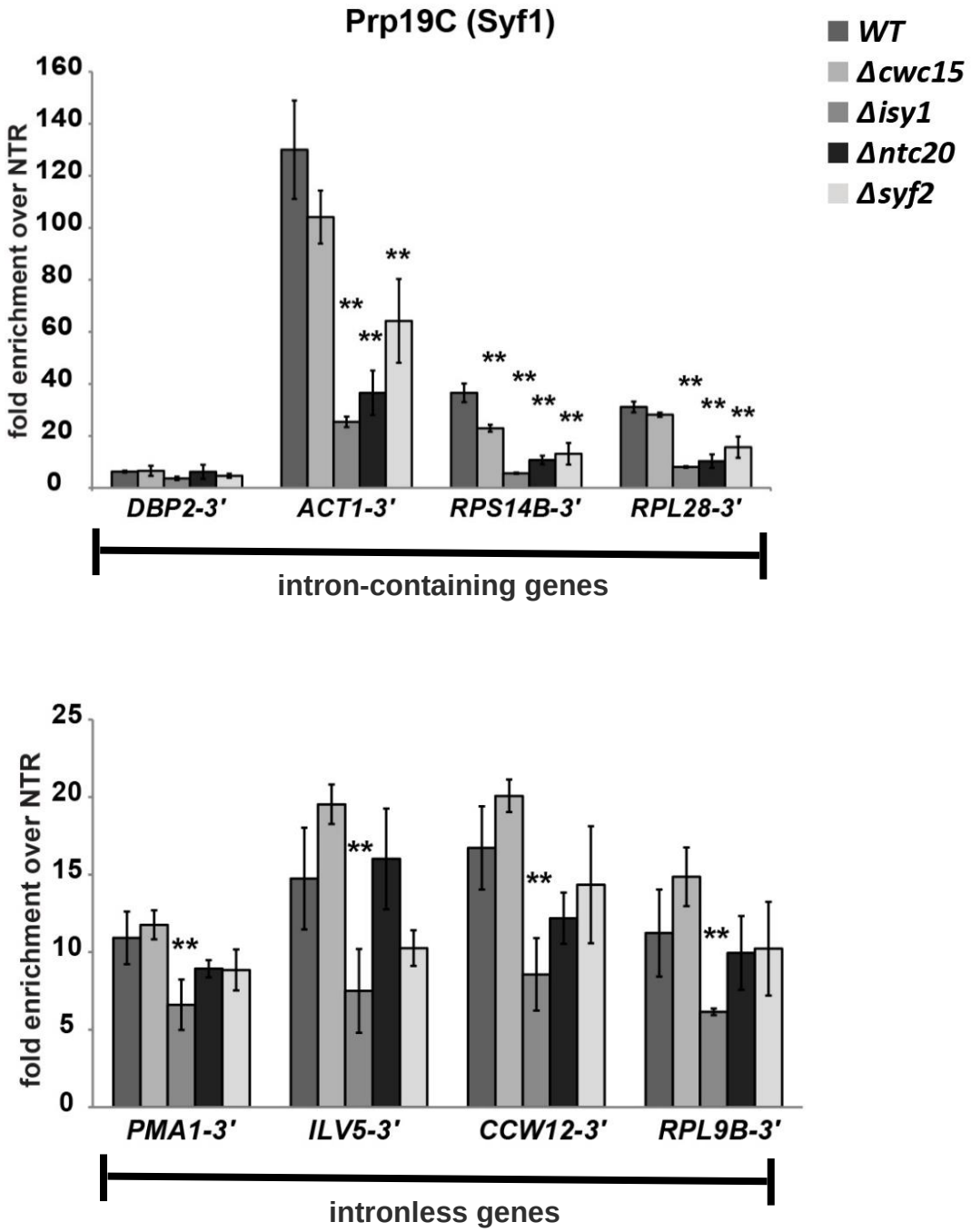


Figure 44. Prp19C occupancy decreases in *Δisy1*, *Δsyf2* and *Δntc20* mutant cells at intron-containing genes and specifically in *Δisy1* mutant at intronless genes.
The occupancy of Syf1 component of Prp19C was assessed in wild-type cells *(WT)* and in Prp19C deletion mutants using ChIP at both intron-containing and intronless genes. ChIPs were quantified by Real Time PCR using the primer pairs as shown in Figure 10.

In order to investigate whether the occupancy of TREX changes in Δ*isy1* mutant cells, ChIP experiments were performed with an endogenously TAP-tagged Hpr1 component of TREX in wild-type cells and the four deletion mutant cells. It was observed that the occupancy of Hpr1

and most likely the whole TREX complex decreases specifically in Δ*isy1* mutants at both intron-containing and intronless genes, although a decrease in TREX occupancy was also observed in Δ*ntc20* mutant cells at certain gene positions (Figure 45).

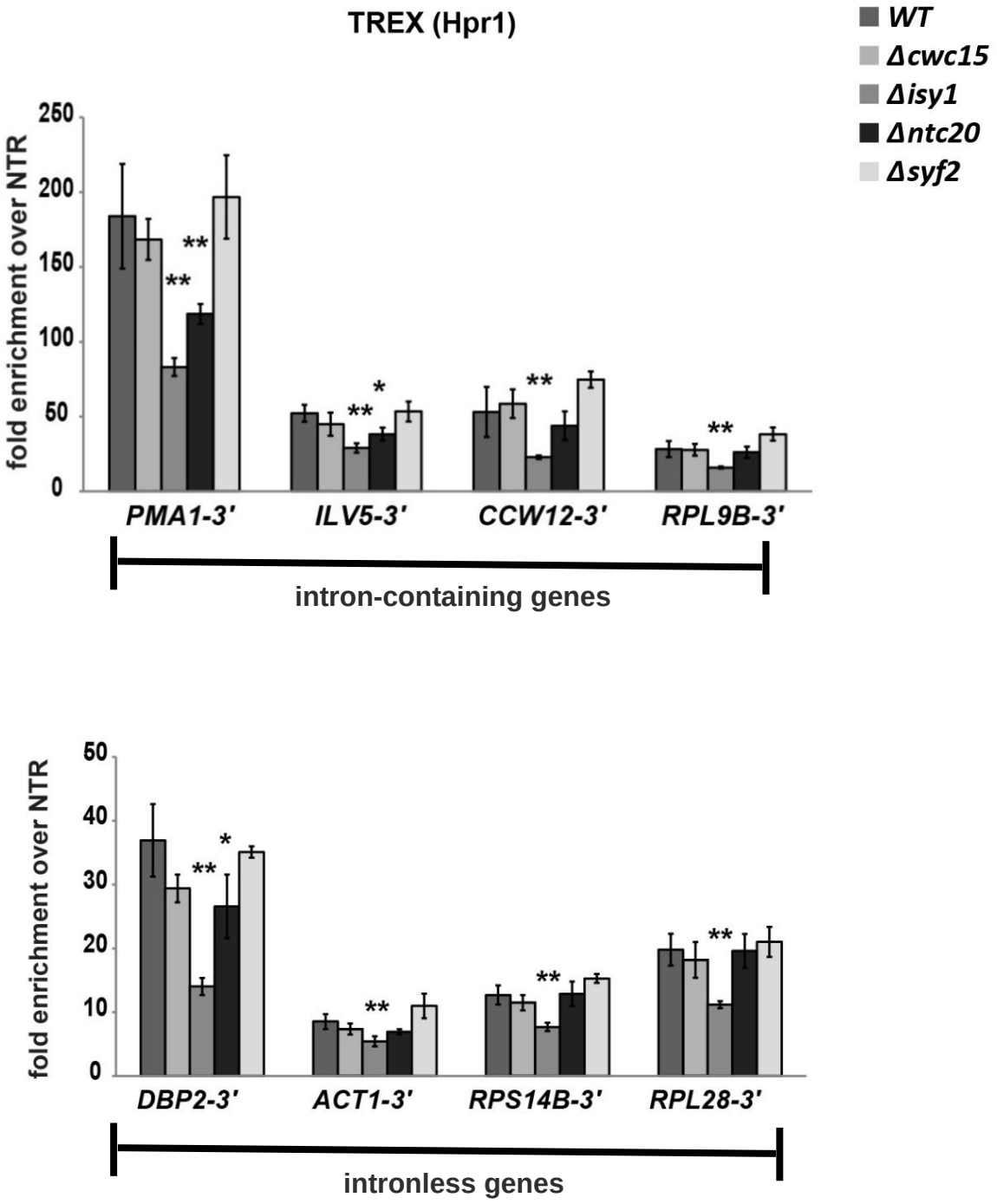


Figure 45. The occupancy of Hpr1 (TREX) decreases in the *Δisy1* mutant cells at both intron-containing and intronless genes.
The occupancy of Hpr1 was assessed in wild-type cells *(WT)* and in Prp19C deletion mutants using ChIP at intron-containing and intronless genes. ChIPs were quantified by Real Time PCR using the primer pairs as shown in Figure 10.

In order to determine if the total cellular levels of Hpr1 and Syf1 change in Δ*cwc15*, Δ*isy1*, Δ*ntc20* and Δ*syf2* mutant cells compared to the wild-type cells, Western blots were performed with the strains used for above ChIP experiments. The TAP-tagged proteins were detected using a Protein A specific PAP (peroxidase-anti-peroxidase soluble complex) reagent (Figure 46). Interestingly, it was observed that total cellular levels of both Hpr1 and Syf1 decrease in Δ*cwc15* and Δ*syf2* mutants, while they remain constant in Δ*isy1* and Δ*ntc20* mutant cells. This indicates that observed decrease of Prp19C and TREX occupancy in Δ*isy1* cells is not caused due to the change in total cellular levels of these complexes. Strikingly, decreased cellular levels of Hpr1 and Syf1 in Δ*cwc15* and Δ*syf2* mutants does not affect their occupancies at genes. Nevertheless, it can be concluded that Isy1 component of Prp19C is required for maintaining the occupancy of both Prp19C and TREX at transcribed genes.

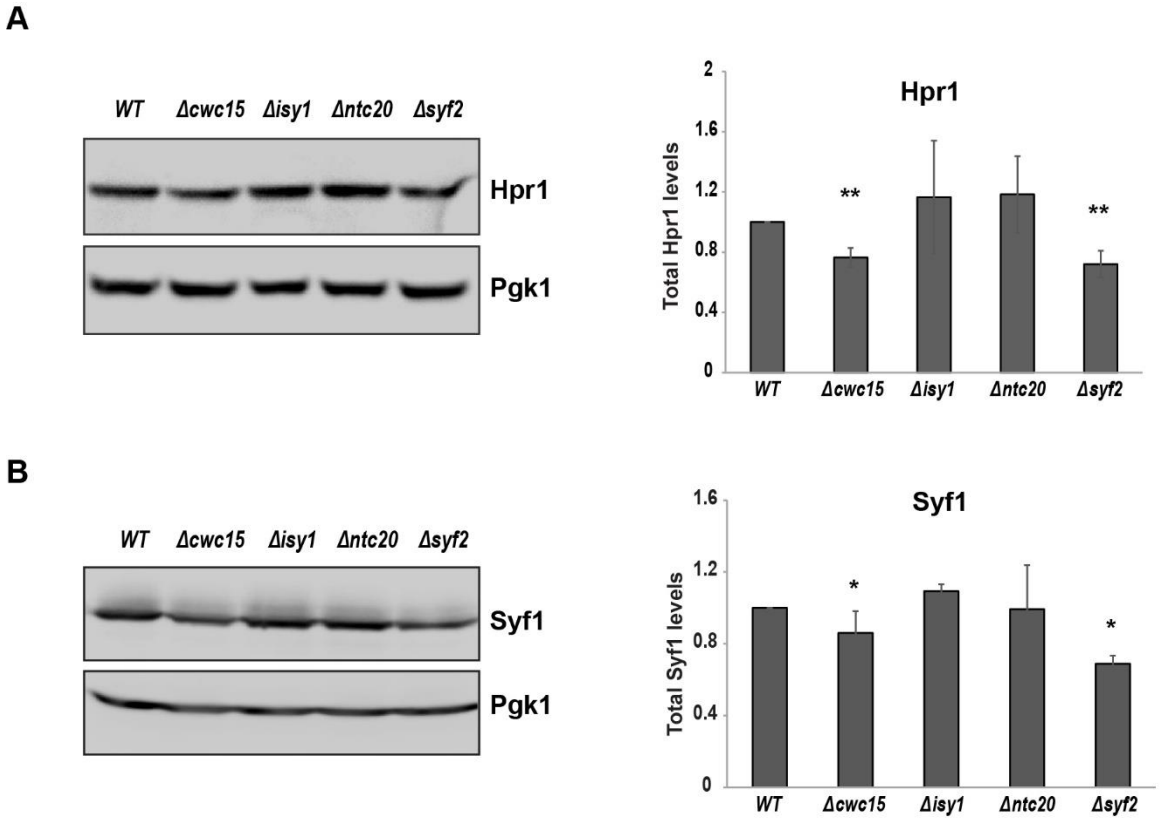


Figure 46. Total cellular protein levels of Hpr1 and Syf1 in wild-type (*WT*) and *Δcwc15*, *Δisy1*, *Δsyf2* and *Δntc20* mutant cells.

Western blots analysis was performed to determine the total cellular levels of Hpr1 and Syf1 in wild-type (*WT*) and *Δcwc15*, *Δisy1*, *Δsyf2* and *Δntc20* mutant cells (A and B left panels). Pgk1 levels served as the loading control for the analysis. Quantification of the Western blots was performed with LabImage 1D software and is shown in the corresponding right panels.

To further confirm that the observed effect is not because of a general decrease in transcription, Rpb1 occupancy was also assessed in these deletion mutants at both intron-containing and intronless genes. Interestingly, RNAPII occupancy as assessed by ChIP using the Rpb1 antibody 8WG16 increases significantly at all the gene positions in Δ*isy1* and Δ*syf2* mutants while it does not change in other deletion mutants (except in Δ*cwc15* mutant cells at two positions) (Figure 47).

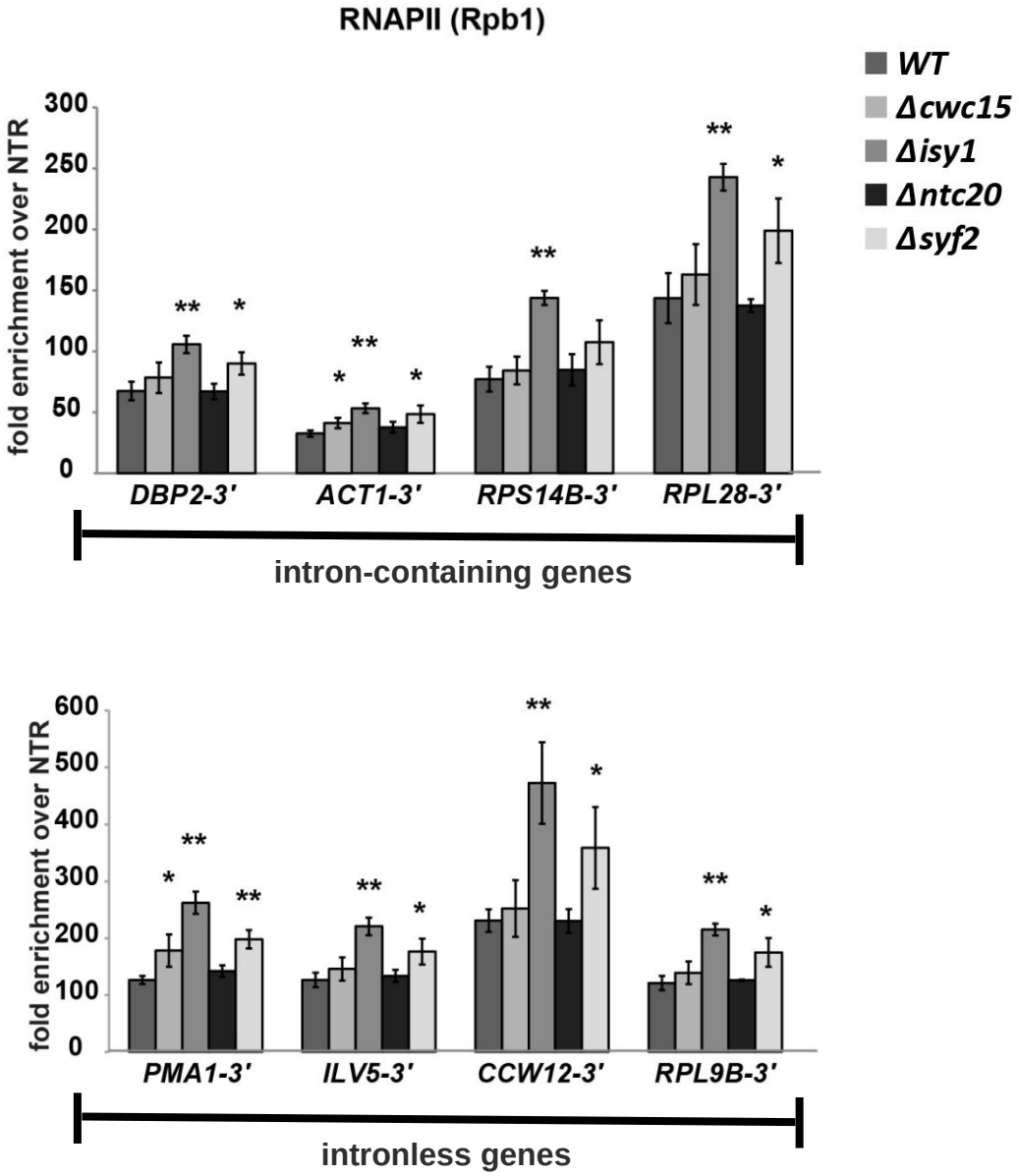


Figure 47. RNAPII occupancy increases in Δ*isy1* and *Δsyf2* mutant cells at both intron-containing and intronless genes.

The occupancy of Rpb1 component of RNAPII was assessed in wild-type cells *(WT)* and in Prp19C deletion mutants using chromatin immunoprecipitation (ChIP) at both intron-containing and intronless genes. ChIPs were quantified by Real Time PCR using the primer pairs as shown in Figure 10.

This suggests that the deletion of Isy1 and Syf2 can recruit even more RNAPII at the genes but the increased amount of RNAPII is still not sufficient to fully maintain Prp19C and TREX at transcribed genes. To understand the reason behind this, more mechanistic insights into how these different components are arranged in the complex and how they are affecting the interaction between Prp19C and TREX are required.

3.2.3 Role of Prp19C non-essential components in Prp19C-TREX interaction *in vivo*

Prp19C interacts with the TREX complex *in vivo* and *in vitro*. However, details of how these two complexes interact and specifically, which subunits from each complex are involved, are still not understood. In order to get more insights into the mechanism of Prp19C-TREX interaction, it was investigated whether deletion of the non-essential components of Prp19C affects its interaction with TREX. In order to assess this, the wild-type as well as the four deletion mutant strains, all expressing *HPR1-TAP* with HA-tagged Syf1 (*SYF1*-HA), were prepared. A strain expressing only *SYF1-HA* (in wild-type cells) served as a negative control for the experiment. Hpr1 was purified by TAP. The copurification of Prp19C was assessed by Western blot analysis using an α-HA antibody which detects the endogenous HA-tagged Syf1. Figure 48 (A, upper panel) shows the EGTA-eluates of the purified proteins in a polyacrylamide gel stained with Coomassie brilliant blue and the corresponding α-HA blots of input and immunoprecipitated samples are depicted in Figure 48 (A, lower panel). As shown in Figure, the TREX complex could be purified from all the deletion mutant strains analyzed. The α-HA signal detected in the Western blot and the signal of Hpr1 in Coomassie stained polyacrylamide gel were quantified. Western blot signal was normalized to the corresponding Hpr1 signal in order to compensate for differences in the purifications from different strains.

As assessed from the normalized blot signals (Figure 48B), the interaction between the TREX and Prp19C decreases in Δ*cwc15*, Δ*ntc20* and Δ*syf2* and increases in Δ*isy1* mutant cells. Interestingly, an extra band was observed in the purified TREX complex from Δ*cwc15* and Δ*syf2* mutant cells (marked by an asterisk). As analyzed by mass spectrometry, it is the degraded Hpr1 component of the complex. Unfortunately, it was difficult to determine if this degradation of Hpr1 is due to a C-terminal truncation or N-terminal truncation. Also, there might be a

possibility that tagging at the C-terminal in these mutants makes the protein unstable and leads to its degradation.

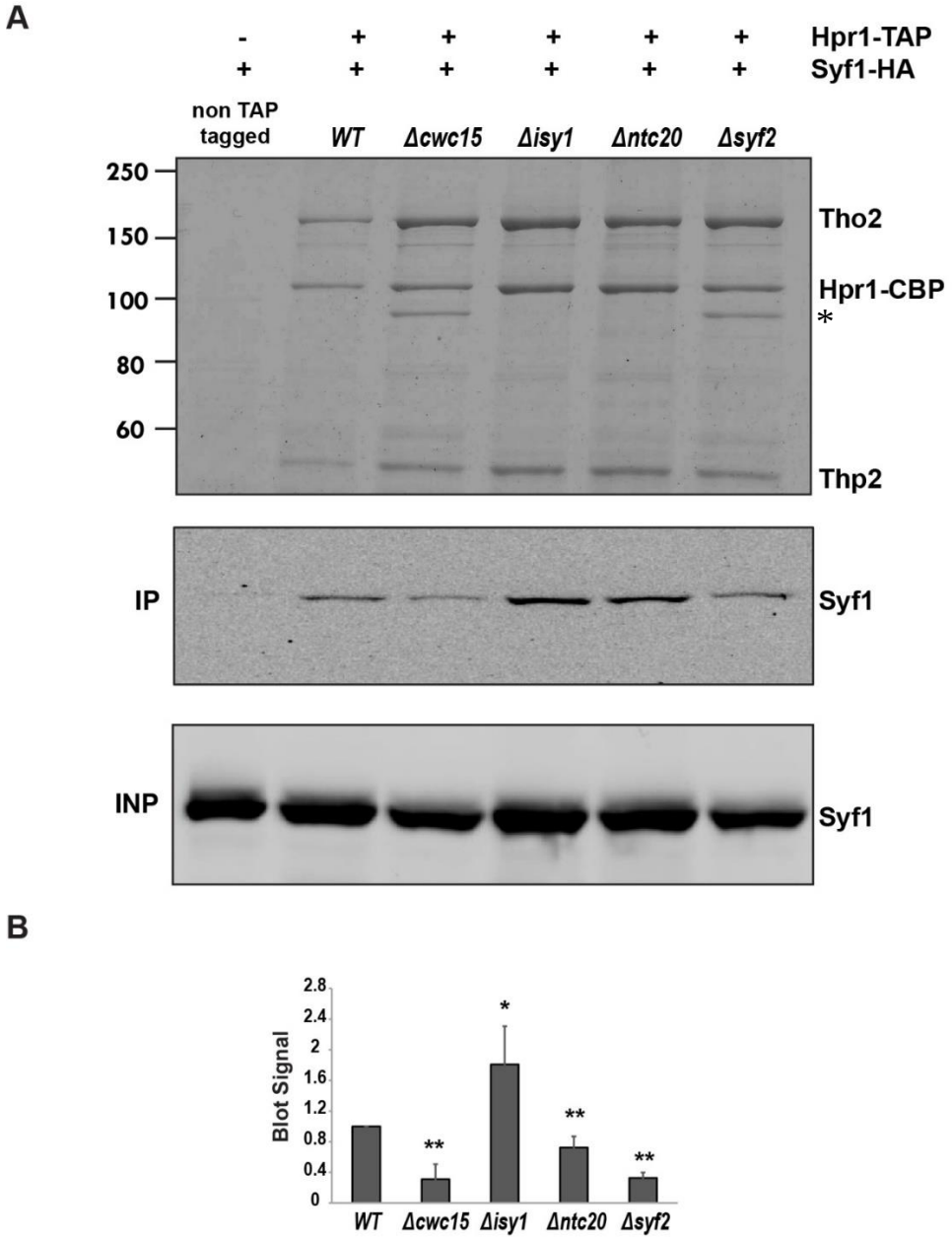


Figure 48. Interaction between the TREX and Prp19C decreases in *Δcwc15*, *Δntc20* and *Δsyf2* and increases in *Δisy1* mutant cells *in vivo.*

(A) Upper panel shows TAP purified EGTA-eluates from the cells expressing *HPR1-TAP* with HA-tagged Syf1 in wild-type and *Δcwc15*, *Δisy1*, *Δntc20* and *Δsyf2* mutants. A strain expressing only *SYF1-HA* served as a negative control. Lower panel shows western blot analysis of Input (INP; whole cell extract) and purified EGTA-eluates (IP; Immunoprecipitated) from the above purifications. **(B)** The quantification of three independent experiments was performed with LabImage 1D software.

In order to test whether the Hpr1 is degraded when TREX is purified with respect to its N-terminal, Hpr1 was TAP-tagged at its N-terminus (*TAP-HPR1*) in the wild-type as well as Δ*cwc15* and Δ*syf2* mutant cells. TREX was purified from these strains and no degradation of Hpr1 was observed in any of the strains (Figure 49A). Also, the expression of Hpr1 did not decrease in Δ*cwc15* and Δ*syf2* mutant cells containing an N-terminally tagged Hpr1 (Figure 49B). To understand the details behind this, it is important to further study the mechanistic details of this interaction.

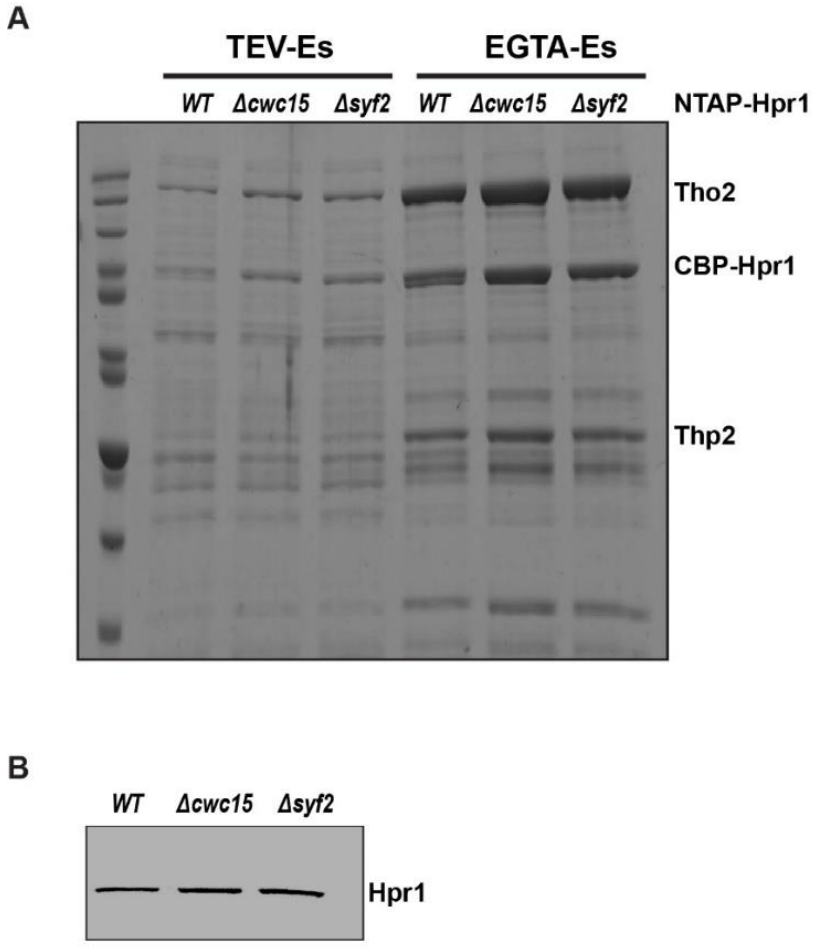


Figure 49. Purification of TREX from N-terminally tagged Hpr1 in wild-type and *Δcwc15* and *Δsyf2* mutant cells.

(A) TEV-eluates (TEV-Es) and EGTA-eluates (EGTA-Es) of the TAP purification from the cells expressing N-terminally tagged Hpr1 in wild-type and *Δcwc15* and *Δsyf2* mutants are shown. (B) Western blots were performed to determine the total cellular levels of Hpr1 in wild-type (*WT*) and *Δcwc15* and *Δsyf2* mutant cells.

3.2.4 Role of Prp19C non-essential components in Prp19C-Mud2 and Prp19C-RNAPII interaction *in vivo*

In order to determine whether the interaction of Prp19C with Mud2 and the interaction of Prp19C with RNAPII is affected when any of the non-essential components of Prp19C including *CWC15* is deleted, Syf1 was C-terminally TAP tagged (*SYF1-TAP*) and Mud2 was HA-tagged (*MUD2-HA*) in the wild-type as well as in the Δ*cwc15*, Δ*isy1*, Δ*ntc20* and Δ*syf2* mutant cells. A strain expressing only *MUD2-HA* served as a negative control for the experiment. To analyze the interaction of Prp19C with Mud2, *SYF1-TAP* was purified from wild-type and mutant strains by TAP. The copurification of Mud2 was assessed by the Western blot analysis using an α-HA antibody which detects the endogenous HA-tagged Mud2. To analyze the interaction of Prp19C with RNAPII, the same strains were TAP purified except that the interaction with RNAPII was assessed using an Rpb1 antibody (8WG16).

Figure 50 (A, upper panel) shows the EGTA-eluates of purified Syf1 in wild-type and all the deletion mutants. The different subunits of the Prp19C could be identified in Coomassie stained polyacrylamide gel. Western blot analysis showed that deletion of any of these factors did not affect the interaction of Prp19C with Mud2 although the interaction of Prp19C with RNAPII decreases in Δ*cwc15* and Δ*syf2* mutants and increases in Δ*isy1* mutant cells (Figure 50A, lower panel and 50B). In conclusion, deletion of non-essential components of Prp19C does not affect Prp19C-Mud2 while differentially affects Prp19C-RNAPII interaction.

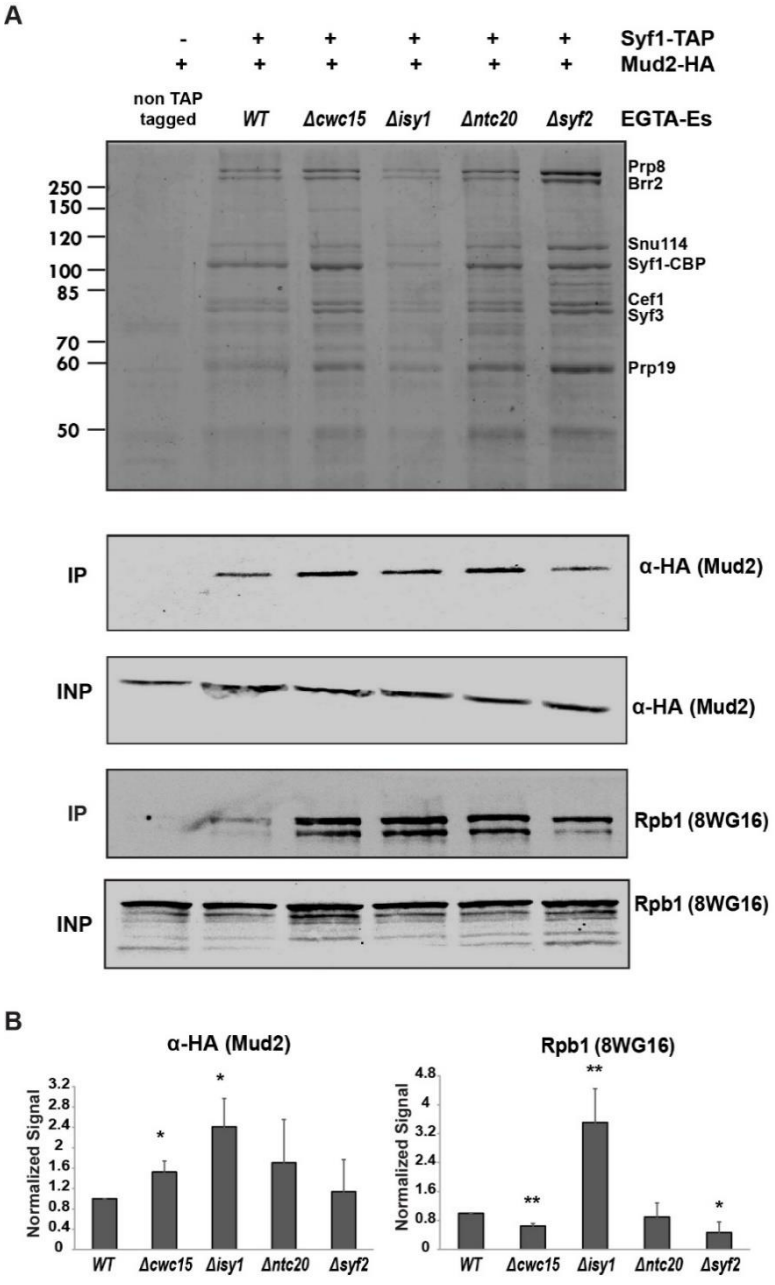


Figure 50. Interaction between the Prp19C and Mud2 is not affected while the interaction between Prp19C and RNAPII is differentially affected in *Δcwc15*, *Δisy1*, *Δntc20* and *Δsyf2* mutant cells *in vivo.* (A) Upper panel shows TAP purified EGTA-eluates from the cells expressing *SYF1-TAP* with HA-tagged Mud2 in wild-type and *Δcwc15*, *Δisy1*, *Δntc20* and *Δsyf2* mutant cells. A strain expressing only *MUD2-HA* served as a negative control. Lower panel shows the Western blot analysis of Input (INP; whole cell extract) and purified EGTA-eluates (IP; Immunoprecipitated) from above purifications using anti-HA and 8WG16 (Rpb1) antibody. (B) The quantification of three independent experiments was performed with LabImage 1D software.

3.2.5 Deletion of non-essential subunits of Prp19C confers 6-azauracil sensitivity to the cells

Prp19C plays a role in transcription elongation but it is still not known which subunits of the complex are needed for its function. To investigate the role of Prp19C subunits in this regard, all the deletion mutant cells in this study were tested for their sensitivity to the drug 6AU, as tested before for *Δmud2* cells. Interestingly, except the Δ*cwc15* mutant cells, all the deletions confer 6AU sensitivity to the cells (Figure 51). Although the Δ*ntc20* mutant cells showed growth defects in the presence of 6AU, Δ*isy1*, Δ*syf2* and Δ*snt309* mutant cells could not grow at all in the presence of the drug. These results indicate an involvement of Prp19C non-essential subunits in transcription elongation.

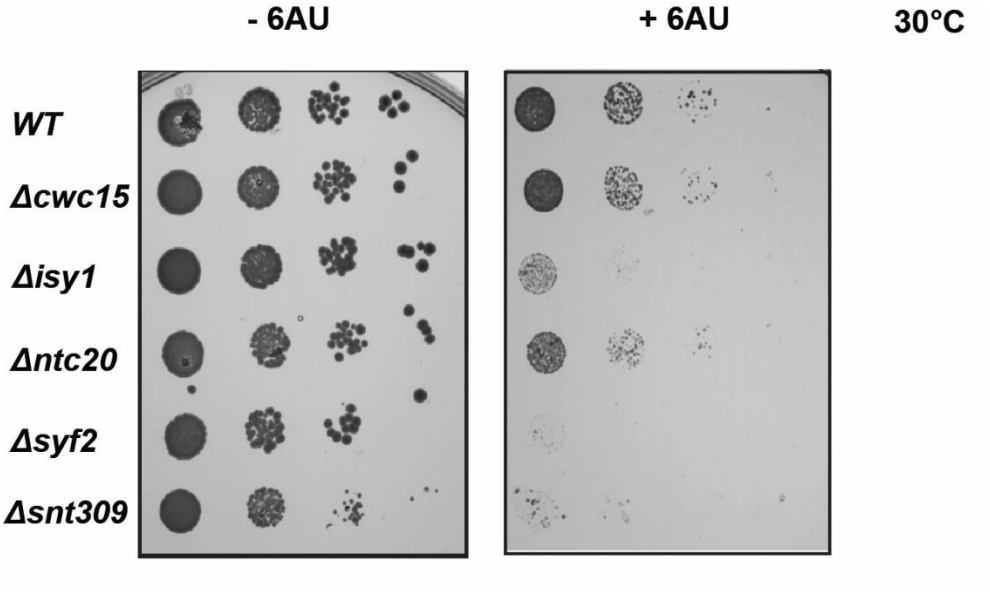


Figure 51. Deletion of *ISY1, NTC20, SYF2* and *SNT309* confers sensitivity to 6-azauracil (6AU).
Dot spot assays with 10 fold serial dilutions of wild-type (*WT*) and *Δcwc15, Δisy1, Δntc20, Δsyf2 and Δsnt309* on SDC-ura plates containing DMSO (as solvent control) or 50 µg/ml of 6AU. The plates were incubated at 30°C for 2 days.

4 Discussion and Scope

4.1 Mud2 functions in transcription by maintaining Prp19C and TREX occupancies at transcribed genes

The splicing protein Mud2 has been known since more than two decades for its role in the early splicing. Mud2 is a part of the pre-mRNA-U1 snRNP complex and interacts with pre-mRNA in a branchpoint dependent manner. It exists as a heterodimer with the branchpoint binding protein Msl5 and is important for the efficient recognition of intronic sequences. Although Mud2 is not required for the splicing reaction *per se*, it is important for the proper assembly of commitment and pre-spliceosomal complexes in early stages of splicing (Abovich, Liao et al. 1994, Rutz and Seraphin 1999, Wang, Zhang et al. 2008).

In this study, it is reported that besides its bonafide role in splicing, Mud2 also has a novel function in the gene expression, namely in transcription. Mud2 directly binds to the S2-phosphorylated CTD of RNAPII and S2 phosphorylation is required for full occupancy of Mud2 at transcribed genes. It also interacts physically with Prp19C in an RNA-independent manner. Deletion of Mud2 causes the loss of Prp19C and hence, also of the TREX complex from target genes. Thus, Mud2 can be classified as a novel factor required for the efficient transcription by RNAPII.

Mud2 is recruited to both intron-containing as well as intronless genes (Figure 11). This is consistent with the PAR-CLIP analysis which shows that Mud2 is not only bound to intron-containing pre-mRNA sequences but also to intronless mRNAs (Baejen, Torkler et al. 2014). Additionally, U2AF50 (Drosophila Mud2) has also been found to be associated with intronless RNAs (Blanchette, Labourier et al. 2004). Furthermore, recruitment of Mud2 to transcribed genes is dependent on active transcription (Figure 12). This implies a role of Mud2 in the transcription.

The conserved TREX complex performs a key role in mRNP biogenesis and nuclear export. In *S. cerevisiae*, recruitment of TREX is mediated by direct binding of the THO subcomplex to the Ser2/Ser5 di-phosphorylated CTD (Meinel, Burkert-Kautzsch et al. 2013). Moreover, Prp19C is required to maintain TREX occupancy at transcribed genes. The C-terminal domain

of Syf1 subunit of Prp19C is needed for the recruitment of Prp19C as well as for stabilizing TREX at actively transcribed genes. The C-terminus of Syf1 also mediates the interaction of Prp19C with RNAPII (Chanarat, Seizl et al. 2011). Using *in vitro* binding experiments with CTD peptides and purified Mud2, an interaction of Mud2 with the S2-phosphorylated CTD of RNAPII was shown (Figure 18). Moreover, S2 phosphorylation is necessary to maintain the full occupancy of Mud2 at transcribed genes (Figure 13). Therefore, like TREX, Mud2 also interacts with transcriptional machinery via the phosphorylated CTD of RNAPII (Figure 52). It further marks the importance of RNAPII CTD and its phosphorylation patterns in coordinating and associating various components of mRNP biogenesis.

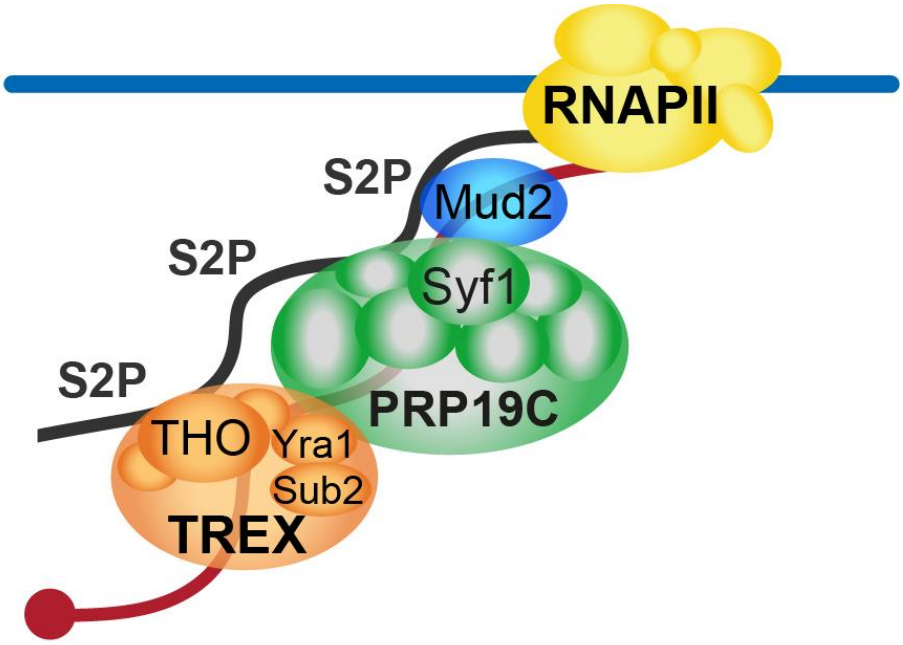


Figure 52. Model of Mud2's function in transcription by ensuring Prp19C and TREX occupancy at transcribed genes.
Mud2 interacts with the transcriptional machinery via S2-phosphorylated CTD of RNAPII. It also interacts with Prp19C and is required to maintain its occupancy at transcribed genes. Furthermore, Prp19C maintains TREX occupancy, which also interacts with the S2-phosphorylated CTD. Deletion of Mud2 leads to a decrease of mRNA levels *in vivo* and *in vitro*. Thus, Mud2 functions in transcription by ensuring Prp19C and TREX at actively transcribed genes.

Mud2 interacts with Prp19C *in vivo* in an RNA-independent manner (Figure 20). This is consistent with the previous finding which showed an interaction of Mud2 with the Syf3 subunit of Prp19C (Chung, McLean et al. 1999). Furthermore, when *MUD2* is deleted, Prp19C occupancy at transcribed genes decreases approximately by 50 percent (Figure 21). This

decrease of Prp19C is not due to a general decrease in the transcription as the occupancy of Rpb1 and S2P Rpb1 remains constant in the *MUD2* deleted cells (Figure 23 and 24). Importantly, the occupancies of two transcription elongation factors Paf1 and Spt5 are also not affected in *MUD2* deleted cells (Figure 26). This further implies that loss of Prp19C is not because of any non-specific loss of transcription factors but is a specific effect of *MUD2* deletion. This strongly suggests that Mud2 might recruit Prp19C or regulate the occupancy of Prp19C at the genes. If this is the case, the decrease of Mud2 occupancy in *rpb1-S2A* mutants should also affect Prp19C occupancy at actively transcribed genes. Indeed, loss of Mud2 from genes also leads to decrease of Prp19C occupancy in these mutants (Figure 14). Thus, Mud2 interacts with the transcriptional machinery via S2P CTD and Prp19C (Figure 52).

Prp19C interacts with the TREX *in vivo* as well as ensures TREX occupancy through the C-terminus of Syf1 (Chanarat, Seizl et al. 2011). As expected, decrease of Prp19C occupancy in *Δmud2* cells also leads to the decrease of TREX occupancy in these mutants (Figure 22). These results suggest that Prp19C associates with the transcriptional machinery via at least two modes: one is through Mud2 and the other is through the C-terminus of Syf1 (Figure 52). Furthermore, the C-terminus of Syf1 also interacts with the RNAPII either directly or indirectly (Chanarat, Seizl et al. 2011). Similarly, TREX also interacts with the transcriptional machinery via RNAPII CTD and Prp19C/Mud2 (Figure 52). Besides, TREX should associate via two different interaction sites with Prp19C and Mud2. Consistent with this possibility, it was observed that *MUD2* deletion combined with a C-terminal truncation of Syf1 (both of which lead to the loss of Prp19C as well as TREX from target genes), results in a synthetically lethal phenotype (Figure 40). Thus, all these results provide a possibility that Mud2, Prp19C and TREX associate closely – directly or indirectly – with each other as well as with the transcriptional machinery and regulate gene expression. Furthermore, Mud2 must act upstream of Prp19C and TREX, as the occupancy of Mud2 does not change when the TREX subunit *HPR1* is deleted or the C-terminus of *SYF1* is truncated (Figure 27 and 28). Taken together, Mud2 might be a recruiting factor for both Prp19C and TREX to transcribed genes.

Deletion of *MUD2* confers 6AU sensitivity to cells (Figure 29). Sensitivity to the drug 6AU has been used as a marker previously for indicating that a protein is involved in transcription elongation. For instance, deletion of *DST1*, a gene which codes for a widely studied transcription elongation factor TFIIS, has been shown to confer 6AU sensitivity to cells (Mason and Struhl 2005). Interestingly, it was observed that deletion of *MUD2* combined with the deletion of *DST1* leads to a synthetically enhanced growth retardation in the presence of 6AU

(Figure 30). This provides a genetic evidence that Mud2 should be involved in a similar or related pathway as Dst1. Furthermore, the sensitivity of *Δmud2* cells to 6AU is not a secondary effect of inefficient splicing because deletion of other splicing factors which are quite closely related to Mud2 does not cause 6AU sensitivity (Figure 29).

Both TREX and Prp19C function in transcription elongation. Mud2 is required to ensure TREX and Prp19C occupancy at transcribed genes. Notably, consistent with the above facts, Mud2 is also required for efficient mRNA synthesis *in vivo* and *in vitro* (Figure 36-39). The function of Mud2 in transcription is presumably direct, as indicated by an increase in the transcriptional activity of *Δmud2* extracts with the addition of purified Mud2 (Figure 39B and 39C). However, the stimulatory effect of Mud2 addition is quite low. This might be due to the fact that purified Mud2 is not folded in the optimal manner required for its function. Another possible explanation for this could be that *Δmud2* cells lack another factor required for their full activity. Nevertheless, these results corroborate a novel function of Mud2 in transcription.

The role of Mud2 in transcription is most likely during transcription elongation. This is because of the fact that *MUD2* deleted cells are 6AU sensitive and Mud2 ensures Prp19C and TREX at genes, both of which are required for efficient transcription elongation. Also, as observed from dot spot assays shown in Figure 34, all the three RRM domains of Mud2 are required to complement the 6AU sensitivity of *Δmud2* cells. This supports the hypothesis that only the N-terminal domain of Mud2 might be dispensable for its function in transcription elongation. Taken together, Mud2 most likely functions in transcription elongation by interacting with the transcriptional machinery through its RRM domains and also by ensuring TREX and Prp19C occupancy at transcribed genes. Deletion of *MUD2* leads to a decrease in transcriptional activity of RNAPII, which is also reflected by the slight decrease of total cellular levels of RNAPII in *Δmud2* cells (Figure 25). This decrease of RNAPII is not reflected in the ChIP occupancy of Rpb1 or S2P phosphorylated Rpb1 in *Δmud2* cells. This infers that even though the total cellular levels of Rpb1 decrease in *Δmud2* cells, the amount of RNAPII is still sufficient for its function in transcription. The decrease in total protein levels of Rpb1 might also be due to the fact that Rpb1 degrades in *Δmud2* cells, consistent with the degradation of Rpb1 observed previously in some other transcription elongation mutants (Karakasili, Burkert-Kautzsch et al. 2014). The difference between the sensitivity of ChIP and Western blotting technique might also cause a little difference observed between the two results. Nevertheless, Mud2 has a novel function in the gene expression i.e. in the transcription.

4.2 Role of RNA Recognition Motifs of Mud2 in transcription

Biochemical characterization has shown that in mammalian cells, the RNA recognition motifs RRM1 and RRM2 (Figure 33) of U2AF65 are required to recognize the polypyrimidine tract in the intronic sequence of pre-mRNA (Kielkopf, Lucke et al. 2004, Agrawal, Salsi et al. 2016). In *S. cerevisiae*, there is a poly-U tract instead of the polypyrimidine tract in the upstream of 3' splice site (Ma and Xia 2011). This explains why in yeast Mud2, only the RRM3 domain is well characterized while other two RRMs are not so well defined. However, it was found that all the three RRM domains of Mud2 are required to complement the 6AU sensitivity of *Δmud2* cells (Figure 34). This indicates that albeit less conserved, RRM1 and RRM2 domains of yeast Mud2 (marked as M-A and M-B in Figure 33) might also be important for Mud2's function in transcription. It still remains to be unraveled using *in vivo* and *in vitro* assays, whether all the three RRMs of Mud2 are required for efficient mRNA synthesis by RNAPII.

4.3 Mud2 is not required for mRNA export

TREX is a major complex in mRNP biogenesis that is also required for efficient mRNA export (Strasser, Masuda et al. 2002, Katahira and Yoneda 2014). In contrast, Prp19C is not required for export of mRNAs (Chanarat, Seizl et al. 2011). Similar to Prp19C, Mud2 is also not required for mRNA export (Figure 41). This might be one of the reasons why Mud2 is not an essential protein in yeast. This is also in agreement with the fact that in yeast, deletion of *MUD2* only leads to a partial decrease of Prp19C and TREX occupancies at the genes (Figure 21 and 22). This means that the remaining amounts of Prp19C and TREX present at genes are still sufficient for efficient transcription elongation and mRNA export. On the other hand, Mud2 is an essential protein in higher organisms (Gama-Carvalho, Krauss et al. 1997). Additionally, Drosophila Mud2 (U2AF50) has been shown to be necessary for nuclear export of intronless mRNAs (Blanchette, Labourier et al. 2004). Likewise, in human cells, U2AF65 recruits its heterodimeric partner U2AF35 and the mRNA export receptor TAP/NXF1 to the mRNP and directly stimulates export of mRNA (Zolotukhin, Tan et al. 2002). Thus, in contrast to *S. cerevisiae*,

Mud2 is required for mRNA export in higher organisms. Taking all these findings into consideration, it can be concluded that yeast Mud2 might be a distant homologue of mammalian Mud2 and diverged during the course of evolution in order to serve the complex requirements of multicellular organisms.

4.4 Function of Mud2 in transcription might be conserved in higher organisms

U2AF65 interacts with Prp19C as well as with the phosphorylated CTD. This interaction leads to an increased recruitment of both Mud2 and Prp19C, which further enhances the splicing reaction (David, Boyne et al. 2011). PAR-CLIP analysis shows that Mud2 is bound to both intron-containing and intronless genes (Baejen, Torkler et al. 2014). In line with that, Mud2 has been found to be associated with intronless mRNAs in *Drosophila* (Blanchette, Labourier et al. 2004). Therefore, despite of the fact that Mud2 is a remote homologue of U2AF65, possibility remains that the prominent function of Mud2 in transcription is conserved in higher eukaryotes.

In mammalian cells, three complexes have been identified harboring subunits which are homologues of yeast Prp19C subunits: - Prp19/CDC5 complex, PRP19-associated complex and the XAB2 complex. In *Drosophila*, two different Prp19 complexes have been identified. Therefore, it is still an open question whether Mud2 is also required to ensure the occupancy of any of these Prp19C complexes found in higher organisms. If this is the case, Mud2 might also play a role in ensuring TREX occupancy at transcribed genes in higher organisms.

4.5 The Mud2 binding partner Msl5 may also be required for efficient transcription

During early spliceosome assembly, U1 snRNP binds to the 5’ splice site in the pre-mRNA. U1 snRNP also interacts with the branchpoint region through the binding of two U1 snRNP proteins, Prp39 and Prp40 with the branchpoint binding protein Msl5/BBP (SF1 in mammalian

cells) (Wang, Zhang et al. 2008). Mud2 has also been found to be associated with Msl5 and this interaction between the two is phylogenetically conserved. Mud2 and Msl5 can be purified as a heterodimer with both the proteins existing in 1:1 stoichiometry. Additionally, Mud2 binds to Msl5 *in vitro* in an RNA-independent manner. This interaction is important for the recognition of branchpoint sequence during pre-spliceosomal assembly. In contrast to Mud2, Msl5 is essential for vegetative growth in yeast (Berglund, Abovich et al. 1998, Rutz and Séraphin 2000, Wang, Zhang et al. 2008).

The N-terminal region of Msl5 (41–141 aa in yeast) has been known for its interaction with Mud2. A conditional mutant of Msl5 (*msl5-2*), which contains two point mutations in this region, shows a strong temperature sensitivity but no splicing defects *in vitro* (Rutz and Séraphin 2000). Although the complete deletion of *MUD2* is not lethal, severe growth defects in *msl5-2* mutants reflect that the N-terminus of Msl5 might be important for a function of Msl5 distinctive than interacting with Mud2. Similar to Mud2, Msl5 also shows crosslinks with intronless RNA in PAR-CLIP analysis (Baejen, Torkler et al. 2014). In addition to this, Msl5 has been found to interact with the nuclear envelope associating protein Mlp1 (Galy, Gadal et al. 2004, Bonnet and Palancade 2015) and plays a role in the nuclear retention of the pre-mRNA (Rutz and Séraphin 2000). All these results support the hypothesis that like Mud2, Msl5 might also be involved in different steps of gene expression, particularly, transcription.

4.6 Mud2's function may couple various events during mRNP biogenesis

Gene expression in eukaryotic cells relies on tight coupling of the individual steps. For example, proteins involved in pre-mRNA processing are already recruited to actively transcribing genes (Moore and Proudfoot 2009, Bentley 2014), mRNA packaging into the mRNP occurs co-transcriptionally (Jensen, Dower et al. 2003) and nuclear export of the mRNA is also coupled to transcription and processing (Lei, Krebber et al. 2001). Mud2 is well known for its role in splicing. In this study, it is shown that Mud2 also functions in transcription by ensuring Prp19C and TREX occupancies at transcribed genes. Since Prp19C plays a role in transcription elongation independent of splicing and TREX couples transcription with mRNA export, there is a possibility that Mud2 is required for packaging of the mRNA into the mRNP. Thus, it is

quite intriguing to further explore whether Mud2 coordinates or regulates transcription, splicing and mRNA export and to further investigate how Mud2's function in maintaining Prp19C and TREX at genes helps in this regulation.

Interestingly, it has been demonstrated that Mud2 interacts genetically and functionally interacts with the TREX subunit Sub2 and Sub2 may function in the removal of Mud2 during pre-mRNA splicing (Kistler and Guthrie 2001). Thus, Mud2 along with Sub2 might couple splicing to transcription and nuclear mRNP biogenesis. Furthermore, Mud2 also interacts genetically and physically with the RNA binding protein Nab2 (Soucek, Zeng et al. 2016). As Nab2 is a multi-functional protein involved in splicing, poly(A) tail length control, nuclear mRNP packaging and mRNA export (Batisse, Batisse et al. 2009, Reuter, Meinel et al. 2015, Schmid, Olszewski et al. 2015, Soucek, Zeng et al. 2016), it is extremely interesting to determine whether this interaction between Mud2 and Nab2 serves as a bridging factor to couple Mud2's function with other processes during mRNP biogenesis. All these findings thus indicate that Mud2 might not only be required for the individual steps of gene expression, but may also be necessary for the coordination and regulation of these steps. In summary, Mud2 might provide a quality control mechanism to ensure that only the correctly processed and packaged mRNPs are targeted for nuclear export.

4.7 The non-essential subunits of Prp19C play differential roles in gene expression

Previous results demonstrate that Prp19C interacts with THO *in vivo* in an RNA-independent manner (Chanarat, Seizl et al. 2011). In this study, a direct interaction between these two complexes was demonstrated. As identified from the Coomassie stained polyacrylamide gel (Figure 43), all core components of Prp19C - Syf1, Cef1, Syf3 and Prp19 - bind to the THO complex. Thus, Prp19C maintains TREX occupancy at transcribed genes by directly binding to its THO subcomplex.

In this study, the role of non-essential subunits of Prp19C and the NTC related protein, Cwc15, in maintaining Prp19C and TREX occupancy was also investigated. It was found that Isy1, Syf2 and Ntc20 components of Prp19C are required for maintaining Prp19C occupancy at the intron-

containing genes (Figure 44). This suggests that these subunits might be needed by Prp19C for splicing. Albeit Isy1, Syf2 and Ntc20 are not required for the cell viability, they are important for the stable association and recruitment of the whole complex at intron-containing genes. If any of these components is not present, Prp19C is lost from the intron-containing genes, which might lead to inefficient splicing.

Interestingly, at the intronless genes, the occupancy of Prp19C decreases in *Δisy1* mutants. Thus, Isy1 represents a key factor which maintains the activity of the whole complex at transcribed genes, whether intron-containing or intronless. This further supports the idea that Isy1 is not only an important factor required for pre-mRNA splicing, but it is also needed for the function of Prp19C in transcription elongation. Moreover, Prp19C occupancy is not affected at intron-containing or intronless genes in *Δcwc15* mutants. This might be due to the fact that Cwc15 is not a core component of Prp19C, but only associates with Cef1 protein of the complex and therefore, deletion of this protein does not affect the occupancy of whole complex. Nevertheless, it can be concluded from these results that the Isy1 subunit of Prp19C plays a pivotal role in maintaining Prp19C occupancy at transcribed genes.

As Prp19C stabilizes TREX at transcribed genes, deletion of *ISY1* should also affect TREX occupancy. Indeed, it was found that TREX occupancy decreases in *Δisy1* mutant cells at both intron-containing and intronless genes (Figure 45). Thus, Isy1 is required for the function of Prp19C in ensuring TREX at the genes. Notably, total cellular levels of Prp19C and TREX, as assessed by its subunits Syf1 and Hpr1 respectively, are not affected in *Δisy1* mutants (Figure 46). This suggests that loss of Prp19C and TREX from transcribed genes is a specific effect of *ISY1* deletion. Importantly, this decrease is also not caused by the loss in the transcriptional activity due to the deletion of an important splicing factor, as the overall occupancy of RNAPII subunit Rpb1 does not decrease in these mutants (Figure 47). Instead, the RNAPII occupancy increases significantly in *Δisy1* mutant cells. Despite of the fact that more RNAPII is recruited at the genes in *Δisy1* mutant cells, the stability of Prp19C and TREX decreases at transcribed genes. Notably, RNAPII occupancy is also increased at the genes in Δ*syf2* mutant cells (Figure 47). This indicates that the Prp19C and TREX occupancy normalized to the RNAPII occupancy should decrease in these mutants. Thus, Syf2 might also contribute to the Prp19C and TREX occupancy although the effect is not as evident as in the case of Isy1. The reason for an increased RNAPII recruitment in *Δisy1* and Δ*syf2* mutants still needs to be examined. Nevertheless, *ISY1* has an additional function in the gene expression, namely in maintaining Prp19C and TREX occupancies at transcribed genes.

Paradoxically, the total cellular levels of Syf1 and Hpr1 decrease in *Δcwc15* and *Δsyf2* mutant cells (Figure 46) despite of the fact that their occupancies, as assessed by ChIP, are not affected at transcribed genes. Thus, although Hpr1 and Syf1 levels decrease in these two mutants, they are still sufficient to maintain full occupancy of Prp19C and TREX at the genes. This might explain why these components are not required for overall cell viability.

In addition to analyzing the role in Prp19C and TREX occupancy, the function of these non-essential proteins in Prp19C-TREX, Prp19C-Mud2 and Prp19C-RNAPII interaction was also investigated. It was found that the interaction between Prp19C and TREX significantly decreases in Δ*cwc15*, Δ*ntc20* and Δ*syf2* and increases in Δ*isy1* mutant cells (Figure 48). The increased interaction between the two complexes in Δ*isy1* mutant cells might be due to the increased recruitment of RNAPII in these mutants. On the other hand, loss of Prp19C-TREX interaction in Δ*cwc15* and Δ*syf2* mutants is consistent with the decrease of total cellular levels of Hpr1 and Syf1 and also with the degradation of Hpr1 observed during the purification of TREX complex in these mutants. It is possible that Hpr1 is not stable enough or fully expressed in Δ*cwc15* and Δ*syf2* mutant cells. Surprisingly, Hpr1 degradation was not observed when Hpr1 was TAP tagged at its N-terminus rather than its C-terminus (Figure 49). Also, the expression of Hpr1 does not decrease in Δ*cwc15* and Δ*syf2* mutants when the Hpr1 is TAP-tagged at its N-terminus (Figure 49). The reason for this behavior is not completely clear and more experiments are required for a deeper understanding. But there might be a possibility that Cwc15 and Syf2 interact with TREX with respect to the C-terminus of Hpr1 and somehow this interaction is interrupted in these two mutant cells. Nevertheless, precise role of these proteins in Prp19C-TREX interaction remain to be resolved.

Surprisingly, the interaction of Prp19C with Mud2 is not affected in any of the deletion mutants (Figure 50). This provides a possibility that Mud2 might be interacting with Prp19C via essential components of the complex. This is consistent with the previous finding that Mud2 interacts with Syf3 (Chung, McLean et al. 1999) and our finding that Mud2 interacts with Prp19 (Figure 20), both of which are essential components of the complex.

Additionally, the interaction of Prp19C with RNAPII decreases in Δ*cwc15* and Δ*syf2* mutants and increases in Δ*isy1* mutant cells. The decrease in Prp19C-RNAPII interaction in Δ*cwc15* and Δ*syf2* mutants might be because of the loss of Syf1 levels in these two mutant cells (Figure 46). The increase of Prp19C-RNAPII interaction in Δ*isy1* mutants might be because of the increased recruitment of RNAPII at transcribed genes in these cells.

This study of non-essential components of Prp19C strongly indicates that even though these components are not necessary for cell survival or vegetative growth, nonetheless, they all play significant and differential roles in gene expression. It also suggests that each of these components might play specific roles in transcription. This is also consistent with the observation that most of these components show severe 6AU sensitivity which indicates their involvement in transcription elongation (Figure 51). Thus, this analysis provides an interesting framework for further investigation of the functions of these factors in gene expression. It also raises some open questions, resolving of which will provide a deeper mechanistic understanding of how Prp19C and TREX interact with each other and how they are involved in transcription. This will further contribute to our understanding of Prp19C and TREX involvement in mRNP biogenesis.

4.8 E3 ligase activity of Prp19

The core component Prp19 of the Prp19C contains a U-box at its N-terminus (Ohi, Vander Kooi et al. 2003). The U-box has been identified as a new class of protein domains which are related to ubiquitin ligase activity. Thus, proteins with a U-box have been identified as functional E3 ubiquitin ligases (Hatakeyama, Yada et al. 2001). For instance, it has been previously reported that Prp3, a component of the U4 snRNP, interacts with Prp19 and is ubiquitylated by Prp19 with non-proteolytic K63-linked ubiquitin chains. Prp8, a component of the U5 snRNP, acts as a receptor for the ubiquitylated form of Prp3. This Prp3-Prp8 interaction promotes the stabilization of the U4/U6/U5 tri-snRNP complex (Song, Werner et al. 2010). Thus, the E3 ligase activity of Prp19 is important for the structural rearrangements during spliceosomal assembly. Apart from this, the U-box of Prp19 is also required for the ubiquitylation of Rpa32 (subunit of Replication protein A) with K63-linked chains in a DNA damage induced manner. This Prp19-mediated ubiquitylation further promotes the recruitment of ATRIP (functional partner of ATR kinase) to the DNA damage sites. *PRP19* mutants that failed to function as an E3 ubiquitin ligase have been found to be defective in the recovery of stalled replication forks at the DNA damage site (Marechal, Li et al. 2014). Thus, E3 ligase activity of Prp19 also plays a key role in DNA damage response.

Athough the role of E3 ubiquitin ligase activity of the Prp19's U-box has been well studied in various cellular events, it still remains elusive whether this activity plays a role in TREX

recruitment and its stabilization at transcribed genes. Notably, the Hpr1 component of the TREX complex has been known to be polyubiquitylated by Rsp5, an ubiquitin ligase from the HECT family. The polyubiquitylated Hpr1 promotes the cotranscriptional recruitment of mRNA export receptor Mex67 by physically interacting with the UBA (ubiquitin associated) domain of Mex67. Deletion of this domain leads to an increased degradation of Hpr1 by the 26S proteasome (Gwizdek, Iglesias et al. 2006). There are no evidences yet regarding the role of U-box of Prp19 in recruitment of mRNA export machinery. It is also extremely intriguing to investigate the role of E3 ligase activity of Prp19 (and the substrates involved) in recruitment of TREX as well as in transcription.

5 References

Abovich, N., X. C. Liao and M. Rosbash (1994). "The yeast MUD2 protein: an interaction with PRP11 defines a bridge between commitment complexes and U2 snRNP addition." Genes Dev **8**(7): 843-854.

Agrawal, A. A., E. Salsi, R. Chatrikhi, S. Henderson, J. L. Jenkins, M. R. Green, D. N. Ermolenko and C. L. Kielkopf (2016). "An extended U2AF(65)–RNA-binding domain recognizes the 3′ splice site signal." Nature Communications **7**: 10950.

Anderson, J. T., S. M. Wilson, K. V. Datar and M. S. Swanson (1993). "NAB2: a yeast nuclear polyadenylated RNA-binding protein essential for cell viability." Mol Cell Biol **13**(5): 2730-2741.

Baejen, C., P. Torkler, S. Gressel, K. Essig, J. Soding and P. Cramer (2014). "Transcriptome maps of mRNP biogenesis factors define pre-mRNA recognition." Mol Cell **55**(5): 745-757.

Batisse, J., C. Batisse, A. Budd, B. Bottcher and E. Hurt (2009). "Purification of nuclear poly(A)-binding protein Nab2 reveals association with the yeast transcriptome and a messenger ribonucleoprotein core structure." J Biol Chem **284**(50): 34911-34917.

Bentley, D. L. (2014). "Coupling mRNA processing with transcription in time and space." Nature reviews. Genetics **15**(3): 163-175.

Berben, G., J. Dumont, V. Gilliquet, P. A. Bolle and F. Hilger (1991). "The YDp plasmids: a uniform set of vectors bearing versatile gene disruption cassettes for Saccharomyces cerevisiae." Yeast **7**(5): 475-477.

Berglund, J. A., N. Abovich and M. Rosbash (1998). "A cooperative interaction between U2AF65 and mBBP/SF1 facilitates branchpoint region recognition." Genes Dev **12**(6): 858-867.

Bjork, P. and L. Wieslander (2017). "Integration of mRNP formation and export." **74**(16): 2875-2897.

Blanchette, M., E. Labourier, R. E. Green, S. E. Brenner and D. C. Rio (2004). "Genome-wide analysis reveals an unexpected function for the Drosophila splicing factor U2AF50 in the nuclear export of intronless mRNAs." Mol Cell **14**(6): 775-786.

Blythe, A. J., B. Yazar-Klosinski, M. W. Webster, E. Chen, M. Vandevenne, K. Bendak, J. P. Mackay, G. A. Hartzog and A. Vrielink (2016). "The yeast transcription elongation factor Spt4/5 is a sequence-specific RNA binding protein." Protein Science : A Publication of the Protein Society **25**(9): 1710-1721.

Bonnet, A. and B. Palancade (2015). "Intron or no intron: a matter for nuclear pore complexes." Nucleus **6**(6): 455-461.

Bradford, M. M. (1976). "A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding." Anal Biochem **72**: 248-254.

Chan, S., E.-A. Choi and Y. Shi (2011). "Pre-mRNA 3'-end processing complex assembly and function." Wiley interdisciplinary reviews. RNA **2**(3): 321-335.

Chan, S. P. and S. C. Cheng (2005). "The Prp19-associated complex is required for specifying interactions of U5 and U6 with pre-mRNA during spliceosome activation." J Biol Chem **280**(35): 31190-31199.

Chan, S. P., D. I. Kao, W. Y. Tsai and S. C. Cheng (2003). "The Prp19p-associated complex in spliceosome activation." Science **302**(5643): 279-282.

Chanarat, S., C. Burkert-Kautzsch, D. M. Meinel and K. Strasser (2012). "Prp19C and TREX: interacting to promote transcription elongationand mRNA export." Transcription **3**(1): 8-12.

Chanarat, S., M. Seizl and K. Strasser (2011). "The Prp19 complex is a novel transcription elongation factor required for TREX occupancy at transcribed genes." Genes & development **25**(11): 1147-1158.

Chanarat, S. and K. Strasser (2013). "Splicing and beyond: the many faces of the Prp19 complex." Biochim Biophys Acta **1833**(10): 2126-2134.

Chavez, S., T. Beilharz, A. G. Rondon, H. Erdjument-Bromage, P. Tempst, J. Q. Svejstrup, T. Lithgow and A. Aguilera (2000). "A protein complex containing Tho2, Hpr1, Mft1 and a novel protein, Thp2, connects transcription elongation with mitotic recombination in Saccharomyces cerevisiae." Embo j **19**(21): 5824-5834.

Chen, C.-H., D.-I. Kao, S.-P. Chan, T.-C. Kao, J.-Y. Lin and S.-C. Cheng (2006). "Functional links between the Prp19-associated complex, U4/U6 biogenesis, and spliceosome recycling." RNA **12**(5): 765-774.

Chen, C. H., W. Y. Tsai, H. R. Chen, C. H. Wang and S. C. Cheng (2001). "Identification and characterization of two novel components of the Prp19p-associated complex, Ntc30p and Ntc20p." J Biol Chem **276**(1): 488-494.

Chen, C. H., W. C. Yu, T. Y. Tsao, L. Y. Wang, H. R. Chen, J. Y. Lin, W. Y. Tsai and S. C. Cheng (2002). "Functional and physical interactions between components of the Prp19p-associated complex." Nucleic Acids Res **30**(4): 1029-1037.

Chen, H. R., S. P. Jan, T. Y. Tsao, Y. J. Sheu, J. Banroques and S. C. Cheng (1998). "Snt309p, a component of the Prp19p-associated complex that interacts with Prp19p and associates with the spliceosome simultaneously with or immediately after dissociation of U4 in the same manner as Prp19p." Mol Cell Biol **18**(4): 2196-2204.

Cheng, H., K. Dufu, C. S. Lee, J. L. Hsu, A. Dias and R. Reed (2006). "Human mRNA export machinery recruited to the 5' end of mRNA." Cell **127**(7): 1389-1400.

Chung, S., M. R. McLean and B. C. Rymond (1999). "Yeast ortholog of the Drosophila crooked neck protein promotes spliceosome assembly through stable U4/U6.U5 snRNP addition." Rna **5**(8): 1042-1054.

David, C. J., A. R. Boyne, S. R. Millhouse and J. L. Manley (2011). "The RNA polymerase II C-terminal domain promotes splicing activation through recruitment of a U2AF65-Prp19 complex." Genes Dev **25**(9): 972-983.

de Almeida, R. A. and R. T. O'Keefe (2015). "The NineTeen Complex (NTC) and NTC-associated proteins as targets for spliceosomal ATPase action during pre-mRNA splicing." RNA Biol **12**(2): 109-114.

Fasken, M. B., M. Stewart and A. H. Corbett (2008). "Functional significance of the interaction between the mRNA-binding protein, Nab2, and the nuclear pore-associated protein, Mlp1, in mRNA export." J Biol Chem **283**(40): 27130-27143.

Galy, V., O. Gadal, M. Fromont-Racine, A. Romano, A. Jacquier and U. Nehrbass (2004). "Nuclear retention of unspliced mRNAs in yeast is mediated by perinuclear Mlp1." Cell **116**(1): 63-73.

Gama-Carvalho, M., R. D. Krauss, L. Chiang, J. Valcarcel, M. R. Green and M. Carmo-Fonseca (1997). "Targeting of U2AF65 to sites of active splicing in the nucleus." J Cell Biol **137**(5): 975-987.

Green, D. M., K. A. Marfatia, E. B. Crafton, X. Zhang, X. Cheng and A. H. Corbett (2002). "Nab2p is required for poly(A) RNA export in Saccharomyces cerevisiae and is regulated by arginine methylation via Hmt1p." J Biol Chem **277**(10): 7752-7760.

Gwizdek, C., N. Iglesias, M. S. Rodriguez, B. Ossareh-Nazari, M. Hobeika, G. Divita, F. Stutz and C. Dargemont (2006). "Ubiquitin-associated domain of Mex67 synchronizes recruitment of the mRNA export machinery with transcription." Proceedings of the National Academy of Sciences of the United States of America **103**(44): 16376-16381.

Hackmann, A., T. Gross, C. Baierlein and H. Krebber (2011). "The mRNA export factor Npl3 mediates the nuclear export of large ribosomal subunits." EMBO Rep **12**(10): 1024-1031.

Hahn, S. (2004). "Structure and mechanism of the RNA Polymerase II transcription machinery." Nature structural & molecular biology **11**(5): 394-403.

Harlen, K. M. and L. S. Churchman (2017). "The code and beyond: transcription regulation by the RNA polymerase II carboxy-terminal domain." Nat Rev Mol Cell Biol **18**(4): 263-273.

Harlen, K. M., K. L. Trotta, E. E. Smith, M. M. Mosaheb, S. M. Fuchs and L. S. Churchman (2016). "Comprehensive RNA Polymerase II Interactomes Reveal Distinct and Varied Roles for Each Phospho-CTD Residue." Cell reports **15**(10): 2147-2158.

Hatakeyama, S., M. Yada, M. Matsumoto, N. Ishida and K. I. Nakayama (2001). "U box proteins as a new family of ubiquitin-protein ligases." J Biol Chem **276**(35): 33111-33120.

Heath, C. G., N. Viphakone and S. A. Wilson (2016). "The role of TREX in gene expression and disease." Biochem J **473**(19): 2911-2935.

Hogg, R., J. C. McGrail and R. T. O'Keefe (2010). "The function of the NineTeen Complex (NTC) in regulating spliceosome conformations and fidelity during pre-mRNA splicing." Biochemical Society transactions **38**(4): 1110-1115.

Hope, I. A. and K. Struhl (1987). "GCN4, a eukaryotic transcriptional activator protein, binds as a dimer to target DNA." EMBO J **6**(9): 2781-2784.

Hsin, J. P. and J. L. Manley (2012). "The RNA polymerase II CTD coordinates transcription and RNA processing." Genes Dev **26**(19): 2119-2137.

Huertas, P. and A. Aguilera (2003). "Cotranscriptionally formed DNA:RNA hybrids mediate transcription elongation impairment and transcription-associated recombination." Mol Cell **12**(3): 711-721.

Hurt, E., M. J. Luo, S. Rother, R. Reed and K. Strasser (2004). "Cotranscriptional recruitment of the serine-arginine-rich (SR)-like proteins Gbp2 and Hrb1 to nascent mRNA via the TREX complex." Proc Natl Acad Sci U S A **101**(7): 1858-1862.

Jaehning, J. A. (2010). "The Paf1 Complex: Platform or Player in RNA Polymerase II Transcription?" Biochimica et biophysica acta **1799**(5-6): 379-388.

Janke, C., M. M. Magiera, N. Rathfelder, C. Taxis, S. Reber, H. Maekawa, A. Moreno-Borchart, G. Doenges, E. Schwob, E. Schiebel and M. Knop (2004). "A versatile toolbox for PCR-based tagging of yeast genes: new fluorescent proteins, more markers and promoter substitution cassettes." Yeast **21**(11): 947-962.

Jensen, T. H., K. Dower, D. Libri and M. Rosbash (2003). "Early formation of mRNP: license for export or quality control?" Mol Cell **11**(5): 1129-1138.

Jimeno, S., R. Luna, M. García-Rubio and A. Aguilera (2006). "Tho1, a Novel hnRNP, and Sub2 Provide Alternative Pathways for mRNP Biogenesis in Yeast THO Mutants." Molecular and Cellular Biology **26**(12): 4387-4398.

Jimeno, S., A. G. Rondón, R. Luna and A. Aguilera (2002). "The yeast THO complex and mRNA export factors link RNA metabolism with transcription and genome instability." The EMBO Journal **21**(13): 3526-3535.

Johnson, S. A., G. Cubberley and D. L. Bentley (2009). "Cotranscriptional recruitment of the mRNA export factor Yra1 by direct interaction with the 3' end processing factor Pcf11." Mol Cell **33**(2): 215-226.

Jonkers, I. and J. T. Lis (2015). "Getting up to speed with transcription elongation by RNA polymerase II." Nature reviews. Molecular cell biology **16**(3): 167-177.

Karakasili, E., C. Burkert-Kautzsch, A. Kieser and K. Strasser (2014). "Degradation of DNA damage-independently stalled RNA polymerase II is independent of the E3 ligase Elc1." Nucleic Acids Res **42**(16): 10503-10515.

Katahira, J. (2012). "mRNA export and the TREX complex." Biochim Biophys Acta **1819**(6): 507-513.

Katahira, J. (2015). "Nuclear Export of Messenger RNA." Genes (Basel) **6**(2): 163-184.

Katahira, J. and Y. Yoneda (2014). "Roles of the TREX complex in nuclear export of mRNA." RNA Biology **6**(2): 149-152.

Kettenberger, H., K. J. Armache and P. Cramer (2003). "Architecture of the RNA polymerase II-TFIIS complex and implications for mRNA cleavage." Cell **114**(3): 347-357.

Kielkopf, C. L., S. Lucke and M. R. Green (2004). "U2AF homology motifs: protein recognition in the RRM world." Genes Dev **18**(13): 1513-1526.

Kistler, A. L. and C. Guthrie (2001). "Deletion of MUD2, the yeast homolog of U2AF65, can bypass the requirement for sub2, an essential spliceosomal ATPase." Genes Dev **15**(1): 42-49.

Kornprobst, M., M. Turk, N. Kellner, J. Cheng, D. Flemming, I. Kos-Braun, M. Kos, M. Thoms, O. Berninghausen, R. Beckmann and E. Hurt (2016). "Architecture of the 90S Pre-ribosome: A Structural View on the Birth of the Eukaryotic Ribosome." Cell **166**(2): 380-393.

Laemmli, U. K. (1970). "Cleavage of structural proteins during the assembly of the head of bacteriophage T4." Nature **227**(5259): 680-685.

Lee, K. M. and W. Y. Tarn (2013). "Coupling pre-mRNA processing to transcription on the RNA factory assembly line." RNA Biol **10**(3): 380-390.

Lei, E. P., H. Krebber and P. A. Silver (2001). "Messenger RNAs are recruited for nuclear export during transcription." Genes & Development **15**(14): 1771-1782.

Lidschreiber, M., K. Leike and P. Cramer (2013). "Cap completion and C-terminal repeat domain kinase recruitment underlie the initiation-elongation transition of RNA polymerase II." Mol Cell Biol **33**(19): 3805-3816.

Lindstrom, D. L. and G. A. Hartzog (2001). "Genetic interactions of Spt4-Spt5 and TFIIS with the RNA polymerase II CTD and CTD modifying enzymes in Saccharomyces cerevisiae." Genetics **159**(2): 487-497.

Lindstrom, D. L., S. L. Squazzo, N. Muster, T. A. Burckin, K. C. Wachter, C. A. Emigh, J. A. McCleery, J. R. Yates, 3rd and G. A. Hartzog (2003). "Dual roles for Spt5 in pre-mRNA processing and transcription elongation revealed by identification of Spt5-associated proteins." Mol Cell Biol **23**(4): 1368-1378.

Loya, T. J. and D. Reines (2016). "Recent advances in understanding transcription termination by RNA polymerase II." F1000Research **5**: F1000 Faculty Rev-1478.

Lucchini, G., A. G. Hinnebusch, C. Chen and G. R. Fink (1984). "Positive regulatory interactions of the HIS4 gene of Saccharomyces cerevisiae." Mol Cell Biol **4**(7): 1326-1333.

Lunde, B. M., S. L. Reichow, M. Kim, H. Suh, T. C. Leeper, F. Yang, H. Mutschler, S. Buratowski, A. Meinhart and G. Varani (2010). "Cooperative interaction of transcription

termination factors with the RNA polymerase II C-terminal domain." Nat Struct Mol Biol **17**(10): 1195-1201.

Ma, P. and X. Xia (2011). "Factors affecting splicing strength of yeast genes." Comp Funct Genomics **2011**: 212146.

Mandel, C. R., Y. Bai and L. Tong (2008). "Protein factors in pre-mRNA 3′-end processing." Cellular and molecular life sciences : CMLS **65**(7-8): 1099-1122.

Marechal, A., J. M. Li, X. Y. Ji, C. S. Wu, S. A. Yazinski, H. D. Nguyen, S. Liu, A. E. Jimenez, J. Jin and L. Zou (2014). "PRP19 transforms into a sensor of RPA-ssDNA after DNA damage and drives ATR activation via a ubiquitin-mediated circuitry." Mol Cell **53**(2): 235-246.

Mariconti, L., B. Loll, K. Schlinkmann, A. Wengi, A. Meinhart and B. Dichtl (2010). "Coupled RNA polymerase II transcription and 3′ end formation with yeast whole-cell extracts." RNA **16**(11): 2205-2217.

Mason, P. B. and K. Struhl (2005). "Distinction and relationship between elongation rate and processivity of RNA polymerase II in vivo." Mol Cell **17**(6): 831-840.

Masuda, S., R. Das, H. Cheng, E. Hurt, N. Dorman and R. Reed (2005). "Recruitment of the human TREX complex to mRNA during splicing." Genes Dev **19**(13): 1512-1517.

Mayer, A., M. Heidemann, M. Lidschreiber, A. Schreieck, M. Sun, C. Hintermair, E. Kremmer, D. Eick and P. Cramer (2012). "CTD tyrosine phosphorylation impairs termination factor recruitment to RNA polymerase II." Science **336**(6089): 1723-1725.

Mayer, A., A. Schreieck, M. Lidschreiber, K. Leike, D. E. Martin and P. Cramer (2012). "The Spt5 C-Terminal Region Recruits Yeast 3′ RNA Cleavage Factor I." Molecular and Cellular Biology **32**(7): 1321-1331.

Meinel, D. M., C. Burkert-Kautzsch, A. Kieser, E. O'Duibhir, M. Siebert, A. Mayer, P. Cramer, J. Soding, F. C. Holstege and K. Strasser (2013). "Recruitment of TREX to the transcription machinery by its direct binding to the phospho-CTD of RNA polymerase II." PLoS Genet **9**(11): e1003914.

Meinel, D. M. and K. Strasser (2015). "Co-transcriptional mRNP formation is coordinated within a molecular mRNP packaging station in S. cerevisiae." Bioessays **37**(6): 666-677.

Moore, M. J. and N. J. Proudfoot (2009). "Pre-mRNA processing reaches back to transcription and ahead to translation." Cell **136**(4): 688-700.

Morris, D. P. and A. L. Greenleaf (2000). "The splicing factor, Prp40, binds the phosphorylated carboxyl-terminal domain of RNA polymerase II." J Biol Chem **275**(51): 39935-39943.

Neve, J., R. Patel, Z. Wang, A. Louey and A. M. Furger (2017). "Cleavage and polyadenylation: Ending the message expands gene regulation." RNA Biology **14**(7): 865-890.

Nissan, T. A., K. Galani, B. Maco, D. Tollervey, U. Aebi and E. Hurt (2004). "A pre-ribosome with a tadpole-like structure functions in ATP-dependent maturation of 60S subunits." Mol Cell **15**(2): 295-301.

Ohi, M. D. and K. L. Gould (2002). "Characterization of interactions among the Cef1p-Prp19p-associated splicing complex." Rna **8**(6): 798-815.

Ohi, M. D., C. W. Vander Kooi, J. A. Rosenberg, W. J. Chazin and K. L. Gould (2003). "Structural insights into the U-box, a domain associated with multi-ubiquitination." Nat Struct Biol **10**(4): 250-255.

Ohi, M. D., C. W. Vander Kooi, J. A. Rosenberg, L. Ren, J. P. Hirsch, W. J. Chazin, T. Walz and K. L. Gould (2005). "Structural and functional analysis of essential pre-mRNA splicing factor Prp19p." Mol Cell Biol **25**(1): 451-460.

Peterlin, B. M. and D. H. Price (2006). "Controlling the elongation phase of transcription with P-TEFb." Mol Cell **23**(3): 297-305.

Pokholok, D. K., N. M. Hannett and R. A. Young (2002). "Exchange of RNA polymerase II initiation and elongation factors during gene expression in vivo." Mol Cell **9**(4): 799-809.

Puig, O., F. Caspary, G. Rigaut, B. Rutz, E. Bouveret, E. Bragado-Nilsson, M. Wilm and B. Seraphin (2001). "The tandem affinity purification (TAP) method: a general procedure of protein complex purification." Methods **24**(3): 218-229.

Reed, R. (2003). "Coupling transcription, splicing and mRNA export." Curr Opin Cell Biol **15**(3): 326-331.

Rehwinkel, J., A. Herold, K. Gari, T. Kocher, M. Rode, F. L. Ciccarelli, M. Wilm and E. Izaurralde (2004). "Genome-wide analysis of mRNAs regulated by the THO complex in Drosophila melanogaster." Nat Struct Mol Biol **11**(6): 558-566.

Reuter, L. M., D. M. Meinel and K. Strasser (2015). "The poly(A)-binding protein Nab2 functions in RNA polymerase III transcription." Genes Dev **29**(14): 1565-1575.

Richard, P. and J. L. Manley (2009). "Transcription termination by nuclear RNA polymerases." Genes Dev **23**(11): 1247-1269.

Riles, L., R. J. Shaw, M. Johnston and D. Reines (2004). "Large-scale screening of yeast mutants for sensitivity to the IMP dehydrogenase inhibitor 6-azauracil." Yeast **21**(3): 241-248.

Rother, S., C. Burkert, K. M. Brunger, A. Mayer, A. Kieser and K. Strasser (2010). "Nucleocytoplasmic shuttling of the La motif-containing protein Sro9 might link its nuclear and cytoplasmic functions." Rna **16**(7): 1393-1401.

Rother, S. and K. Strasser (2007). "The RNA polymerase II CTD kinase Ctk1 functions in translation elongation." Genes Dev **21**(11): 1409-1421.

Ruby, S. W. and J. Abelson (1991). "Pre-mRNA splicing in yeast." Trends Genet **7**(3): 79-85.

Rutz, B. and B. Seraphin (1999). "Transient interaction of BBP/ScSF1 and Mud2 with the splicing machinery affects the kinetics of spliceosome assembly." Rna **5**(6): 819-831.

Rutz, B. and B. Séraphin (2000). "A dual role for BBP/ScSF1 in nuclear pre-mRNA retention and splicing." The EMBO Journal **19**(8): 1873-1886.

Santos-Pereira, J. M., A. B. Herrero, S. Moreno and A. Aguilera (2014). "Npl3, a new link between RNA-binding proteins and the maintenance of genome integrity." Cell Cycle **13**(10): 1524-1529.

Schmid, M., P. Olszewski, V. Pelechano, I. Gupta, L. M. Steinmetz and T. H. Jensen (2015). "The Nuclear PolyA-Binding Protein Nab2p Is Essential for mRNA Production." Cell Rep **12**(1): 128-139.

Schneider, D. A., S. L. French, Y. N. Osheim, A. O. Bailey, L. Vu, J. Dodd, J. R. Yates, A. L. Beyer and M. Nomura (2006). "RNA polymerase II elongation factors Spt4p and Spt5p play roles in transcription elongation by RNA polymerase I and rRNA processing." Proc Natl Acad Sci U S A **103**(34): 12707-12712.

Segref, A., K. Sharma, V. Doye, A. Hellwig, J. Huber, R. Lührmann and E. Hurt (1997). "Mex67p, a novel factor for nuclear mRNA export, binds to both poly(A)+ RNA and nuclear pores." The EMBO Journal **16**(11): 3256-3271.

Seizl, M., L. Lariviere, T. Pfaffeneder, L. Wenzeck and P. Cramer (2011). "Mediator head subcomplex Med11/22 contains a common helix bundle building block with a specific function in transcription initiation complex stabilization." Nucleic Acids Res **39**(14): 6291-6304.

Selenko, P., G. Gregorovic, R. Sprangers, G. Stier, Z. Rhani, A. Kramer and M. Sattler (2003). "Structural basis for the molecular recognition between human splicing factors U2AF65 and SF1/mBBP." Mol Cell **11**(4): 965-976.

Shandilya, J. and S. G. Roberts (2012). "The transcription cycle in eukaryotes: from productive initiation to RNA polymerase II recycling." Biochim Biophys Acta **1819**(5): 391-400.

Sickmier, E. A., K. E. Frato, H. Shen, S. R. Paranawithana, M. R. Green and C. L. Kielkopf (2006). "Structural Basis for Polypyrimidine-Tract Recognition by the Essential pre-mRNA Splicing Factor U2AF(65)." Molecular cell **23**(1): 49-59.

Sikorski, R. S. and P. Hieter (1989). "A system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in Saccharomyces cerevisiae." Genetics **122**(1): 19-27.

Sims, R. J., 3rd, R. Belotserkovskaya and D. Reinberg (2004). "Elongation by RNA polymerase II: the short and long of it." Genes Dev **18**(20): 2437-2468.

Song, E. J., S. L. Werner, J. Neubauer, F. Stegmeier, J. Aspden, D. Rio, J. W. Harper, S. J. Elledge, M. W. Kirschner and M. Rape (2010). "The Prp19 complex and the Usp4Sart3 deubiquitinating enzyme control reversible ubiquitination at the spliceosome." Genes Dev **24**(13): 1434-1447.

Soucek, S., Y. Zeng, D. L. Bellur, M. Bergkessel, K. J. Morris, Q. Deng, D. Duong, N. T. Seyfried, C. Guthrie, J. P. Staley, M. B. Fasken and A. H. Corbett (2016). "The Evolutionarily-conserved Polyadenosine RNA Binding Protein, Nab2, Cooperates with Splicing Machinery to Regulate the Fate of pre-mRNA." Mol Cell Biol.

Squazzo, S. L., P. J. Costa, D. L. Lindstrom, K. E. Kumer, R. Simic, J. L. Jennings, A. J. Link, K. M. Arndt and G. A. Hartzog (2002). "The Paf1 complex physically and functionally associates with transcription elongation factors in vivo." The EMBO Journal **21**(7): 1764-1774.

Strasser, K. and E. Hurt (2000). "Yra1p, a conserved nuclear RNA-binding protein, interacts directly with Mex67p and is required for mRNA export." Embo j **19**(3): 410-420.

Strasser, K. and E. Hurt (2001). "Splicing factor Sub2p is required for nuclear mRNA export through its interaction with Yra1p." Nature **413**(6856): 648-652.

Strasser, K., S. Masuda, P. Mason, J. Pfannstiel, M. Oppizzi, S. Rodriguez-Navarro, A. G. Rondon, A. Aguilera, K. Struhl, R. Reed and E. Hurt (2002). "TREX is a conserved complex coupling transcription with messenger RNA export." Nature **417**(6886): 304-308.

Svejstrup, J. Q. (2004). "The RNA polymerase II transcription cycle: cycling through chromatin." Biochim Biophys Acta **1677**(1-3): 64-73.

Tarn, W. Y., C. H. Hsu, K. T. Huang, H. R. Chen, H. Y. Kao, K. R. Lee and S. C. Cheng (1994). "Functional association of essential splicing factor(s) with PRP19 in a protein complex." Embo j **13**(10): 2421-2431.

Thomas, B. J. and R. Rothstein (1989). "Elevated recombination rates in transcriptionally active DNA." Cell **56**(4): 619-630.

Topisirovic, I., Y. V. Svitkin, N. Sonenberg and A. J. Shatkin (2011). "Cap and cap-binding proteins in the control of gene expression." Wiley Interdiscip Rev RNA **2**(2): 277-298.

Tsai, W. Y., Y. T. Chow, H. R. Chen, K. T. Huang, R. I. Hong, S. P. Jan, N. Y. Kuo, T. Y. Tsao, C. H. Chen and S. C. Cheng (1999). "Cef1p is a component of the Prp19p-associated complex and essential for pre-mRNA splicing." J Biol Chem **274**(14): 9455-9462.

Vander Kooi, C. W., L. Ren, P. Xu, M. D. Ohi, K. L. Gould and W. J. Chazin (2010). "The Prp19 WD40 Domain Contains a Conserved Protein Interaction Region Essential for its Function." Structure (London, England : 1993) **18**(5): 584-593.

Vincent, K., Q. Wang, S. Jay, K. Hobbs and B. C. Rymond (2003). "Genetic interactions with CLF1 identify additional pre-mRNA splicing factors and a link between activators of yeast vesicular transport and splicing." Genetics **164**(3): 895-907.

Wang, Q., L. Zhang, B. Lynn and B. C. Rymond (2008). "A BBP-Mud2p heterodimer mediates branchpoint recognition and influences splicing substrate abundance in budding yeast." Nucleic Acids Res **36**(8): 2787-2798.

Wang, X., Y. Chang, Y. Li, X. Zhang and D. W. Goodrich (2006). "Thoc1/Hpr1/p84 is essential for early embryonic development in the mouse." Mol Cell Biol **26**(11): 4362-4367.

West, M. L. and J. L. Corden (1995). "Construction and analysis of yeast RNA polymerase II CTD deletion and substitution mutations." Genetics **140**(4): 1223-1233.

Wind, M. and D. Reines (2000). "Transcription elongation factor SII." Bioessays **22**(4): 327-336.

Wood, A. and A. Shilatifard (2006). "Bur1/Bur2 and the Ctk complex in yeast: the split personality of mammalian P-TEFb." Cell Cycle **5**(10): 1066-1068.

Zenklusen, D., P. Vinciguerra, J.-C. Wyss and F. Stutz (2002). "Stable mRNP Formation and Export Require Cotranscriptional Recruitment of the mRNA Export Factors Yra1p and Sub2p by Hpr1p." Molecular and Cellular Biology **22**(23): 8241-8253.

Zhang, D. W., J. B. Rodriguez-Molina, J. R. Tietjen, C. M. Nemec and A. Z. Ansari (2012). "Emerging Views on the CTD Code." Genet Res Int **2012**: 347214.

Zhou, H., Q. Liu, T. Shi, Y. Yu and H. Lu (2015). "Genome-wide screen of fission yeast mutants for sensitivity to 6-azauracil, an inhibitor of transcriptional elongation." Yeast **32**(10): 643-655.

Zhou, Q., T. Li and D. H. Price (2012). "RNA Polymerase II Elongation Control." Annual review of biochemistry **81**: 119-143.

Zolotukhin, A. S., W. Tan, J. Bear, S. Smulevitch and B. K. Felber (2002). "U2AF participates in the binding of TAP (NXF1) to mRNA." J Biol Chem **277**(6): 3935-3942.

Abbreviations

°C	degree Celsius
5-FOA	5-fluoroorotic acid
6AU	6-azauracil
Aa	amino acids
APS	ammonium persulfate
Bp	base pairs
C-	carboxy-terminal
CBP	calmodulin binding peptide
cDNA	complementary DNA
ChIP	Chromatin Immunoprecipitation
CTD	C-terminal domain
CTR	C-terminal region
DAPI	4',6-diamidino-2-phenylindole
ddH_2O	double-distilled water
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
dNTP	deoxyribonucleoside triphosphate
Ds	double stranded
DTT	dithiothreitol
E. coli	*Escherichia coli*
EDTA	ethylene diamine tetraacetic acid
H	hour/hours
HA	hemagglutinin
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
IgG	immunoglobulin G
kDa	kilo dalton
L	liters
M	molar
min	minute/minutes
mRNA	messenger ribonucleic acid
mRNP	messenger ribonucleoprotein
N-	amino-terminal
NPC	nuclear pore complex
nt	nucleotides
OD	optical density

ORF	open reading frame
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PDB	Protein Data Bank
PEG	polyethylene glycol
pH	potential of hydrogen
PMSF	phenylmethylsulfonyl fluoride
poly(A)	poly-adenosine
Prp19C	Prp19 complex
qPCR	quantitative PCR
RNAPII	RNA Polymerase II
RNase	Ribonuclease
rpm	revolutions per minute
RRM	RNA Recognition Motif
RT	room temperature
S. cerevisiae	*Saccharomyces cerevisiae*
Ser-2/ S2	Serine 2
Ser-5/S5	Serine 5
S2A	Serine to Alanine
Ser2P/ Ser5P/ Tyr1P/ Thr4P	phosphorylated Serine 2/ Serine 5/ Tyrosine 1/ Threonine 4
SDS	sodium dodecyle sulphate
sec	Seconds
TAP	Tandem Affinity Purification
TCA	trichloro-acetic acid
TEMED	tetramethylethylenediamine
TEV	Tobacco etch virus
TEV-E	TEV-eluate
TREX	transcription and export complex
TRIS	tris(hydroxylmethyl)aminomethane
v/v	volume per volume
w/v	weight per volume
WT	wild-type cells
Tyr-1/ Y1	Tyrosine 1
µ	Micro

Figures

Tables