Function and mechanism of the unconventional ubiquitin-like protein Hub1

K KIRAN KUMAR

A thesis submitted for the partial fulfillment
of the degree of Doctor of Philosophy

Department of Biological Sciences
Indian Institute of Science Education and Research (IISER) Mohali
Knowledge city, Sector 81, SAS Nagar, Manauli PO, Mohali 140306, Punjab, India.

November 2019

Dedicated to my family

Declaration

The work presented in this thesis has been carried out by me under the guidance of Dr. Shravan Kumar Mishra at the Indian Institute of Science Education and Research Mohali. This work has not been submitted in part or full for a degree, a diploma, or a fellowship to any university or institute. Whenever contributions of others are involved, every effort is made to indicate this clearly, with due acknowledgement of collaborative research and discussions. This thesis is a bona fide record of original work done by me and all sources listed within have been detailed in the references.

Date

Place

K Kiran Kumar

In my capacity as the supervisor of the candidate's thesis work, I certify that the above statements by the candidate are true to the best of my knowledge.

Date

Place

Dr. Shravan Kumar Mishra

Acknowledgments

Learning is a lifelong process. In this process, the Ph.D. journey has taught me many beautiful learning experiences. Penning down this acknowledgement, I take the opportunity to express my gratitude to my mentors, colleagues, family and friends, who helped me accomplish this task that seemed prodigious but ended as a gratifying journey.

First and foremost, I would like to thank my research supervisor Dr. Shravan Kumar Mishra, for introducing me to this exciting field of science and for his motivation throughout my Ph.D. I thank him for his exceptionally versatile guidance, valuable suggestions, critical appreciation and keen interest that led to successful compilation of this thesis. I'm grateful to him for endowing me with knowledge from his research experience. His passion and commitment to science are truly inspiring and motivating. I have learned to perform high-quality work under his guidance.

My particular words of thanks will always go Dr. Kavita Babu and Dr. Samrat Mukhophadhyay for their continuous support, motivation, guidance, and helping me to overcome difficulties during this journey. I sincerely express my gratitude for their support.

I want to thank Dr. Sudip Mandal for his inputs and support with the project.

I am thankful to Prof. Anand Kumar Bachhawat for his timely motivation and encouragement.

I express my immense gratitude to all faculties and students of the Department of Biological Sciences, IISER Mohali, for providing me access to departmental facilities and for the scientific discussions during seminars and presentations.

I am grateful to Dr. Pallavi Sharma, Mr. Nagesh Kadam, and Worm lab members for their support with *the C. elegans* project.

Dr. Partima Pandey and Dr. Yogesh were very supportive and helped with scientific discussions. I express my sincere gratitude for their support.

I am thankful to our UBL lab members Dr. Prashanth, Dr. Poonam, Poulami Choudhary, Rakesh Pandian, Anupa, Karan, Bala, Nivedha, Bhargesh, Sumanjith, Manu, Sandra, Amjad, Arundathi, Pravar, Vidhya and other lab members, people with whom I share a lifelong learning experience, for their generous support, care and helpwhile at bench work

Dr. Sivaranjan, Dr. Chiranjeevi, and Dr. Raju Attada are the three people with whom I share my sorrows, happiness, joy, excitement, and cup of tea almost every day. I could not have imagined my stay at IISER without them. They were always beside me whenever I was in need. Without their prompt help, support and concern, it would have not been possible for me to accomplish this task.

I thank Dr. Bapaiah and Dr. Visakhi deeply from my heart for the personal care they bestowed on me and for being my guardians almost.

I am very thankful to our collaborators Dr. Arashdeep and Dr. Kuljeet, for the support with Hub1 project.

I profoundly thank our collaborator, Dr. Jeffrey A. Pleiss, at Cornell University, Ithaca, the USA, for sharing the design of the splicing-sensitive microarray, helping us to perform the experiment and for the analysis of data.

I am very thankful to our collaborators Dr. Ranabir Das and his lab members for the support with Hub1 project

I would like to thank Indian Council of Medical Research (ICMR), Government of India, for the fellowship during my Ph.D. tenure, IISER Mohali, Centre for Protein Science, Design and Engineering (CPSDE), Ministry of Human Resource and Development (MHRD), DST, Government of India, Max Planck Society, Germany, for providing financial support and RNA Society, DBT-CTEP for providing international travel award.

I am thankful to my friends Krishna, Gayathri, Muskan, Kanchan, Manpreet, Daneesh, Vinay, Billa, Ramarao, Sai, Rajkumar, SubbbaRao, Krishnakanth, Budha, Sandeep and many juniors, seniors etc.

No choice of words would suffice to adequately register my love and gratitude to my grandparents and parents-in-law. I thank them for their unconditional love, support, care and encouraging me to follow my dreams. A special thanks to my sister Harika, and my brother Yogesh, for their love and support.

It is impossible to express thanks in any words to my better half Dr. Sree Lalitha for her support and belief in my abilities. I would not have been able to complete much of what I have done and become who I am. Words are short of expressing my gratitude to her. I am thankful to my daughter Purvi for giving me happiness during my last year of a Ph.D.

Last but not least, I am genuinely thankful to the almighty God for giving me such a beautiful life.

All may not be mentioned, but no one is forgotten.

(K Kiran Kumar)

ABSTRACT

Ubiquitin-like proteins (UBLs) control various cellular processes such as protein degradation, DNA repair, autophagy, transcription, RNA splicing, and immune responses. Hub1 is one such UBLs that have been reported to regulate pre-mRNA splicing. However, the function and the mechanism of Hub1 in intron-rich eukaryotes has not been studied yet. We aimed to understand the role of Hub1 in RNA splicing in an intron-rich unicellular eukaryote *Schizosaccharomyces pombe,* and a multicellular eukaryote, *Caenorhabditis elegans.* This study demonstrates a conserved genome-wide role of Hub1 in pre-mRNA splicing in *S. pombe*. Hub1 alters the protein composition of the spliceosome selectively. It promotes splicing of pre-mRNAs that are synthesized faster. It is likely that rapidly synthesizing transcripts require Hub1 to couple transcription with splicing. We identified a functionally conserved Hub1 surface centered at two positively charged residues critical for splicing in *S. pombe*. We also show that Hub1 directly binds to the Kreb's cycle enzyme, Fum1, through another surface. This Hub1 surface also regulates pre-mRNA splicing. Additionally, our study showed the potential role of Hub1 in trans-splicing in the multicellular *C. elegans*. In summary, Hub1 employs multiple surfaces to facilitate binding of specific factors for its function in pre-mRNA splicing. My findings are highly relevant not only for regulatory and tissue-specific gene expression, but also for understanding new mechanisms of alternative splicing.

Synopsis

Function and mechanism of the unconventional ubiquitin-like protein Hub1

Introduction

The genetic information present in DNA is transcribed into precursor-messenger RNA, which is translated into protein (Poetsch and Yoshida, 2018). During transcription, pre-mRNA undergoes several processing steps such as 5'-capping, splicing, and 3' polyadenylation. Most eukaryotic genes are interrupted by introns and exons. RNA splicing is a process in which introns are precisely excised out to generate mature mRNA (Maniatis and Reed, 2002). This process is catalyzed by the spliceosome, a large ribonucleoprotein (RNP) complex consisting of five snRNPs and numerous proteins. Multiple RNA and protein networks serve to align the reactive groups of the pre-mRNA for catalysis in a highly regulated and dynamic manner. Some exons are spliced constitutively, that is, exons are part of every mRNA generated from a given pre-mRNA. Whereas, many exons are alternatively spliced to generate multiple mRNAs from a single pre-mRNA (Will and Luhrmann, 2011). In addition to the core components of the spliceosome, regulatory factors in association with RNA binding proteins (RBP) direct spliceosomes on to pre-mRNAs to perform alternative splicing (Ast, 2004). Regulators like RNP modifying enzymes (ATPases, helicases), RNA binding proteins (RBP), and ubiquitin-like proteins (UBLs) are thereby required for optimal splicing.

Ubiquitin and UBLs proteins contain a globular β-grasp fold and flexible C-terminal tail. They are synthesized as inactive precursors and processed by UBL-specific proteases after conserved di-glycine (GG) motif (Taherbhoy et al., 2012). UBLs control various cellular processes such as protein degradation, DNA repair, chromatin remodeling, cell cycle regulation, endocytosis, and kinase signaling pathways (Turcu et al., 2009). Four UBLs, Ubiquitin, SUMO, Hub1, and Sde2 have been reported to regulate pre-mRNA splicing (Chinnarat and Mishra, 2018). The

ubiquitin-like protein Sde2 is synthesized as a precursor of an N-terminal ubiquitin fold, ($Sde2_{UBL}$), an invariant GG-KGG motif, and a C-terminal domain Sde2-C. Processed Sde2 generates active ^{K}Sde2-C, which is then incorporated into the spliceosome and promotes the splicing of selected introns from a subset of pre-mRNAs (Thakran et al., 2018).

The ubiquitin-like protein Hub1 binds Snu66, a component of the U4/U6.U5 tri-snRNP complex, and promotes alternative splicing of *SRC1* and *PRP5* genes in *S. cerevisiae.* Hub1 also interacts with the DEAD-box helicase Prp5, a regulator of pre-spliceosome assembly, and stimulates its ATPase activity to enhance splicing by relaxing spliceosome's fidelity. Higher levels of Hub1 cause missplicing by allowing usage of suboptimal splicesites (SSs) and branch-point sequence (BPS) (Karaduman et al., 2017). In contrast to *S. cerevisiae,* where Hub1 is not essential for viability, Hub1 is essential for viability in *S. pombe and* human cells. A plausible reason for this difference could be higher prevalence of splicing and intron diversity in higher eukaryotes where Hub1 likely affects splicing of several pre-mRNAs (Dittmar et al., 2002; Wilkinson et al., 2004; Yashiroda and Tanaka, 2004; Mishra et al., 2011; Ammon et al., 2014). Though the role of Hub1 in pre-mRNA splicing has been studied by RNAi-mediated knockdown approaches in human cell lines, the protein's surfaces relevant for its splicing function has not been elucidated in higher eukaryotes (Ammon et al., 2014; Oka et al., 2014). In addition, the mechanism of Hub1 in intron-rich eukaryotes has not been studied.

The HIND (Hub1 Interacting Domain) elements are located in the homologs of RNA splicing factor Snu66/SART1 in most eukaryotes. In plants, HIND is not observed in Snu66 homolog, the absence in Snu66 might be compensated by its presence in another splicing factor Prp38 (Mishra et al., 2011). Interestingly, in certain organisms, including *C. elegans*, HINDs are observed in homologs of both Snu66 and Prp38. At present, implications of more than one splicing factor associating with Hub1 are not clear. It has been shown that HUB1 knockdown worms did not show any splicing defects (Benedetti et al., 2006). Also, Hub1's splicing function is not studied in multi-cellular eukaryotes.

Further, human Hub1 promotes the Fanconi anemia (FA) pathway for repair of DNA interstrand crosslinks (ICLS) by stabilizing FA pathway component FANCI (Oka et al., 2015). This function is proposed to be independent of Hub1's splicing function. In the nematode *Caenorhabiditis elegans,* Hub1 (referred to as UBL-5) is reported to play a role in mitochondrial unfolded-protein response (UPRmt) (Benedetti et al., 2006). These observations possibly indicate that the function of the Hub1 may not be confined to splicing.

Objectives

Therefore, in our current study, we aimed to study the following

I. To understand the role of Hub1 in RNA splicing in an intron-rich organism.
II. To understand the other functions of Hub1.
III. The role of Hub1 in a multi-cellular eukaryote, *C. elegans.*

Results

I. The ubiquitin-like protein Hub1 is a conserved regulator of pre-mRNA splicing and alternative splicing. In *Saccharomyces cerevisiae*, Hub1 binds to the spliceosomal protein Snu66 through its Asp-22 surface (surface I), and to the DEAD-box helicase Prp5 through its His-63 surface (surface II). Hub1 is not essential for viability in *S. cerevisiae* (Dittmar et al., 2002), but the protein becomes essential in *Schizosaccharomyces pombe,* possibly because Hub1 is required for splicing of a larger number of introns. By using microarray and RT-PCR, we show that Hub1 is crucial for pre-mRNA splicing in *S. pombe*. The protein specifically modulates the spliceosome. It possibly facililates assembly of acitivated spliceosome by removing a specific set of essential splicing factors. Strikingly, however, mutations in both surfaces resulted in only mild growth and splicing defects, thereby suggesting that Hub1 might work in pre-mRNA splicing thorugh other unidentified surface(s).

By using approaches of NMR, mutagenesis, and complementation of a *hub1*-knockout *S. pombe* strain, we have identified a new Hub1 surface, surface III,

containing Arg-9 and Arg-41 residues (referred to as hub1-R941). This surface is functionally conserved in all eukaryotes and alteration of this surface resulted in temperature sensitive growth and splicing defects in *S. pombe*. Further, SPCPB16A4.06C, a Schizosaccharomyces specific gene, was identified as a high-copy suppressor of the temperature sensitive *hub1* surface III mutant. We further show that Hub1 promotes splicing of the transcripts that are synthesized faster. It is likely that rapidly synthesizing transcripts require Hub1 to couple transcription with splicing. Altogether, this part of my study revealed that, despite its small globular size of only 73 amino acids (~8.4 kDa), Hub1 employs multiple surfaces for its function in pre-mRNA splicing. These surfaces facilitate binding of specific (known and certain unknown) factors.

II. The ubiquitin-like protein Hub1 functions in pre-mRNA splicing by binding non-covalently to the spliceosomal proteins Snu66 and Prp5. We have found that *Schizosaccharomyces pombe* Hub1 also binds to a mitochondrial enzyme of the citric acid cycle, fumarase (Fum1). The enzyme binds to a conserved surface of Hub1 centred at a solvent-exposed tryptophan residue. This surface is absent in *S. cerevisiae* Hub1; however, an introduction of tryptophan at analogous position restores its affinity with Fum1. Hub1-Fum1 complex precipitates *in vitro* indicating mutually inhibitory activities of the two partners. In support of potential inhibitory activities, elevated levels of both Fum1 and *fum1ΔSS* (cytosolic fumarase) are more toxic in *hub1-W47G* mutant cells compared to the wild-type *S. pombe* cells. *hub1-W47G* mutant cells show genetic interaction with splicing factor mutants and also exhibit inefficient excision of introns from selected pre-mRNAs. *fum1Δ* mutant also shows genetic interaction with certain splicing factor mutants. Higher levels of *fum1ΔSS* is more toxic in splicing factor mutants, compared to the wild-type cells. Thus, Fum1 protein which is not imported to mitochondria could regulate pre-mRNA splicing through Hub1. Since fumarase is frequently mutated in multiple diseases in humans, phenotypes of some of the mutants may be due to their influence on Hub1-dependent pre-mRNA splicing.

III. The ubiquitin-like protein Hub1 is a conserved member of the UBL family but functions distinctly from ubiquitin. The protein modifies spliceosomes non-covalently by binding to HIND (Hub1 INteraction Domain) containing pre-mRNA splicing factor Snu66 and promotes alternative splicing. However, the splicing function of Hub1 remains unexplored in multicellular eukaryotes. Hub1 has been reported to play a role in mitochondrial unfolded protein response (UPRmt) in *Caenorhabditis elegans* and worms depleted of HUB1 did not show any defects in pre-mRNA splicing. In the present study, we show that *C. elegans* Hub1 rescues *S. pombe hub1Δ* lethality. It also complements *S. pombe* Hub1 splicing function. We further show that *C. elegans* Hub1 interacts with HIND-containing splicing factors Snu66 and Prp38. The mode of Hub1 interaction through salt bridge formation with Snu66 is conserved in *C. elegans*. Hub1 also associates with the components of spliceosome. It is further shown that *HUB1* knockout worms were lethal at Larval 3 stage, which suggests an essential role of HUB1 in development. HUB1 transcript levels are stage-specific and at L3-L4 stage levels are higher than other stages of *C. elegans* life cycle. By using splicing-sensitive microarray, we show that *HUB1* knockout led to accumulation of outron-containing transcripts. These results suggest that Hub1 might play role in RNA trans-splicing.

Conclusion

This study demonstrates a conserved genome-wide role of Hub1 in pre-mRNA splicing in *S. pombe*. Hub1 alters the protein composition of the spliceosome selectively. It promotes splicing of pre-mRNAs that are synthesized faster. It is likely that rapidly synthesizing transcripts require Hub1 to couple transcription with splicing. We identified a functionally conserved Hub1 surface centered at two positively charged residues critical for splicing in *S. pombe*. We also show that Hub1 directly binds to the Kreb's cycle enzyme, Fum1, through another surface. This Hub1 surface also regulates pre-mRNA splicing. This Hub1 surface may promote associations with other factors, part from Fum1, to perform its splicing function. Additionally, our study showed the potential role of Hub1 in trans-splicing in the multicellular, *C. elegans*.

In summary, Hub1 employs multiple surfaces to facilitate binding of specific factors for its function in pre-mRNA splicing. My findings are highly relevant not only for regulatory and tissue-specific gene expression, but also for understanding new mechanisms of alternative splicing.

References:

1. Schneider-Poetsch, T. & Yoshida, M. Along the Central Dogma—Controlling Gene Expression with Small Molecules. *Annu. Rev. Biochem.* **87**, 391–420 (2018).
2. Maniatis, T. & Reed, R. An extensive network of coupling among gene expression machines. *Nature* **416**, 499–506 (2002).
3. Will, C. & Lührmann, R. Spliceosome structure and function. Cold Spring Harb Perspect Biol doi: 10.1101/cshperspect.a003707. *Cold Spring Harb Perspect Biol* **3**, (2011).
4. Ast, G. How did alternative splicing evolve? *Nat. Rev. Genet.* **5**, 773–782 (2004).
5. Taherbhoy, A. M., Schulman, B. A. & Kaiser, S. E. Ubiquitin-like modifiers. *Essays Biochem.* **52**, 51–63 (2012).
6. Reyes-Turcu, F. E., Ventii, K. H. & Wilkinson, K. D. Regulation and Cellular Roles of Ubiquitin-Specific Deubiquitinating Enzymes. *Annu. Rev. Biochem.* **78**, 363–397 (2009).
7. Chanarat, S. & Mishra, S. K. Emerging Roles of Ubiquitin-like Proteins in Pre-mRNA Splicing. *Trends Biochem. Sci.* **43**, 896–907 (2018).
8. Thakran, P. *et al.* Sde2 is an intron-specific pre- mRNA splicing regulator activated by ubiquitin-like processing. *EMBO J.* **37**, 89–101 (2018).
9. Wilkinson, C. R. M. *et al.* Ubiquitin-like Protein Hub1 Is Required for Pre-mRNA Splicing and Localization of an Essential Splicing Factor in Fission Yeast. **14**, 2283–2288 (2004).
10. Yashiroda, H. & Tanaka, K. Hub1 is an essential ubiquitin-like protein without functioning as a typical modifier in fission yeast. *Genes to cells* **9**, 1189–1197 (2004).
11. Mishra, S. K. *et al.* Role of the ubiquitin-like protein Hub1 in splice-site usage and alternative splicing. *Nature* **474**, 173–178 (2011).
12. Ammon, T. *et al.* The conserved ubiquitin-like protein Hub1 plays a critical role in splicing in human cells. *J. Mol. Cell Biol.* **6**, 312–323 (2014).
13. Oka, Y. *et al.* UBL5 is essential for pre- mRNA splicing and sister chromatid cohesion in human cells. *EMBO Rep.* **15**, 1330–1330 (2014).
14. Benedetti, C., Haynes, C. M., Yang, Y., Harding, H. P. & Ron, D. Ubiquitin-like protein 5 positively regulates chaperone gene expression in the mitochondrial unfolded protein response. *Genetics* **174**, 229–239 (2006).
15. Karaduman, R., Chanarat, S., Pfander, B. & Jentsch, S. Error-Prone Splicing Controlled by the Ubiquitin Relative Hub1. *Mol. Cell* **67**, 423-432.e4 (2017).

Table of Contents

Abbreviations

AD	Activation domain
APEX2	Ascorbate peroxidase
AS	Alternative splicing
ATP	Adenosine triphosphate
BD	Binding domain
BP	Biotin-phenol
BPS	Branch-point sequence
CBs	Cajal bodies
CD	Circular dichroism
CEC	Capping enzyme complex
Co-IP	Co-immunoprecipitation
CPSF	Cleavage and polyadenylation specificity factor
CSTF	Cleavage stimulating factor
CTD	Carboxy-terminal domain
DDR	DNA damage response
DSEs	Downstream sequence elements
ESCs	Embryonic stem cells
ESEs	Exonic splicing enhancers
FA	Fanconi anemia
FOA	Fluororotic acid
HIF	Hypoxia-inducible transcription factor
HIND	Hub1 Interacting Domain
HLRCC	Hereditary leiomyomatosis and renal cell cancer
HR	Homologous recombination
HU	Hydroxyurea
ICLS	DNA interstrand crosslinks
IDRs	Intrinsically disordered regions
IPTG	Isopropyl β- d-1-thiogalactopyranoside
KH	K-homology
m7G	7-methylguanosine
MoaD	Molybdopterin

MTS	Mitochondrial signal sequence
NEDD8	Neural precursor cell expressed, developmentally down-regulated 8
NMR	Nuclear magnetic resonance
PAB2	Polyadenylated binding factor
PAP	Polyadenylate polymerase
PAS	PolyA signals
PCI	Phenol: chloroform: isoamyl alcohol
PE	Phosphatidylethanolamine
PHD	Prolyl hydroxylase enzymes
PPT	Polypyrimidine tract
Pre-mRNA	Precursor messenger RNA
RBP	RNA binding proteins
RNAP II	RNA polymerase II
RNP	Ribonucleoprotein complex
RRM	RNA recognition motif
RT-PCR	Reverse transcriptase polymerase chain reaction
RT-qPCR	Real-time quantitative PCR
RT-qPCR	Real-time quantitative PCR
SL	Splice leader
snRNAs	Spliceosomal small nuclear RNAs
SnRNPs	Small-nuclear ribonucleoprotein
SR	Serine-arginine
SRPKs	SR protein kinases
SSs	Splicesites
SUMO	Small ubiquitin-related modifier
TCA	Tricarboxylic acid cycle
TCA	Trichloroacetic acid
ThiS	Thiamine biosynthesis
TMG	Trimethylguanosine
UBLs	Ubiquitin-like proteins
UPRmt	Mitochondrial unfolded-protein response
WB	Western blot

WT	Wild-type

Chapter 1

Introduction

1.1 Transcription and pre-mRNA processing

Eukaryotes transcribe their genomes by three highly related nuclear RNA polymerases I, II, and III. The transcription of protein-coding genes into mRNA is carried out by RNA polymerase II (RNAP II). During this process, pre-mRNA has to undergo several processing steps, such as 5'-capping, splicing, and polyadenylation. The mature RNA can be exported from the nucleus to the cytoplasm for translation into proteins.

The general transcription initiation factors TFIIA, -B, -D, -E, -F, and -H assemble on the core promoter with RNAP II to initiate transcription. Transcription initiation begins with TFIID recognizing and binding tightly to the TATA box elements. TFIID-TATA-box guides remaining general initiation factors and RNAP II to form the pre-initiation complex or PIC. Further, PIC complex is completed by the joining of TFIIH and TFIIE complexes (Bentley, 1999). Chromatin remodeling enzymes provide the accessibility and open the densely packed chromatin to allow productive transcription. Among these enzymes, acetyl- or methyltransferases are crucial to post-translationally modify nucleosomal histones for subsequent reorganization of chromatin structure (Ho and Crabtree, 2010). DNA methylation and histone modifications remodel chromatin structures. Upon histone acetylation (histone 3 at K9 and K14, histone 4 at K16) and methylation (histone 3 at K4), RNAP II exits from its promoter to start productive mRNA synthesis.

The largest subunit of RNAP II contains a unique domain at its C-terminus, termed as the carboxy-terminal domain (CTD). In humans, the CTD contains 52 heptapeptide repeats of Tyr-Ser-Pro-Thr-Ser-Pro-Ser residues. These residues act as a regulatory interaction platform for various mRNA processing factors, thereby couple transcription to subsequent mRNA maturation (Hirose and Manley, 2000).

CTD promotes each step of pre-mRNA processing events. RNAP II containing unmodified CTD is referred to as RNAP IIA, and with hyperphosphorylated CTD is referred as RNAP IIO. The phosphorylation of CTD is a prerequisite for the transition of RNAP II for preinitiation complex to stable elongation complex. While RNAP II is mostly phosphorylated at serine S5 during transcrption initiation, during elongation phase the serine S2 is predominantly phosphorylated. In case of paused or terminating RNA pol II complexes, serine S7 is found to be phosphorylated (Dahmus, 1996). Upon RNA synthesis, the 5'- capping enzyme complex (CEC) directly and selectively binds to the nascent chain and the CTD S5 phosphorylation mark. This association guides transfer of 7-methylguanosine (m7G) cap to the emerging 5' end of the newly synthesized transcript (Anderson and Parker, 1998). The coupling among the capping and transcription machinery ensures that nascent pre-mRNA is protected from degradation (Hirose and Manley, 2000; Maniatis and Reed, 2002). The serine S2 phosphorylated RNAP II CTD recruits auxiliary splicing factors like ASF/SF2 or SC35, which promote pre-mRNA splicing (Millhouse and Manley, 2005). Pre-mRNA splicing (detailed in section 1.2) is coupled to transcription elongation through interaction between the splicing machinery and the transcription elongation factor, TAT-SF1. Another transcription elongation factor, TFIIS associates with high molecular mass complex containing RNAP II and several splicing factors. The SR protein family of splicing factors binds to specific exon sequences, known as splicing enhancers and helps to recruit splicing machinery to the 5' splice site (SS), and 3' splice site (SS). SR proteins couple pre-mRNA splicing with transcription (Maniatis and Reed, 2002). Different promoters recruit SR protein family members, which in turn promotes alternative splicing (Cramer et al., 1999). After splicing reactions are completed, the nascent transcripts undergo polyadenylation co-transcriptionally. Most protein-coding genes contain polyadenylation elements [polyA signals (PAS)] followed by U- or GU-rich downstream sequence elements (DSEs) at their 3' region. These signals are required for both 3'-end cleavage and polyadenylation of pre-mRNA (Proudfoot, 2011). Most of the human genes possess more than one poly (A) site, thereby generates transcript isoforms with alternative 3' ends through alternative polyadenylation. Cleavage and polyadenylation specificity factor (CPSF) and cleavage stimulating factor (CSTF) are two-multi-peptide

complexes that recognize the PAS and DSEs respectively to promote cleavage and polyadenylation. Polyadenylate polymerase (PAP) in association with polyadenylated binding factor (PAB2) adds polyadenosine tail to free 3' end of the mRNA (Elkon et al., 2013). Coupling between transcription and different pre-mRNA processing steps is prerequisite for optimal gene expression (Maniatis and Reed, 2002).

1.2 RNA Splicing

Most eukaryotic genes are interrupted by introns which are precisely excised out by a process called RNA splicing. Some exons are spliced constitutively, that is exons are part of every mRNA generated from a given pre-mRNA. Whereas, many exons are alternatively spliced to generate multiple mRNAs from a single pre-mRNA species (Will and Luhrmann, 2011). Alternative splicing is prevalent in higher eukaryotes, and it increases the complexity of the organism (Nilsen and Graveley, 2010). In the human genome splicing is essential as almost all the genes contain introns. The average human transcript contains around 8.8 short exons (120 nucleotides) and long introns with average size of more than 5400 nucleotides (Sakharkar et al., 2004). In contrast, in *Saccharomyces cerevisiae (S. cerevisiae)* only three percent of genes carry introns, of which only six genes with two introns (Barrass and Beggs, 2003). Although the fission yeast *Schizosaccharomyces pombe (S. pombe)* genome with 43% of intron-containing genes, but both yeasts contain short introns which average length around 40-75 nucleotides (Ast, 2004). Despite the difference in splicing prevalences between lower eukaryotes *S. cerevisiae* and humans but basic biochemistry of pre-mRNA splicing is conserved (Nilsen, 2003). This process carried out by a large multi-megadalton ribonucleoprotein complex known as spliceosome via two transesterification reactions (Moore et al., 1993). Schematics of RNA splicing is depicted in the (Figure. 1.1).

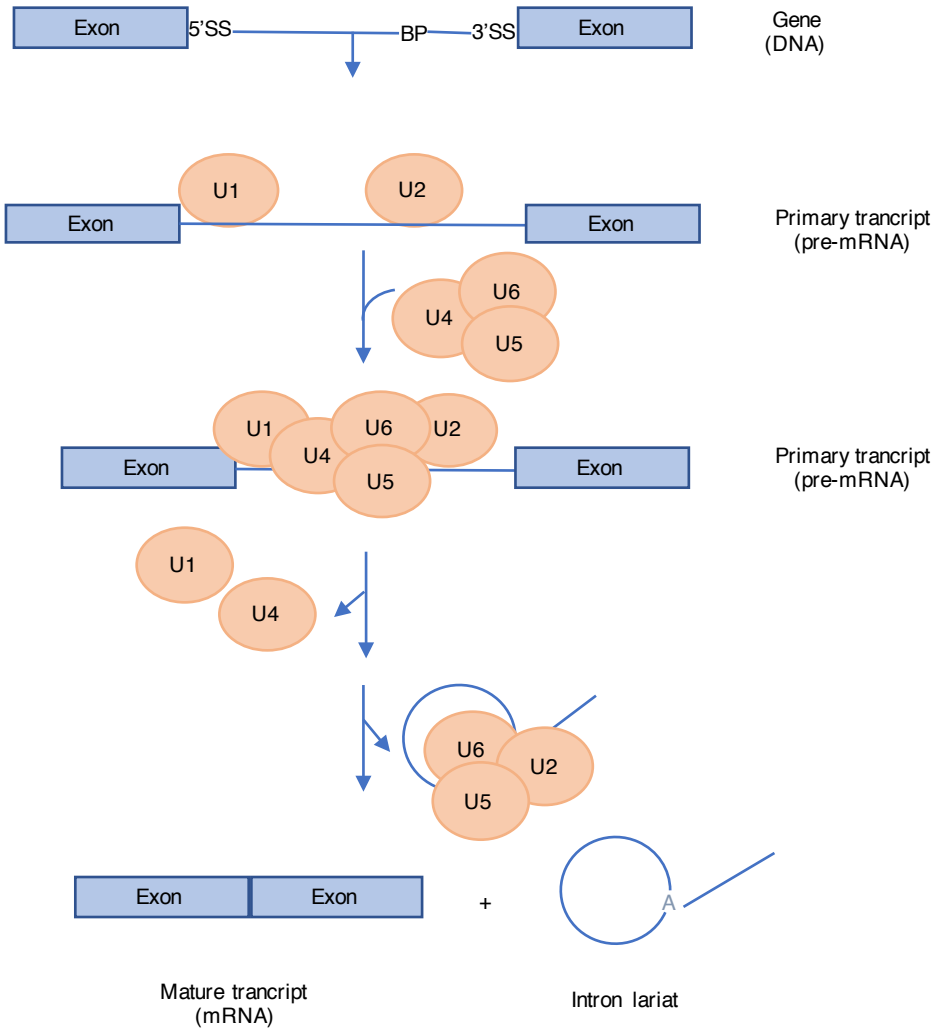


Figure 1.1: RNA splicing

Schematic of intron excision and two adjacent exons ligation. The steps shown are the assembly of small nuclear ribonucleoprotein complexes (snRNPs), active spliceosome and formation of mature transcript and intron lariat (adapted from Dvinge et al., 2016).

1.3 Components of the spliceosome

The mechanistic similarities between pre-mRNA splicing and autocatalytic excision of group II introns led to the hypothesis that pre-mRNA splicing reaction may be RNA mediated catalysis (Weiner, 1993; Wise, 1993). Spliceosome consists of five spliceosomal small nuclear RNAs (U1, U2, U4, U5, and U6 snRNAs), their respective ribonucleoproteins snRNPs and associated non-snRNP protein complexes (Nilsen, 1994). All non-U6 snRNAs are transcribed by RNA polymerase II

and exported to the cytoplasm for maturation (Urlaub et al., 2001). In contrast, U6 snRNA is transcribed by RNA polymerase III and matures within the nucleus (Karijolich and Yu, 2010).

During spliceosome assembly, the 5' SS is initially recognized via base pairing with U1 snRNA. In metazoans, U1 snRNA/5' SS interactions are stabilized by U1-70 K, U1-C and serine-arginine (SR) protein family (Will and Luhrmann, 2006). SR proteins are general metazoan splicing factors containing essentially an arginine/ serine-rich (RS) domain, which helps in stabilizing U6/5' SS and U5/5' SS base pairing (Shen and Green, 2007). The U5 protein 220k / Prp8 association with 5' SS, 3' SS, branchpoint sequence (BPS), and PPT (Grainger and Beggs, 2005) helps in aiding first and second step of splicing (Will and Luhrmann 2011). In metazoans, branch point adenosine is initially recognized by auxiliary splicing factor SF1/BBP in co-ordination with heterodimer U2AF. U2AF subunits, U2AF65 and U2AF35 help in recruiting U2 snRNP to BPS upstream of 3' SS. The stable association with BPS is supported by twelve U2 snRNP proteins including A'', B'', SF3a and SF3b complex to form a prespliceosomal complex A (Will and Luhrmann, 2011). U5 snRNP consists of eight U5 proteins, DEAD-box helicase hPrp28, the GTPase Snu114, DExD/H-box protein hBrr2, and the multi-domain protein Prp8. The complementary domains in U4 and U6 snRNAs help in RNA base-pairing to form U4/U6-snRNP complex proteins (hPrp3, hPrp4, hPrp31, CypH, and 15.5 K) (Schneider et al., 2002). The U5 snRNP is pre-assembled and associates with U4/U6 snRNPs to generate U4/U6.U5 tri-snRNP complex, it undergoes conformational and structural rearrangements to form the catalytic core complex of the spliceosome (Gottschalk et al., 1999; Wahl et al., 2009).

In addition to the snRNP complexes, the spliceosome contains non-snRNA containing RNPs. The major non-snRNP is PRP19/ CDC5L or Nineteen complex (NTC) in *S. cerevisiae*. It associates with the spliceosome during precatalytic complex B formation (Ajuh et al., 2000). PRP19/ CDC5L complex consists of at least seven proteins; hPRP19 (Prp19), CDC5L (Cef1), PRL1 (Prp46) and SPF27 (Snt309) are conserved, the subunits human AD002, CTNNBL1 (β-catenin-like 1) and HSP73

are not found in yeast. Immunodepletion and complementation experiments show that the Prp19 complex is essential for the first catalytic step of splicing (Makarova, 2004). Prp19, a member of U-box and WD-repeat family of E3 ubiquitin ligase, is an essential factor required for activation and stabilization of spliceosome (Vander Kooi et al., 2010). Prp19 complex promotes K63-linked ubiquitination of Prp3 (U4 snRNP component) and increases affinity to interact with Prp8 (U5 snRNP component). Thereby facilitates stable assembly of U4/U6.U5 snRNP complex. Prp3 is reversibly deubiquitinated by Usp4 and its substrate targeting factor Sart3 to promote ejection of U4 proteins from the spliceosome during maturation of its active site. This reversible modification pathway is required for efficient pre-mRNA splicing (Song et al., 2010). PRP19/CDC5L complex stays with the spliceosome till the splicing reaction is completed and disassociates from the post-spliceosomal complex with the release of lariat intermediate.

In addition to PRP19/CDC5L complex, a large number of splicing factors transiently associate with the spliceosome to mediate rearrangements during splicing (Wahl et al., 2009). During this reaction, a large number of RNA-RNA and RNA-protein interaction networks undergo rearrangements that are mediated by co-ordinated action of various enzymes. U4/U6-U5 tri-snRNP requires unwinding of U4/U6 duplex snRNA for the spliceosome activation, which requires RNA helicase Brr2, Prp8, and GTPase subunit Snu114 (U5 components of tri snRNP) (Hacker et al., 2008). Other helicases such as Sub2 and Prp5 helps in recognition of BPS by U2 snRNP during the formation of pre-spliceosome (Chang et al., 2013). Another group of non-snRNP proteins is serine/arginine rich-SR proteins and hnRNP (heterogeneous nuclear ribonucleoproteins). Both the family of proteins are central regulators of alternative splicing. SR proteins are a large family of RNA binding proteins with RS domain (arginine/serine-rich domain). Numerous studies indicate that multiple SR proteins bind to exonic splicing enhancers (ESEs) to promote splice-site selection. It also promotes exon inclusion and exon skipping depending on the interacting sequences on pre-mRNA (Zhou and Fu, 2013). hnRNPs contain K-homology (KH) domains accompanied by the more divergent functional domains such as RGG boxes, glycine-, acidic-, proline-, rich domains. Different hnRNPs bound to distinct pre-

mRNA sequences can enhance or repress the spliceosome assembly (Krecic and Swanson, 1999).

1.4 Assembly of the spliceosome

Spliceosome assembly occurs by the sequential binding of spliceosomal snRNPs and several other splicing factors (Staley and Woolford, 2009). The exon/intron architecture of the gene determines the usage of the splicesites via spliceosome across exon or intron. Majority of splice site recognition in higher eukaryotes occurs across the exon, whereas in lower eukaryotes recognition occurs across the intron (Berget, 1995). In the event of introns less than 200-250 nucleotides, spliceosome assembles across the introns (Figure. 1.2) (Fox-Walsh et al., 2005).

In the case of cross-intron assembly, U1 snRNP interact with 5' SS and non-snRNP factors, SF1 and U2AF bind to BPS and polypyrimidine tract (PPT) respectively. (i.e., the E complex). In later steps of assembly, U2 snRNP stably associates with BPS forming pre-spliceosome (i.e., the A complex). U4/U6.U5 tri-snRNP associates with the pre-mRNA to generate pre-catalytic B complex. Major RNA-RNA and RNA-protein rearrangements lead to disassociation of U1 and U4 snRNPs to form activated spliceosome (i.e., the B^{act} complex). Further catalytic activation by DEAH-box helicase Prp2, generates B* complex which carries out first transesterification reaction, it yields C complex which undergoes additional rearrangements to perform second transesterification reaction. The spliceosome then disassociates, and with additional remodeling, releases snRNPs for further rounds of splicing (Will and Luhrmann, 2011).

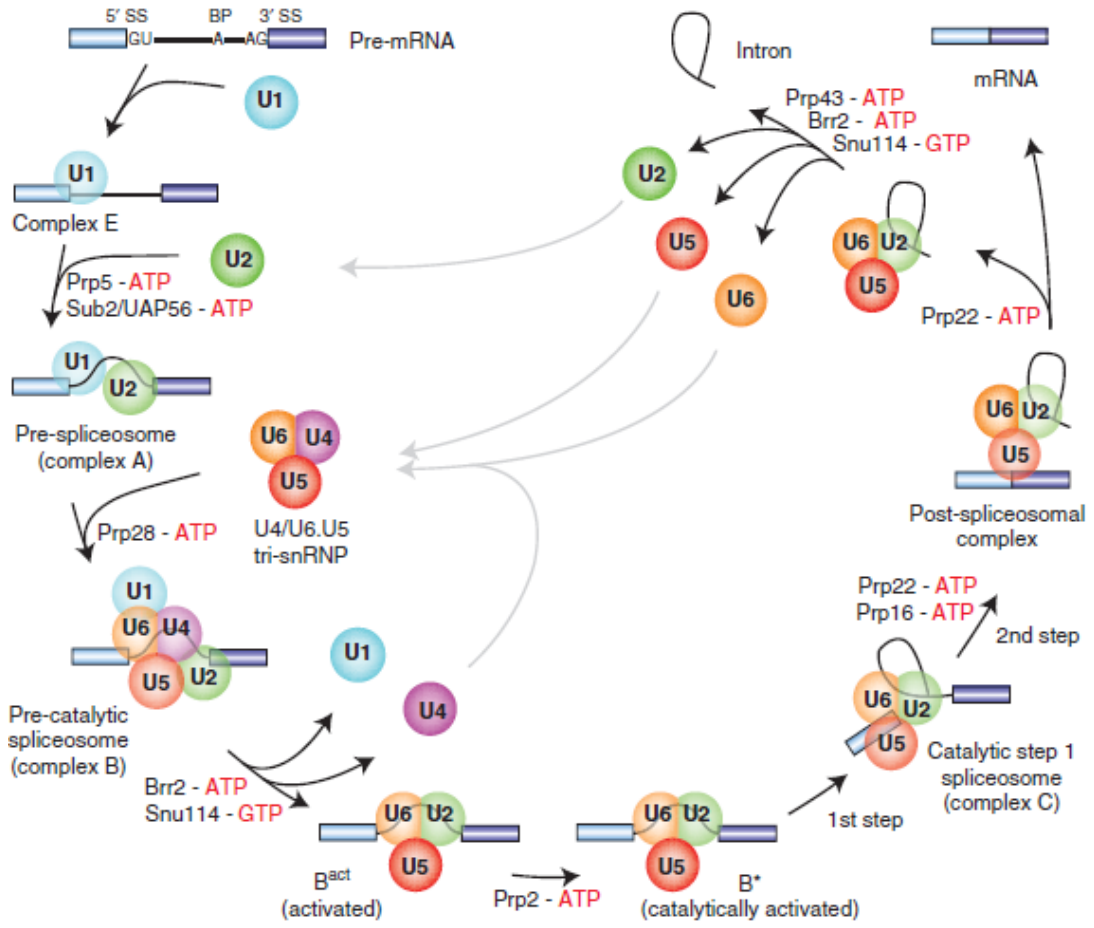


Figure 1.2: Assembly and disassembly of the spliceosome

The snRNPs (indicated by circles) recognize the intronic sequences and mediate spliceosome assembly. Various spliceosomal complexes such as DExH/D-box RNA ATPases/helicases Prp5, Sub2/UAP56, Prp28, Brr2, Prp2, Prp16, Prp22, and Prp43, or the GTPase Snu114, act to facilitate the conformational changes during spliceosome formation. (adapted from Will and Luhrmann, 2011).

1.5 Alternative splicing

Alternative splicing is a process by which multiple isoforms are generated from a single pre-mRNA (Soller, 2006). Initially, strength of 5' SS, BPS, and 3' SS determines the efficiency of splicing. In *S. cerevisiae,* alternative splicing occurs rarely. Whereas, in humans, around 95% of intron-containing transcripts undergo alternative splicing (Pan et al., 2008). The occurrence of alternative splicing correlates with the high degree degeneracy of splice sites found in metazoan genome (Ast, 2004). Several types of alternative splicing (AS) events (schematics of

alternative splicing events are depicted in the (Figure. 1.3)) involve exon skipping, alternative 3' SS selection, alternative 5' SS selection, and intron retention (Keren et al., 2010). In addition to strength of splice sites, cis-acting RNA sequence elements can also govern the splice site usage. Splicing enhancer elements recruit SR proteins such as SC35, ASF/SF2, and SRp20 to positively regulate the splicing (Sheprad and Hertel, 2009). In contrast, hnRNPs bound to distinct pre-mRNA sequences can repress or enhance the spliceosome assembly (Krecic and Swanson, 1999). Another layer of regulation of alternative splicing is through post-translational modification of SR proteins and hnRNPs in human cells. The arginine-serine rich (RS) domain-containing SR protein ASF2/SF2 is phosphorylated by SR protein kinases (SRPKs) and Clk/Sty kinases. Phosphorylation of ASF2/SF2 by SRPKs at N-terminal part of RS domain is essential for its recruitment to the splicing speckles. Whereas, phosphorylation of ASF2/SF2 by Clk/Sty kinases at C-terminal part of RS domain is required for its release from speckles. Phosphorylation of ASF2/SF2 is critical for its subcellular distribution (Ngo et al., 2005). Both brain and testis of chimpanzee, mouse, and human tissues show large number of alternative splicing events and a large number of splicing factor genes are differentially expressed. Among the splicing factor genes, snRNPs and SRPKs are most highly differentially expressed in a particular tissue (Grosso et al., 2008). As mostly splicing occurs co-transcriptionally, strength of the promoters, RNA polymerase processivity, pausing and post-translational histone modifications regulate alternative splicing events (Kornblihtt et al., 2013; Luco et al., 2011). The transcription factor positioned in the promoter of processing transcript associate with SR protein. This associated transcription factor binds with the transcription initiation complex to promote loading of SR proteins with the Pol II moving complex, which inturn regulates alternative splicing (Cramer et al., 1997). Mutations in cis-acting elements or trans-acting factors leads to abberant splicing and abnormal protein production in many diseases. Hence reagents regulating alternative splicing could be used as potential therapeutics (Garcia-Blanco et al., 2004).

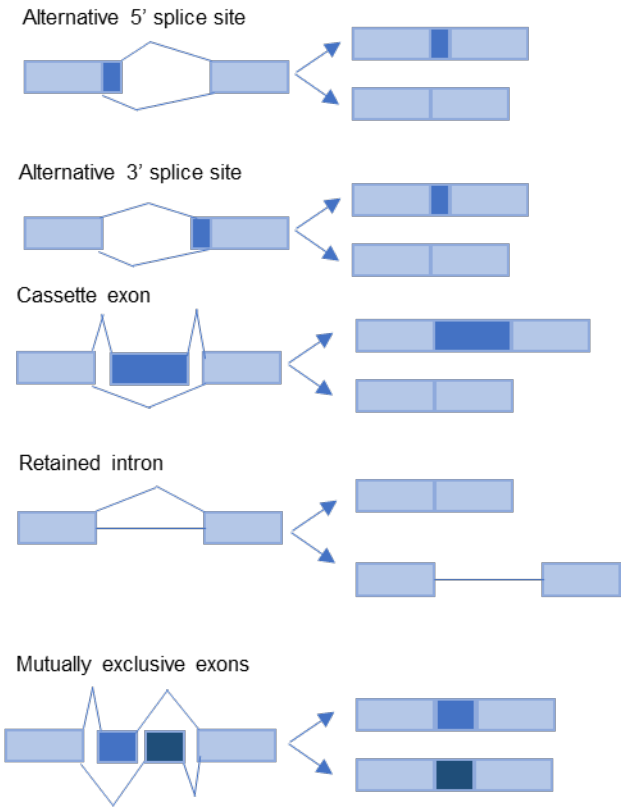


Figure 1.3: Events of alternative splicing

Schematics of alternative splicing events. Light blue: constitutive exon sequence that always forms part of mature mRNA; mid-blue or dark blue: an alternative sequence that can be included or excluded in the mature mRNA. (adapted from Dvinge et al., 2016).

1.6 Ubiquitin and ubiquitin-like proteins

The origin of ubiquitin-like proteins (UBLs) begins from prokaryotic biosynthetic pathways. Proteins involved in molybdopterin (MoaD) and thiamine biosynthesis (ThiS) pathways are structural homologs of UBLs. MoaD and ThiS do not post-translationally modify proteins or other macromolecules. These proteins are modified transiently themselves carrying sulfur atoms at the C-termini, and promoting sulfur transfer in their pathways (Taherbhoy et al., 2012).

Ubiquitin and UBLs contain a globular β-grasp fold and flexible C-terminal tail, are synthesized as inactive precursors and are processed by UBL-specific proteases after conserved di-glycine (GG) motif (Taherbhoy et al., 2012). Structure of UBLs are shown below (Figure. 1.4). In the early 1980s, post-translational modification of

proteins by ubiquitin targets for intracellular proteolysis. Later it was shown that ubiquitination and its reversal, deubiquitination acts as a targeting signal to control various cellular processes, such as protein degradation, chromatin remodeling, immune system regulation, endocytosis/trafficking, DNA repair, and kinase signaling pathways (Turcu et al., 2009).

Ubiquitination is catalyzed by the sequential action of three enzymes such as a ubiquitin-activating enzyme, E1; a ubiquitin-conjugating enzyme, E2; and a ubiquitin ligase, E3. The Ubiquitin-activating enzyme activates ubiquitin through an ATP-dependent step forming a thiol-ester linkage between the C-terminus of ubiquitin and the active cysteine of the E1. The activated ubiquitin is later transferred to active-site cysteine of the ubiquitin-conjugating enzyme, which together with ubiquitin-ligase transfers the ubiquitin to a lysine residue in the target protein (Turcu et al., 2009).

The UBL protein SUMO (small ubiquitin-related modifier) is conserved from yeast to humans. Although human SUMO-1 protein is 18% identical to ubiquitin, both contain a globular β-grasp fold. The conjugation of SUMO to protein substrates referred to as SUMOylation requires an enzyme cascade similar to ubiquitin (Jentsch and Pyrowolakis, 2000). SUMO (Smt3 in *S. cerevisiae*) is conjugated to the lysine residues on the protein substrates through an enzyme cascade. It involves SUMO E1 (AOS1/Uba2) and E2 Ubc9 in association with E3 ligases like PIAS or the nucleoporin RanBP2 (Flotho and Melchior, 2013). SUMOylation can be reversed by SUMO-specific proteases (SENP1-7 in humans) (Mukhopadhyay and Dasso, 2007). SUMOylation of proteins helps in stabilization of protein substrates and localization to subcellular complexes, which are implicated in various cellular pathways such as transcription, chromatin remodeling, DNA repair, nucleo-cytoplasmic transport, mitosis and stress response (Muller et al., 2001).

Another ubiquitin-like protein, Rub1 (NEDD8 in metazoans, Neural precursor cell expressed, developmentally down-regulated 8), shows 53% identity with ubiquitin. Processing of NEDD8 precursor is catalyzed by UCH-L3 and NEDP1, conjugation to the protein substrates is catalyzed through E1 enzyme (ULA1 and UBA3), E2

conjugating enzyme Ubc12 and Ube2F in co-ordination with several E3 ligases like Rbx1, Dcn1. The targets of NEDD8 were cullin family of proteins (CUL4A in human cells). Neddylation is critical for regulating cullin function (Rabut and Peter, 2008). The covalent NEDD8 conjugation can be reversed by NEDD8 isopeptidases in a process known as deneddylation (Cope and Deshaies, 2003). The neddylation and deneddylation of cullin-based ubiquitin ligases are essential for the activity and its stability (Bosu and Kipreos, 2008).

Another ubiquitin-like protein of the ATG family, Atg8 and Atg12, play crucial roles in the autophagy pathway. Unlike most UBLs, Atg12 is synthesized in a mature form containing C-terminal glycine residue ready for activation by an E1 enzyme, Atg7. The E2 enzyme, Atg10, transfers Atg12 to Atg5. The Atg5-Atg12 complex acts as the E3 enzyme for the Atg8. Atg8 is unique among UBLs, and it is the only UBL gets ligated to lipid PE (phosphatidylethanolamine). Atg8 ligation is catalyzed by E1 activating enzyme Atg7, E2 conjugating enzyme Atg3 in co-ordination with E3 Atg5-Atg12 complex. Atg8 can be removed from PE by Atg4 (Taherbhoy et al., 2012).

ISG15 is another UBL, which conjugates to the proteins by a process known as ISGylation, similar to ubiquitylation. The enzyme cascade for ISGylation consists of UbeL1 (E1), UbcH8 (E2), and Herc5 or EFP as E3 ligases. However, the molecular mechanism of its action remains unclear (Sgorbissa and Brancolini, 2012). In interferon-stimulated human cells, ISG15 is highly induced and mediates ISGylation of host and viral proteins. ISG15 is a component of the innate immune response, which exhibits antiviral activity against several types of viruses (Durfee et al., 2010).

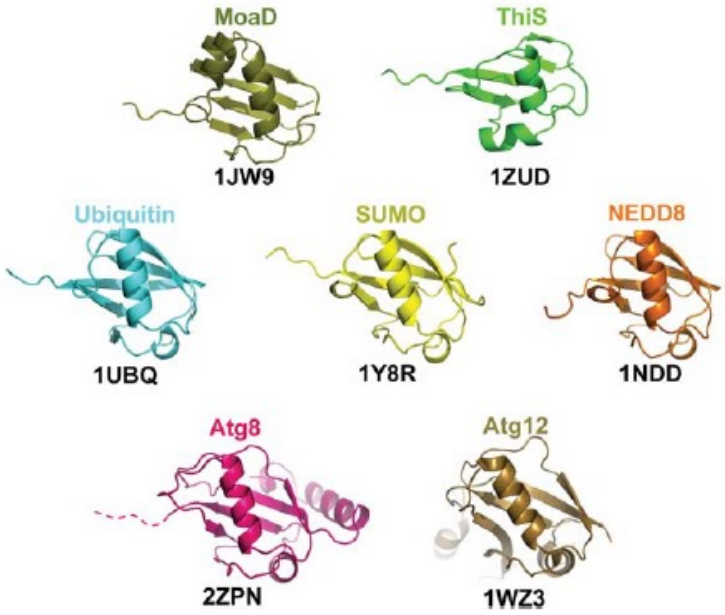


Figure 1.4: Structure of UBLs

Structural comparison between ubiquitin-like proteins with their respective PDB codes. The Atg8 C-terminus is represented by a broken line (not observed in the structure). (adapted from Taherbhoy et al., 2012).

Table 1.1: Summary of UBLs

UBL	Function	Substrates	Comments
Ub	26S proteasomal degradation, chromatin remodeling, immune system regulation, endocytosis/trafficking	Short-lived proteins, histones, signaling proteins, membrane proteins (ion channels and receptors)	Diversity in functions due to its ability to form polyubiquitin chains using one of its seven lysine residues or N-terminus and also from the recognition of numerous distinct forms of ubiquitin
SUMO	DNA repair, nuclear transport, transcription	Ran GTPase-activating protein I, activators and repressors of transcription	SUMO primarily has nuclear-related functions
NEED8	Activation of cullin E3 Ub ligases	Cullins	Cullin E3s are the largest class of ubiquitin ligases that regulate cell

			cycle
Atg12	Autophagy	Atg5	Atg12-Atg5 act as E3 for Atg8 pathway
Atg8	Autophagy	PE (phosphatidylethano-lamine)	Mammals have Atg8 homologs

(adapted from Taherbhoy et al., 2012).

Table 1.2: The following list represents the details of the protein and the methods used to obtain their structures.

UBL (PDB-ID)	Name	Method	Organism
1JW9	MoaD	X-diffraction	*E. coli*
1ZUD	ThiS	X-diffraction	*E. coli*
1UBQ	Ubiquitin	X-diffraction	*Homo sapiens*
1Y8R	SUMO	X-diffraction	*Homo sapiens*
1NDD	NEDD8	X-diffraction	*Homo sapiens*
2ZPN	Atg8	X-diffraction	*S. cerevisiae*
1WZ3	Atg12	X-diffraction	*Arabidopsis thaliana*

1.7 Ubiquitin and ubiquitin-like proteins in pre-mRNA splicing

Ubiquitin and three other ubiquitin-like proteins such as SUMO, Sde2, and Hub1 are implicated in pre-mRNA splicing. Ubiquitin binds both covalently and non-covalently to the spliceosomal core Prp8 protein. These non-degradative associations imply its role in splicing regulation. Further, Prp3 and Prp31 get ubiquitinated by Prp19 which is essential for spliceosomal activation. Similar to ubiquitin, SUMOylation of many splicing factors, serine/arginine (SR) proteins and members of hnRNP family upon stress response were observed. SUMO conjugation to spliceosomal proteins strengthens the spliceosomal assembly and enhances splicing efficiency. Crosstalk between SUMOylation and ubiquitination may also exist as Prp3 gets modified by both SUMOylation and ubiquitination (Chanarat and Mishra, 2018).

Another ubiquitin-like protein Sde2 is synthesized as a precursor of an N-terminal ubiquitin fold, $Sde2_{UBL}$; an invariant GG_KGG motif; and a C-terminal domain, Sde2-

C. Processed Sde2 generates active $_{KGG}$Sde2-C which is then incorporated into the spliceosome and promotes splicing of selected introns from a subset of pre-mRNAs. In *S. pombe* the Sde2 splicing functions have been attributed to its ability to recruit Cactin/Cay1 to the spliceosome. *S. pombe* Sde2 mutants show defects in telomeric silencing and genomic stability (Thakran et al., 2018). The ubiquitin-like protein Hub1 role is summarized in the section below.

Table 1.3: List of UBLs in pre-mRNA splicing

UBLs	Mode of action	Function
Ubiquitin	Covalent conjugation for proteasomal degradation, protein-protein interactions and non-covalent binding for protein-protein interactions	Assembly and disassembly of the spliceosome
SUMO	Covalent conjugation with splicing factor to facilitate protein-protein interaction with splicing factors containing SUMO binding motifs	Spliceosome assembly
Hub1	Non-covalent binding for mediating protein-protein interaction and stimulates ATPase activity of RNA helicase Prp5	Spliceosome assembly, activation, and splicing fidelity
Sde2	Processed Sde2 facilitates association with the spliceosome and spliceosome specific factors	Spliceosome assembly and specificity

(adapted from Chanarat and Mishra, 2018).

1.8 The Ubiquitin-like protein Hub1

Hub1 (also called UBL5) is a highly conserved member of the UBL family of proteins in eukaryotes. But, different from UBLs, Hub1 interacts with substrates non-covalently as it lacks any enzyme cascade and the UBLs' characteristic di-glycine motif, both essential for covalent conjugation (Luders et al., 2003; Yashiroda and Tanaka, 2004; Mishra et al., 2011). Hub1 associates with the spliceosomal protein Snu66, a protein of the U4/U6.U5 small nuclear-ribonucleic particle (tri-snRNP) to promote alternative splicing (Wilkinson et al., 2004; Mishra et al., 2011). Hub1 is not essential for viability in *S. cerevisiae* (Dittmar et al., 2002). It is not required for general splicing, but moderately affects the splicing of few transcripts. It also functions in usage of non-canonical 5' SS in splicing *SRC1* and *PRP5* transcripts,

both the transcripts are alternatively spliced in *S. cerevisiae* (Mishra et al., 2011). Hub1 interacts with the DEAD-box helicase Prp5, a regulator of pre-spliceosome assembly, and stimulates its ATPase activity to enhance splicing and relaxing fidelity of the spliceosome. High levels of Hub1 causes missplicing by tolerating suboptimal SSs and BPS. Importantly, Prp5 itself is regulated through Hub1-dependent negative feedback loop to have a check on missplicing (Karaduman et al., 2017). Since Hub1 is essential, temperature-sensitive *hub1-1* (*hub1I42S*) cells with growth at 30°C and restricted growth at 37°C were identified in *S. pombe*. The *hub1-1* cells showed genetic interaction with Snu66 and splicing defects at 37°C (Yashiroda and Tanaka, 2004). Other temperature-sensitive *hub1-4* (*hub1M70K*) cells which grow at 30°C and growth-restricted at 32°C were identified. Similar to *hub1-1*, *hub1-4* cells showed defects in pre-mRNA splicing at the restrictive temperature (Wilkinson et al., 2004).

In human cell lines, Hub1 is essential for viability, possibly because it affects splicing of many intron-containing genes (Ammon et al., 2014). It is a part of the spliceosomal complex B just before the activation but after incorporation of tri-snRNP (Makarova et al., 2004). Hub1 interacts with coilin (core component of Cajal bodies), colocalizes with Cajal bodies (CBs), a subnuclear domain where assembly and/or modification of spliceosomal components occur (Sveda et al., 2013). Other than snu66, Hub1 makes additional contacts with the SR-protein kinases Cdc2/Cdc28-like kinases (Kantham et al., 2003). In this process, Hub1 exhibits splicing independent function in stabilizing FANCI protein (Oka et al., 2015). As of now, Hub1-Snu66 interaction is well established and studied in molecular details in humans and *S. pombe* (Mishra et al., 2011; Ammon et al., 2014). In human cell lines, *hub1* mutant with defective Snu66 binding complemented growth and splicing defects. Firstly, complementation might occur due to overexpression of *hub1* mutant from plasmid-borne constructs. Secondly, Hub1 might interact with Snu66 through other surfaces, and/or thirdly, Hub1 makes additional contacts with other proteins of the spliceosome. Though the role of Hub1 in pre-mRNA splicing has been studied by RNAi-mediated knockdown approaches in human cell lines, the protein's surfaces relevant for its splicing function has not been elucidated in higher eukaryotes

(Ammon et al., 2014; Oka et al., 2014). In addition, the mechanism of Hub1 in intron-rich eukaryotes has not been studied.

The HIND (Hub1 Interacting Domain) elements are located in the homologs of RNA splicing factor Snu66/SART1 in most eukaryotes. In plants, HIND is not observed in Snu66 homolog, the absence in Snu66 might be compensated by its presence in another splicing factor Prp38 (Mishra et al., 2011). Interestingly, in certain organisms, including *C. elegans*, HINDs are observed in homologs of both Snu66 and Prp38. At present, implications of more than one splicing factor associating with Hub1 are not clear. It has been shown that HUB1 knockdown worms did not show any splicing defects (Benedetti et al., 2006). Also, Hub1's splicing function is not studied in multi-cellular eukaryotes.

Further, human Hub1 promotes the Fanconi anemia (FA) pathway for repair of DNA interstrand crosslinks (ICLS) by stabilizing FA pathway component FANCI (Oka et al., 2015). This function is proposed to be independent of Hub1's splicing function. In the nematode *Caenorhabiditis elegans,* Hub1 (referred to as UBL-5) is reported to play a role in mitochondrial unfolded-protein response (UPRmt) (Benedetti et al., 2006). These observations possibly indicate that the function of the Hub1 may not be confined to splicing.

Objectives

Therefore, in our current study, we aimed to study the following

I. To understand the role of Hub1 in RNA splicing in an intron-rich organism.
II. To understand the other functions of Hub1.
III. The role of Hub1 in a multi-cellular eukaryote, *C. elegans.*

Chapter 2

Materials and Methods

2.1 Materials

2.1.1 Chemicals and plastic wares

All chemicals used in the study were obtained from commercial sources. Which were of either of analytical or molecular biology grades. Media components, fine chemicals, and reagents were purchased from Sigma Aldrich, USA, HiMedia, India, Merck. Ltd, USA, Difco, USA and Formedium, UK. All plastic wares used for yeast, bacteriological and molecular biological works were procured from Abdo's lab tech, India and Tarsons, India.

2.1.2 Molecular biology reagents

Enzymes (restriction enzymes, T4 DNA ligase, Alkaline phosphatase (CIP), Phusion DNA polymerase, Pfu turbo, Taq DNA polymerase, Vent DNA polymerase, and other modifying enzymes, buffers, dNTPs, DNA and protein molecular weight markers were purchased from New England Biolabs, Invitrogen, Sigma Aldrich, USA. Gel-extraction and plasmid miniprep kits were obtained from Bioneer, Korea. RNA isolation kits were obtained from Invitrogen, USA.

2.1.3 Antibodies and antibody-coupled beads

The antibodies anti-Myc (polyclonal), anti-FLAG M2 (clone M2), anti-haemagglutinin (HA, clone HA-7), anti-HA (polyclonal),anti-rabbit-HRP and anti-mouse-HRP (raised in Goat) were obtained from Sigma Aldrich. Antibody-coupled beads Anti-HA-rabbit (H6908), anti-MYC (A7470) and anti-FLAG (A2220) were also from the same source. Rabbit serum enriched with polyclonal antibody (MERCK), against bacterially purified 6xHis-Ce Hub1 was used.

2.1.4 Media

All the media buffers and stock solutions were prepared using Millipore distilled water and were sterilized as recommended, either by autoclaving at 15 psi pressures at 121°C for 15 minutes or by using membrane filters (HiMedia, India) of pore size 0.22-0.45μ (for heat-labile compounds)

i. Luria-Bertani (LB) medium: 10 g tryptone, 5 g yeast extract, and 10 g sodium chloride dissolved in 1 liter of distilled water and autoclaved

ii. Luria-Bertani (LB) Agar: 10 g tryptone, 5 g yeast extract and 10 g sodium chloride in 0.5-liter water and 20 g agar dissolved in 0.5 liters of distilled water and autoclaved separately and were mixed together to make LB Agar. Desires antibiotics (ampicillin (100 μg/ml) or kanamycin (50 μg/ml) were added as per requirements.

iii. Yeast-extract (YES) medium: 5 g yeast extract, 2 g casamino acids, 30 g glucose, 0.1 g adenine, 0.1 g uridine, 0.1g leucine, 0.1 g histidine and 20 g agar per liter of medium. When selecting for the kan-MX4 marker using G418/geneticin resistance (Sigma G-5013), G418 plates were made by dissolving 200 μg/ml. For nat-NT2 cassette, Nat plates with 75 μg/ml and for Hph-NT1, hygromycin plates with 100 μg/ml

iv. Synthetic defined (SD) media: Per liter: 6.7 g yeast nitrogen base, 2 g required supplement dropout mixtures for auxotrophies (e.g. leucine and uracil), 20 g glucose, and for making plates 20 g agar was used. Expression constructs under the nmt promoter were induced in the absence of thiamine, and 5 μg/ml thiamine was used to repress the promoter

v. Edinburg minimal medium (EMM) medium: Per liter: 3 g potassium hydrogen phthalate, 2.2 g Na_2HP0_4, 5 g ammonium chloride, 2 g required supplements mixture for auxotrophies are added as required, 20 g glucose,

20 ml of 50x salt stock, 2 ml of 500x vitamins and minerals stock, and for plates 20 g agar was used. Expression constructs under the nmt promoter were induced in the absence of thiamine, and 5 μg/ml thiamine was used to repress the promoter.

2.1.5 Buffers and stock solutions

i. Cell lysis buffer: 2 % Triton X-100, 1% SDS, 100 mM NaCl, 10 mM Tris-Cl (pH-8.0), and 1mM EDTA.

ii. 10x Tris-EDTA (TE) buffer (pH 8.0): 10 mM Tris-HCl (pH 8.0), and 1mM EDTA

iii. 50x TAE: Per Liter: 242 g Tris base, 57.1 ml glacial acetic acid, and 100 ml of 500 mM EDTA (pH 8.0). Autoclaved and stored at room temperature.

iv. High Urea (HU) buffer: 8M urea, 5% SDS, 200 mM Tris pH 6.8, 1 mM EDTA, with bromophenol blue, and 1.5% dithiothreitol (DTT) was added before use.

v. 30% Acrylamide mixture: 29% acrylamide and 1% methylbisacrylamide

vi. Resolving Gel (12%): 1.25 ml 1.5 M Tris-HCL (pH 8.8), 50 μl 10% SDS, 2 ml 30% acrylamide solution, 50 μl 10% ammonium persulphate (APS), 5 μl Tetramethylethylenediamine (TEMED), and 1.7 ml water.

vii. Stacking Gel (4%): 0.5 ml 0.5 M Tris-HCl (pH 6.8), 20 μl 10% SDS, 0.26 ml 30% acrylamide solution, 40 μl 10% APS, 4 μl TEMED, and 1.2 ml water.

viii. 10x SDS buffer (pH 8.3): Per liter: 30 g Tris base, 144 g glycine, and 10 g SDS

ix. 20x MOPS buffer (pH 7.7): 50 mM MOPS, 50 mM Tris base, 0.1% SDS, and 1 mM EDTA.

x. 10x Semi-dry transfer buffer: Per liter: 29.3 g Glycine, 58.2 g Tris base, and 4 g SDS. For transfer 1x buffer with 5% methanol used.

xi. 10x Tris-Buffered Saline (TBS) buffer (pH 7.6): 24.2 g Tris base, and 80 g NaCl per liter:

xii. For washing, 1x TBS with 0.1% tween 20 was used.

xiii. IP Cell-lysis buffer: 150 mM Nacl, 5 mM $Mgcl_2$, 50 mM Tris-Cl (pH-7.5), 5%-10% glycerol, 1% Triton X-100, 1 mM PMSF, two phosphatase inhibitors tablets per 50 ml buffer (Roche) and one protease inhibitors tablet per 50 ml buffer (Thermo Scientific).

xiv. IP Wash buffer 1: Cell lysis buffer with 1% Triton X-100 and 1 mM PMSF

xv. IP Cell lysis buffer: 1% Triton X-100, 1 mM PMSF and without protease inhibitor.

xvi. SORB (pH 8): 100 mM lithium acetate, 10 mM Tris-HCl (pH 8), 1 mM EDTA (pH 8), 1 M sorbitol, filter sterilized and stored at room temperature.

xvii. 40% PEG: 100 mM lithium acetate, 10 mM Tris-HCl (pH 8), 1 mM EDTA (pH 8), 40% PEG, filter sterilized and stored at 4°C.

xviii. Salmon sperm DNA (10 mg/ml): It was denatured at 100°C for 10 minutes, cooled on ice and stored at -20°C. For yeast competent cell preparation, 40 μl of denatured salmon sperm DNA was used per 50 ml of culture.

xix. AES buffer composition: 50mM NaoAC (PH: 5.3), 10 mM EDTA, 1% SDS.

xx. Lysis buffer for purification of GST-tagged proteins: 100 ml Phosphate Buffered Saline, 5% Glycerol, 1 mM PMSF, 1 tablet of protease inhibitor (Add Lysozyme - 1-4 mg/ml at the time of use)

xxi. High salt buffer**:** 500 mM NaCl, 50 mM Tris (pH - 7.5), 1 mM MgSO4, 1 mM EDTA, 1 mM PMSF

xxii. Elution buffer**:** 100 mM NaCl, 50 mM Tris (pH 8), 10 mM Glutathione

xxiii. 10 mM Phosphate Buffered Saline (pH 7.4): 0.26 g KH2PO4, 2.17 g Na2HPO4-7H20, 8.71 g NaCl, 0.201 g KCl, 800 mL dH20. Adjust pH to 7.4 and bring volume to 1 L with dH20.

xxiv. Ni-NTA lysis buffer: 50 mM NaH2PO4, 300 mM NaCl, 10 mM Imidzole, 1 mM PMSF, 1 tablet of protease inhibitor (for 40 ml buffer), adjust PH to 8.0 using NaOH (Add lysozyme 4 mg/ml at the time of lysis)

xxv. Ni-NTA wash buffer: 50 mM NaH2PO4, 300 mM NaCl, 20 mM Imidzole, 1 mM PMSF, adjust PH to 8.0 using NaOH

xxvi. Ni-NTA elution buffer: 50 mM NaH2PO4, 300 mM NaCl, 250 mM Imidzole, adjust PH to 8.0 using NaOH

2.1.6 *S. pombe* strains and plasmids

A complete list of strains and plasmids utilized in this study is given in Tables 2.1 and 2.2.

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

Strain	Relevant genotype	Reference
SP1	*h- ade6-M216, leu1, ura4-D18*	Tanaka's lab

Strain	Relevant genotype	Reference
SP9	PEM2 (E-MAP study)	Mishra et al., 2011
SP10	PEM2 *hub1-I42S::Nat-NT2*	This study
SP13	*JY741 Δhub1::aur1R pUR19-hub1+*	Yashiroda and Tanaka, 2004
SP20	*h+ JY741 Δsde2:: Nat-NT2*	Thakran et al., 2018
SP37	*PEM2 hub1-I42S::Nat-NT2 prp19-6HA::KanMX4*	This study
SP42	*h+ cdc5-6HA::KanMX4*	This study (made by Prashant A. Pandit)
SP44	*h+ prp19-6HA::KanMX4*	This study (made by Prashant A. Pandit)
SP75	PEM2 *hub1-I42S::Nat-NT2 cdc5-6HA::KanMX4*	This study
P2	PEM2 *hub1H63L::Nat-NT2*	This study
P3	PEM2 *hub1D22A::Nat-NT2*	This study
K2	PEM2 *hub1D22A H63L::Nat-NT2*	This study
P8	*h+ prp19-6HA::KanMx4 mug161-9MYC::HphNT1*	This study
P10	PEM2 *hub1-I42S::Nat*-NT2 *prp19-6HA::KanMx4 mug161-9MYC::HphNT1*	This study
P46	PEM2 *hub1R9R41A::Nat-NT2*	This study
P50	*h- hub1W47G::Nat-NT2*	This study
	h+ Δcwf18::KanMX4	BIONEER
	h+ Δiss9::KanMX4	BIONEER
PN1	Wild-type	This study (made by Nivedha Balaji)
PN2	*hub1W47G::Nat-NT2*	This study (made by Nivedha Balaji)
PN3	*Δcwf18::KanMX4*	This study (made by Nivedha Balaji)
PN4	*Δcwf18::KanMX4 hub1W47G::Nat-NT2*	This study (made by Nivedha Balaji)
PN5	Wild-type	This study (made by Nivedha Balaji)
PN6	*hub1W47G::Nat-NT2*	This study (made by Nivedha Balaji)

Strain	Relevant genotype	Reference
PN7	*Δiss9::KanMX4*	This study (made by Nivedha Balaji)
PN8	*Δiss9::KanMX4 hub1W47G::Nat-NT2*	This study (made by Nivedha Balaji)
PN9	*h- Δfum1::Nat-NT2*	This study (made by Nivedha Balaji)
PN23	Wild-type	This study (made by Nivedha Balaji)
PN24	*Δfum1::Nat-NT2*	This study (made by Nivedha Balaji)
PN25	*Δiss9::KanMX4*	This study (made by Nivedha Balaji)
PN26	*Δfum1::Nat-NT2 Δiss9::KanMX4*	This study (made by Nivedha Balaji)
PN47	Wild-type	This study (made by Nivedha Balaji)
PN48	*Δfum1::Nat-NT2*	This study (made by Nivedha Balaji)
PN49	*Δcwf18::KanMX4*	This study (made by Nivedha Balaji)
PN50	*Δfum1::Nat-NT2 Δcwf18::KanMX4*	This study (made by Nivedha Balaji)

(PN1, PN2, PN3, PN4), (PN5, PN6, PN7, PN8), (PN47, PN48, PN49, PN50), and (PN23, PN24, PN25, PN26) are individual spores (tetratype(T)) from a single tetrad separated into rows on solid media by microdissection.

Table 2.2: *S. cerevisiae* strain used in this study

Strain	Relevant genotype	Reference
PJ69-7A	*trp1-901 leu2-3,112 ura3-53 his3-200 gal4 gal80 GAL1::HIS3 GAL2- ADE2 met2::GAL7-lacZ*	James et al., 1996

Table 2.3: Plasmid clones used in this study

Plasmid No	Name	Description	Reference
D008	*pET28a 6xHIS–ScHUB1*	*S. cerevisiae HUB1* in pET28a	This study
D025	*pET28a 6xHIS–hub1*	*S. pombe hub1* in pET28a	This study
D026	*pGBUC1-hub1*	*S. pombe hub1* in pGBDUC1	This study
D027	*pGBUC1-hub1G47W*	*S. pombe hub1-G47W* in pGBDUC1	This study
D029	*pGBDUC1-snu66*	*S. pombe snu66* in pGBDUC1	This study
D030	*pGADC1-fum1*	*S. pombe fum1* in pGADC1	This study
D036	*pPROEx-fum1*	*S. pombe fum1* in pROPEx	This study
D121	pREP81x-*Ce HUB1*	*C. elegans HUB1* in pREP81X	This study (made by Pallavi Sharma)
D122	pREP81x-*Sp hub1*	*S. pombe hub1* in pREP81x	This study
D123	pGBDUC1-*Ce HUB1*	*C. elegans HUB1* in pGBUC1	This study (made by Pallavi Sharma)
D127	pET28a-*6xHIS-Ce HUB1*	*C. elegans HUB1* in pET28a	This study (made by Pallavi Sharma)
D128	pGBDUC1-*Ce SNU66(HIND)*	*C. elegans HIND* (aa 1-80) in pGBUC1	This study (made by Pallavi Sharma)
D131	pGADC1-*Ce HUB1*	*C. elegans HUB1* in pGDAC1	This study (made by Pallavi Sharma)

Plasmid No	Name	Description	Reference
D150	pGEX-5x-1-*Ce SNU66(HIND)*	*C. elegans HIND* (aa 1-80) in pGEX-5x-1	This study (made by Pallavi Sharma)
D151	pGADC1-*Ce SNU66(HIND)*	*C. elegans* HIND (aa 1-80) in pGADC1	This study (made by Pallavi Sharma)
D163	pGBDUC1-*Ce HUB1D22A*	*C. elegans HUB1D22A* in pGDAC1	This study (made by Pallavi Sharma)
D169	pGEX-5x-1-*Ce PRP38(HIND)*	*C. elegans HIND (aa 291-320)* in pGEX-5x-1	This study (made by Pallavi Sharma)
D133	pET28a-*hub1-DD*	*S. pombe hub1* with DD extensions at C-terminus in pET28a	This study
D172	pGADC1-*Ce SNU66(HINDR62A)*	*C. elegans HIND-R62A* in pGADC1	This study
D173	pREP81x-*Ce HUB1I42S*	*C. elegans HUB1-I42S* in pREP81X	This study
D174	*pREP81x-hub1I42S*	*S. pombe hub1-I42S* in pREP81x	This study
D280	*p49.26-3XFLAG-CeHUB1*	*C. elegans HUB1 in p49.26*	This study
D288	*pREP81x-hub1-DD*	*S. pombe hub1-DD* in pREP81x	This study
D289	*pREP81x-hub1-GG*	*S. pombe hub1-GG* in pREP81x	This study
D299	*pHub1-3MYC-hub1*	*S. pombe hub1* with 3MYC tag at N-terminus in pREP81x with *hub1* promoter	This study
D373	*pFA6a.Nat-NT2-hub1*	*S. pombe hub1* genomic DNA including 500 bp upstream of start codon of the gene and 100bp downstream of the stop codon of the gene, followed by Nat-	This study

Plasmid No	Name	Description	Reference
		NT2 cassette and 101-600bp downstream of the stop codon of the gene	
D374	*pFA6a.Nat-NT2-hub1H63L*	Similar to D373 construct with H63L mutation	This study
D375	*pFA6a.Nat-NT2-hub1D22A*	Similar to D373 construct with D22A mutation	This study
D376	*pFA6a.Nat-NT2-hub1D22A H63L*	Similar to D373 construct with D22A H63L mutation	This study
D377	*pFA6a.Nat-NT2-hub1R9A R41A*	Similar to D373 construct with R9A R41A mutation	This study
D407	*pHub1-Ce HUB1*	*C. elegans HUB1* in pREP81x with *HUB1* promoter	This study
D413	*pREP81x-3MYC-hub1I42S*	*S. pombe hub1I42S* in pREP81x with 3MYC tag at N-terminus	This study
D415	*pHub1-3MYC-hub1-R9R41A*	*S. pombe hub1-R9R41A* with 3MYC tag at N-terminus in pREP81x with *HUB1* promoter	This study
D416	*pHub1-3MYC-hub1-W47G*	*S. pombe hub1-W47G* with 3MYC tag at N-terminus in pREP81x with *HUB1* promoter	This study
D417	*pREP81x-3MYC-hub1*	*S. pombe hub1* in pREP81x with 3MYC tag at N-terminus	This study
D420	*pREP81x-3MYC-hub1-W47G*	*S. pombe hub1-W47G* in pREP81x with 3MYC tag at N-terminus	This study

Plasmid No	Name	Description	Reference
D476	*pFA6a.Nat-NT2-hub1W47G*	Similar to D373 construct with W47G mutation	This study
D477	*pREP3x-fum1Δss*	*S. pombe fum1* lacking N-terminal mitochondrial localization sequence in pREP3x	This study
D480	*pREP4x-fum1Δss-6HA*	*S. pombe fum1* lacking N-terminal mitochondrial localization sequence in pREP4x with 6HA tag at C-terminus	This study
D481	*pREP4x-fum1-6HA*	*S. pombe fum1* in pREP4x with 6HA tag at C-terminus	This study
D482	*pREP3x-fum1Δss-6HA*	*S. pombe fum1* lacking N-terminal mitochondrial localization sequence in pREP3x with 6HA tag at C-terminus	This study
D483	*pREP3x-fum1-6HA*	*S. pombe fum1* in pREP3x with 6HA tag at C-terminus	This study
D489	*pHub1-3MYC-hub1-R9A*	*S. pombe hub1-R9A* with 3MYC tag at N-terminus in pREP81x with *HUB1* promoter	This study
D490	*pHub1-3MYC-hub1-R40A*	*S. pombe hub1-R40A* with 3MYC tag at N-terminus in pREP81x with *HUB1* promoter	This study
D491	*pHub1-3MYC-hub1-R41A*	*S. pombe hub1-R41A* with 3MYC tag at N-terminus in pREP81x with *HUB1* promoter	This study
D492	*pHub1-3MYC-hub1-R9A R40A*	*S. pombe hub1-R9A R40A* with 3MYC tag at N-terminus in	This study

Plasmid No	Name	Description	Reference
		pREP81x with *HUB1* promoter	
D493	*pHub1-3MYC-hub1-R40A R41A*	*S. pombe hub1-R40A R41A* with 3MYC tag at N-terminus in pREP81x with *HUB1* promoter	This study
D494	*pHub1-3MYC-hub1-R9A R40A R41A*	*S. pombe hub1-R9A R40A R41A* with 3MYC tag at N-terminus in pREP81x with *HUB1* promoter	This study
D495	*pHUB1-1XFLAG-APEX2*	1XFLAG-APEX2 tag at N-terminus in pREP81x with *HUB1* promoter	This study
D496	*pHUB1-1XFLAG-APEX2-hub1*	*S. pombe hub1* with 1XFLAG-APEX2 tag at N-terminus in pREP81x with *HUB1* promoter	This study
D497	*pHUB1-1XFLAG-APEX2-hub1-DD*	*S. pombe hub1-DD* with 1XFLAG-APEX2 tag at N-terminus in pREP81x with *HUB1* promoter	This study
D498	*pHub1-Ce HUB1 R9A K41A*	*C. elegans HUB1 R9A K41A* in pREP81x with *HUB1* promoter	This study
D499	*psod2-3MYC-gsod2*	*S. pombe sod2* with 3MYC tag at N-terminus in pREP81x with *sod2* promoter and terminator	This study
D500	*pmug37-3MYC-gsod2*	*S. pombe sod2* with 3MYC tag at N-terminus in pREP81x with *mug37* promoter and *sod2* terminator	This study
D501	*pmug37-3MYC-gmug37*	*S. pombe mug37* with 3MYC tag at N-	This study

Plasmid No	Name	Description	Reference
		terminus in pREP81x with *mug37* promoter and terminator	
D502	*psod2-3MYC-gmug37*	*S. pombe mug37* with 3MYC tag at N-terminus in pREP81x with *sod2* promoter and *mug37* terminator	This study
D503	*pGADC1-hub1*	*S. pombe hub1* in pGADC1	This study
D504	*pGADC1-hub1R9A R41A*	*S. pombe hub1R9A R41A* in pREP81x	This study
D505	*pREP1X-DR9*	*S. pombe SPCPB16A4.06C* highy copy suppressor clone	This study
	pREP3X-SPCPB16A4.06C	*S. pombe SPCPB16A4.06C in pREP3X*	This study
	pGBUC1-ScHUB1G47W	*S. cerevisiae hub1-G47W* in pGBDUC1	This study
	pREP3x-fum1 K379A N381A-6HA	*S. pombe fum1 K379A N381A* in pREP3x	This study

2.1.7 Primers

All primers used in this study are mentioned in table 2.3

Table 2.4: List of primers used in this study (F-forward primer; R-reverse primer) (Int-intron; Ex-exon)

Number	Name	Sequence (5'-3')
SKM_PR 13	*act1* F	CCCCTAGAGCTGTATTCCC
SKM_PR 14	*act1* R	CAGTGGTACGACCAGAGG
SKM_PR 317	*SPBC354.07c Ex5 R*	GTGCCATAAACCACAGTCCTTC
SKM_PR 318	*SPBC354.07c Ex4 F*	GAGTTTCCTGAGCTATTCGTAAC
SKM_PR 495	*rap1 Ex1 F*	AAAAACTTTGAACATATTAGGGG
SKM_PR 502	*rap1 Ex3 R*	CTTATAATGTTGCCGCCAGG

SKM_PR 635	*pla1 Ex1 F*	GGCGTTACTCTTGAAGTCGC
SKM_PR 636	*pla1 Ex2 R*	CGTAAAATTCTTGGGTAGGCC
SKM_PR 637	tcg1 Ex2 F	CGTCCTGTCCAAGAGCAACC
SKM_PR 638	tcg1 EX3 R	ACTTGAAGGAGCTGTTGCGG
SKM_PR 1235	*gnd1 exon 4 F*	CCGTACAACTTCCAGAGTTGACGAG
SKM_PR 1236	*gnd1 exon 5 R*	CAAATTCCTCAAGGGAGTGAGCACC
SKM_PR 1290	*MYC-F*	AGCTGTCGACCGAGATGGGTGAACA AAAG
SKM_PR 1463	*mug161 Ex2 F*	CAAGCGTACAACTAGTGCGG
SKM_PR 1464	*mug161 Ex3 R*	AATGGACTCTGGCAAACCAGC
SKM_PR 1971	*sod2-R*	GAGAATGATTGATGTGACCACCG
SKM_PR 1972	*mug37-R*	CTTCAAAGTAGTAGAGGATGAC
SKM_PR 2280	*hri2 Ex2 F*	GCGGATGCTTTTAACTGCTTTG
SKM_PR 2281	*hri2 Ex3 R*	TCAAATACATTGGTGGGATCGG
SKM_PR 2282	*mms1 Ex1 F*	GCAACTCCCAAGAGATTACTTG
SKM_PR 2283	*mms1 Ex3 R*	GCGAAGTTCTATAGCATTGCTG

2.2 Methods

2.2.1 *S. pombe* strain maintenance

Strains were cultured to stationary phase (around 1.5 OD_{600nm}) in YEL media at 30°C for 24 hours with shaking at 250 rpm, then mixed 1:3 with 50% (v/v) sterile glycerol and immediately stored at -80°C. For experiments, strains were revived from glycerol stocks on YES plates and maintained at standard growth conditions. The transformed yeast strains were selected and maintained on SD medium with supplements as per the requirement to keep the selective pressure on the plasmid. Strains used in this study are listed in Table 2.1.

2.2.2 Yeast genomic DNA isolation

Strains were grown to saturation phase in 15 ml YEL media at 30°C for 16-18 hours at 250 rpm. Cells were harvested by centrifugation at 3000 rpm for 5 min at room

temperature and washed with distilled water. Cells were lysed by glass beads method in the presence of 200 µl phenol: chloroform: isoamyl alcohol (PCI) by vigorous vortexing for 30 secs and 30 secs on ice repeated 8 times. After vortexing, cells were suspended in 200 µl of TE buffer (pH 8.0), vortexed for 30 seconds and centrifuged at maximum speed for 5 min at room temperature. The aqueous layer was transferred to a fresh tube with 0.7 ml of 100% ethanol, the tube was gently mixed, stored at -20°C for 15- 20 min and centrifuged at maximum speed for 10 min at 4°C, after discarding the supernatant, the pellet was dried at 37°C for 10-15 min and dissolved in 100 µl TE buffer with 1 µl RNase A (20 mg/ml) and the tube was incubated at 37°C for 1 hr, followed by addition of 1 µl proteinase k (100 µg/ml) for 1 hr at 37°C. Then, 200 µl phenol: chloroform: isoamyl alcohol was added, vortexed for 1 min and 1 min on ice-repeated twice, centrifuged at 12000 rpm for 5 min and then washed with 100 µl of chloroform twice. Then, 10 µl 4M ammonium acetate with 1 ml 100% ethanol was added to the tube and incubated at -20°C for 1 hr. Following precipitation, centrifugation was performed at maximum speed for 10 min. The pellet was washed with 1 ml 70% ethanol, dried in vacuum concentrator, resuspended in 50 µl TE buffer and stored at -20°C.

2.2.3 Preparation of *S. pombe* competent cells

Competent cells preparation was done following published protocols (Knop et al.1999). Briefly, *S. pombe* cultures were grown in YEL media at 30°C with shaking for 24 hours and then reinoculated in fresh YEL media to an OD_{600nm} around 0.1-0.2. Then cells were allowed to grow till the OD_{600nm} 0.5-1.0. The cells were harvested by centrifugation at 3000 rpm, for 5 min, room temperature then washed with sterile water and once with 0.1X sterile SORB. After centrifugation, SORB was removed, and then pellet was resuspended in 360 µl SORB and 40 µl denatured salmon sperm DNA (10 mg/ml stock) per 50 ml culture. Aliquots were made, and competent cells were stored at -80°C.

2.2.4 Transformation of *S. pombe*

The transformation of *S. pombe* strains was carried out by lithium acetate method (Knop et al.1999). 10 µl of competent cells was mixed with 1 µl of plasmids in a

sterile microcentrifuge tube and six-fold sterile 40% PEG was added. After vortexing, cells were incubated at 30°C for 30 minutes. Heat shock was given at 42°C for 5 min and cells were plated on selection plates and kept at 30°C for 4-5 days.

2.2.5 Complementation assays

To complement growth defect phenotypes of *hub1-1* strain, competent cells were transformed with respective plasmids and plated on selection plates. The plates were incubated at 30°C till the growth of transformants was observed. After growth, the transformants were resuspended in sterile water, and OD_{600nm} was measured. The transformants were diluted at five-fold serial dilution in a microtiter plate, and cells corresponding to OD_{600nm} of 5 (~7.5×10^7 cells) were five-fold serial diluted and spotted on plates with 5 μg/ml thiamine and without thiamine. Following spotting, the plates were kept at 30°C and 37°C until growth was observed.

2.2.6 QuikChange site-directed mutagenesis

All point mutations on plasmids were created by QuikChange site-directed mutagenesis (Agilent) using specific-primers and high fidelity Pfu DNA polymerase. Both primers harbor the desired mutations flanked by unmodified nucleotide sequences and anneal to the same sequence on opposite strands of the plasmid. The PCR was carried for 18 cycles using 10-50 ng of the template with the annealing at 55°C for 1 min and extension at 68°C for 2 min/kb of plasmid length. Then, the PCR product was treated with 1μl Dpn1 at 37°C for 3 hours to digest the parent template. Subsequently, 10 μl of PCR product was transformed in 100 μl of DH5α competent cells and plated on selective antibiotic plate.

2.2.7 Protein isolation by trichloroacetic acid (TCA) precipitation

Cells were grown to log phase in selection media, and cells corresponding to OD_{600nm} of 1 were harvested by centrifugation at 13000 rpm for 30 sec. According to a published protocol (Knop et al. 1999), the pellets were then resuspended in freshly prepared mixture of 900 μl of 2N NaOH solution and 100 μl of β-mercaptoethanol solution. Cells were vortexed and kept on ice for 10 min. Then 200 μl of 55% trichloroacetic acid (TCA) was added and further vortexed and incubated on ice for

10 min. TCA precipitation was followed by centrifugation at maximum speed for 15 min at 4°C. The supernatant was discarded, and again centrifugation was performed for 2-3 min. Residual liquid was removed with vaccusip, and protein extraction was done in 50 μl HU buffer with 1.5% DTT by heating at 65°C, 10 min, 1400 rpm. After centrifuging it at 14,000 rpm, 1 min, room temperature; 10 μl of the isolated protein lysates was used for immunoblot assays.

2.2.8 Western blot (WB) assays

For immunoblot assays, cells corresponding to OD_{600nm} of 1 from exponentially growing culture were harvested. Primary cultures were grown in synthetic defined media till saturation at optimum conditions; then diluted to secondary culture in EMM media (without thiamine) to induce protein expression from the thiamine-repressible nmt81 promoter. Total proteins were isolated by TCA precipitation (Knop et al. 1999). 10 μl of the isolated proteins was loaded on SDS-PAGE and transferred on PVDF membrane for two and half hour at 110 mA; blocked with 5% skimmed milk for 1 hour at room temperature. The membranes were then incubated with primary antibody for 3 hours at room temperature with gentle shaking, followed by 30 minutes washing with 1x TBST buffer and incubation of HRP-conjugated secondary antibody for 2 hours at room temperature. Blots were again washed for 30 min with 1x TBST buffer and visualized using chemiluminescence detection reagents from Pierce.

2.2.9 Co-immunoprecipitation (Co-IP) assays

i. Cells harvesting: Cells were grown to log phase, OD_{600nm} around 0.6-0.8 and total cells corresponding to OD_{600nm} of 50 were harvested by centrifugation at 3000 rpm for 5 min at 4°C. After centrifugation supernatant was discarded and pellets were resuspended in cell-lysis buffer with PMSF, phosphatase and protease inhibitors, snap froze in liquid nitrogen and stored at -80°C. The assay was described previously (Mishra et al. 2011).

ii. Cell-lysis: Total cell lysates were prepared by mechanical grinding of frozen pellets with liquid nitrogen in the presence of cell-lysis buffer. The total cell

suspension volume was 1 ml. Lysates were pre-cleared two times by centrifugation at 10,000 x g for 10 min at 4°C

iii. Immunoprecipitation (IP): After pre-clearing, the supernatant was transferred to new micro-centrifuge tube and immunoprecipitation was done using appropriate antibody tagged beads (15 μl/50 OD_{600nm} cells) for 3 hours at 4°C on slow speed rotator. After immunoprecipitation, unbound proteins were washed away by centrifugation at 3000 rpm, 4°C for 2 min, first with diluted lysis buffer; then three times with wash buffer 1 and finally by wash buffer 2 (without triton X-100). The supernatant was discarded thoroughly using vaccusip and both inputs (2%), as well as immunoprecipitated proteins, were extracted by heating at 65°C for 10 min in the presence of 25 μl HU buffer. After centrifuging it at 14,000 rpm, 5 min, room temperature; 10 μl of eluted proteins was loaded on NU-PAGE, and co-immunoprecipitated proteins were analyzed by immunoblotting with respective antibodies (The experiment was performed once only).

Co-immunoprecipitation (Co-IP) assays (Large Scale)

i. Cells harvesting: Cells were grown to log phase, OD_{600nm} around 0.8 and total cells corresponding to OD_{600nm} of 1600 were harvested by centrifugation at 3000 rpm for 10 min at 4°C. After centrifugation, supernatant was discarded and pellets were resuspended in cell-lysis buffer with PMSF, phosphatase and protease inhibitors, snap froze in liquid nitrogen and stored at -80°C. The assay was described previously (Mishra et al. 2011).

ii. Cell-lysis: Total cell lysates were prepared by mechanical grinding of frozen pellets with liquid nitrogen in the presence of cell-lysis buffer. The total cell suspension volume was 10 ml. Lysates were pre-cleared two times by centrifugation at 10,000 x g for 10 min at 4°C.

iii. Immunoprecipitation (IP): After pre-clearing, the supernatant was transferred to a new 15 ml falcon tube and immunoprecipitation was done using appropriate antibody tagged beads (200 μl) for 6 hours at 4°C on slow speed rotator. After immunoprecipitation, unbound proteins were washed away by centrifugation at 3000 rpm, 4°C for 2-3 min, first with diluted lysis buffer; then three times with wash buffer 1 and finally by wash buffer 2 (without triton X-100). The supernatant was discarded thoroughly using vaccusip and both inputs (~2%), as well as immunoprecipitated proteins, were extracted by heating at 65°C for 10 min in the presence of 25 μl HU buffer. After centrifuging it at 14,000 rpm, 5 min, room temperature; 20 μl of eluted proteins was loaded on NU-PAGE, and co-immunoprecipitated proteins were subjected for mass spectrometry. From the identified proteins in mass spectrometry, we specifically selected all the splicing factors and the table represents the number of unique peptides in mass spectrometry for selected proteins.

2.2.10 Chromosomal tagging of splicing factors

Chromosomal tagging and gene deletion was done following published protocols (Janke et al., 2004). Briefly, sequence encoding 6HA epitope tag with kan-MX4 cassette for resistance against G418 antibiotic was inserted chromosomally at the C-terminus of selected splicing factors. The cassette was amplified using long primers that contain sequences of homology to the genomic target location with a mixture of Tag and vent polymerase. The PCR product was ethanol-precipitated and 10 μl of it was transformed in *S. pombe* strains. After transformation, the strains were revived for 16-18 hours on shaking followed by selection on YES-G418 plates. The transformants were then screened by immunoblotting assays using anti-HA antibody.

2.2.11 Chromosomal mutagenesis of *hub1* gene

Chromosomal mutagenesis was done following published protocols (Janke et al., 2004). A plasmid was made which contained a NotI insert comprising of the *S. pombe hub1* genomic DNA including 500 bp upstream of start codon of the gene and

100bp downstream of the stop codon of the gene, followed by Nat-NT2 cassette and 101-600bp downstream of the stop codon of the gene. The plasmid was mutagenized by site-directed mutagenesis to obtain versions of *hub1* mutants. The NotI digested inserts from these plasmids were purified and transformed in *S. pombe* strains. The revival in YEL media was done for not more than 8-12 hrs. The revived cells were plated on YES-Nat plates. The chromosomal mutations were screened by growth phenotypes and sequencing of genomic PCR fragments covering the mutated regions.

2.2.12 Splicing-sensitive microarray

S. pombe splicing-sensitive microarray design, experimental procedure, and data analysis were reported earlier (Inada and Pleiss, 2010) and the experiment was carried out in the laboratory of Jeffrey A. Pleiss at Cornell University. Briefly, total RNA was isolated from logarithmically growing cultures of wild-type and *hub1-1* strains grown at 30°C and 37°C for 15 min by hot acid phenol method. Subsequently, total RNA was converted to cDNA using reverse transcriptase (RT) and random nine-mer primers. cDNAs from both strains were labeled with Cy3 and Cy5 dyes. Mixtures of Cy3-labeled wild-type sample and Cy5-labeled *hub1-1* sample and from dye-swapped samples were hybridized to the splicing-sensitive microarrays having introns-, exons- and junctions- specific probes for nearly all intron-containing genes of *S. pombe*. The relative abundance of transcripts was compared between the wild-type and *hub1-1* strains.

2.2.13 RNA isolation and RT-PCR

i. Cells harvesting: RNA isolation and cDNA synthesis were done as described previously (Inada and Pleiss 2010). Briefly, cells corresponding to OD_{600nm} of 5.0 in the logarithmically growing phase were harvested at 30°C (untreated control) or after 15 minutes of heat shock at 37°C by filtration and pellets were stored at -80°C after snap freezing with liquid nitrogen.

ii. RNA isolation: Total RNA was isolated by the hot acid phenol method using 15 ml phase-lock gel heavy tubes. Briefly, pellets were

resuspended in acid phenol: chloroform and AES buffer by vortexing. Then the pellets were transferred to 65°C water bath for 7-10 min and vortexed thoroughly once every minute. After lysis, cells were incubated on ice for 5 min, and entire organic and aqueous phase was transferred to pre-spun 15 ml phase-lock gel tubes. The tubes were centrifuged at 3000 x g at 4°C for 5 min. Then, PCI was added to the gel tubes, and again centrifugation was done. Subsequently, chloroform was added to the supernatant, and after centrifugation, the aqueous phase was transferred into a new 15 ml conical tube with isopropanol and 3 M sodium acetate. The conical tubes were vortexed thoroughly and 2 ml isopropanol slurry was centrifuged at maximum speed for 20 min at 4°C. After centrifugation, the supernatant was discarded, and RNA pellets were washed two times with 70% ethanol. The RNA pellets were dried in a vacuum concentrator and finally resuspended in nuclease-free water. It was followed by DNase I treatment for 15 min at room temperature and Zymo-Spin II column was used for clean-up of RNA. Total RNA was quantified with a spectrophotometer by measurement at OD260/280nm (Nanodrop).

iii. cDNA synthesis and RT-PCR: cDNA synthesis from 3 μg total RNA was done using RT and random-hexamer primer at 42°C for 16 hours followed by RT-PCR using target-specific primers and products were analyzed by agarose gel electrophoresis. Primers used in RT-PCR assays of splicing targets are listed in table 2.3.

2.2.14 APEX labeling in *S. pombe*

i. Cells harvesting: Cells corresponding to OD_{600nm} of 50 were harvested, washed with water and 1.2 M sorbitol solution. The cell pellet was dissolved in 1 ml of 1.2 M sorbitol solution, followed by addition of 2.5 mM biotin-phenol (BP). It was mixed on a roller for 30 min-1 hr at room temperature.

ii. *In vivo* biotin labeling: Added 1 mM H_2O_2 for 30 sec-1 min to initiate biotin-labeling, quickly spun at 4000 rpm for 1.5-2 min and aspirated off

the solution. Washed 3 times with 1 ml quenching solution (5 mM Trolox, 10 mM sodium azide and 10 mM sodium ascorbate in 1.2 M sorbitol dissolved in H_2O) and finally washed the pellet with 1.2 M sorbitol.

iii. Protein extraction: Cells were lysed by the addition of 27 mM NaoH, 1% (v/v) 2-mercaptoethanol and incubated on ice for 10 min. Added trichloroacetic acid (20% w/v), incubated on ice for 10 min and spin down at 4°C. The protein pellet was washed with 0.5-1 ml ice-cold acetone once, added 500-700 μl lysis buffer (25 mM Tris pH 7.4, 150 mM NaCl, 2% SDS) containing complete protease inhibitor EDTA free and denatured by heating at 75°C for 15 min.

iv. Biotinylated protein enrichment: For streptavidin pulldown, sample were dialyzed in dialysis buffer (25 mM Tris pH 7.4, 150 mM NaCl, 0.2% SDS) using a 3.5K MWCO dialysis tubing for 3 hrs at room temperature. 5 mg of protein sample was incubated with 200 μl Streptavidin-coated magnetic bead for 2 hr at room temperature, followed by 10-times washing with wash buffer (25 mM Tris pH 7.4, 150 mM NaCl, 0.3% SDS), 2-times wash with 2 M urea/50 mM Tris pH 7.4 and final wash with wash buffer. Biotinylated proteins were eluted by incubating the beads with 30 μl 2X SDS sample loading buffer (4% (w/v) SDS, 0.2% (w/v) bromophenol blue, 20% (v/v) glycerol, 100 mM Tris pH6.8, 10% (v/v) 2-mercaptoethanol) and heating at 75°C for 15 min. 15 μl of samples was subjected to mass spectrometry.

2.2.15 High-copy suppressor screen

For high-copy suppressor screen, hub1 surface III mutant competent cells were freshly prepared, and cells were transformed with a high-copy pREP1 vector-based S. pombe cDNA library. The transformants were on plated on Emm-Leu plates, incubated at 30°C for 6 hours. Further cells were shifted to 37°C till the growth of transformants appeared. Further plasmids were shuttled out from the transformants which grew at 37°C . The identity of the clones was established by DNA sequencing.

2.2.16 Yeast two-hybrid screen

For yeast two-hybrid screening of whole cDNA libraries, potential binding factors were expressed as Gal4 activation domain fusion proteins (AD-fusion), whereas the bait protein carries Gal4 DNA binding domain. In case of physical interaction between the two fusion proteins reporter gene expression (HIS3 and ADE2) was induced which allows growth on selection medium (SC-Ura-Leu-His: 3.5% yeast nitrogen base, 2% glucose and 0.2% amino acid mix). Here, Hub1 served as the BD-fusion bait protein for screening with S. pombe cDNA library. Two-hybrid yeast strain PJ69-7A was co-transformed with pGBDU-Hub1 construct and *S. pombe* cDNA library to identify interactors of Hub1. Transformants monitored for interaction on SC-Ura-Leu-His/Ade plates. Positives clones were further retransformed, and interaction is validated by analysis growth on SC-Ura-Leu-His/Ade plates.

2.2.17 Purification of recombinant proteins from *E. coli*

E. coli BL21(DE3) cell pre-cultures were inoculated in 50 ml antibiotic-containing LB media and cultured at 37^{o}C overnight. For recombinant protein purification, these cultures served as inoculants for larger volumes starting at OD_{600nm} at 0.2. E. coli cells reached OD_{600nm} at 0.4-0.5 were induced with 1 mM IPTG for 3-5 h at RT.

I. In the case of GST-fusion proteins, the pelleted cells (6000xg, 10 min, 4^{o}C) were resuspended in 40 ml of lysis buffer (with 1 mM PMSF, 200 μl Protease inhibitor cocktail solution, 0.1 μl/ml Nuclease) and enzymatic cell wall digestion by lysozyme on ice for 30 min. Followed by sonication (10-15 Amplitude, 10 secs on sonication &10 secs on cooling), lysates were centrifuged at 5000 rpm for 10 min at 4^{o}C. Cleared supernatants were loaded onto ethanol-free glutathione agarose beads (Qiagen) and incubated for 60 min on ice with continuous shaking. It is transferred to the column (10 ml capacity) and subsequently, columns were washed 3-4 times with 10 ml of lysis buffer, 2 times with 10 ml of high salt buffer (with 300 mM NaCl), once with 10 ml of detergent buffer (with 1% TX-100) and finally with 10 ml of PBS. Finally, affinity matrix-bound proteins were eluted with 500 μl of GST elution buffer. Total 6 such fractions were eluted, which were analyzed by protein concentration measurements and SDS PAGE, to ensure successful and efficient purification. The

eluted was dialyzed against PBS (0.1 mM DTT + 0.2 mM PMSF) for 2-3 hrs at 4^{o}C and later shifted to glycerol-containing PBS buffer for 16 hrs (0.1 mM DTT + 0.2 mM PMSF + 10% glycerol) prior to freezing in liquid N_2 and storage at -80^{o} C.

II. Similarly, 6xHis-tagged fusion proteins were purified from *E. coli* lysates after induction, expression, and harvest. Here cell lysates were resuspended Ni-NTA lysis buffer (with 1 mM PMSF, 200 μl Protease inhibitor cocktail solution, 0.1 μl/ml Nuclease) before lysis via lysozyme digestion and sonication (see above). After centrifugation cleared lysates were incubated with Ni-NTA agarose columns (Qiagen) and incubated for 1 hr at 4^{o}C. Followed by 4 times washing with 10 ml of wash buffer, once with 10 ml of wash buffer containing 1% TX-100 and final wash with 10 ml of wash buffer. Finally, affinity matrix-bound proteins were eluted with 500 μl of Ni-NTA elution buffer. Total 6 such fractions were eluted, which were analyzed by protein concentration measurements and SDS PAGE, to ensure successful and efficient purification. The eluted was dialyzed against PBS (0.1 mM DTT + 0.2 mM PMSF) for 2-3 hrs at 4^{o}C and later shifted to glycerol-containing PBS buffer for 16 hrs (0.1 mM DTT + 0.2 mM PMSF + 10% glycerol) prior to freezing in liquid N_2 and storage at -80^{o} C.

III. Similarly, in the case of insoluble 6xHis-tagged fusion proteins, denaturing conditions were applied for its purification from *E. coli* lysates after induction, expression, and harvest. Here cell lysates were resuspended Ni-NTA lysis buffer before lysis via lysozyme digestion and sonication (see above). After centrifugation, pellet fractions were washed with 1xPBS and resuspended in high urea buffer (8M urea + 1XPBS) for 2-3 hrs at 37^{o}C. After centrifugation cleared lysates were incubated with Ni-NTA agarose columns (Qiagen) and incubated for 1 hr at 4^{o}C. Followed by 4-5 times washing with 10 ml of wash buffer (20 mM imidazole, 8M urea in 1xPBS), Finally, affinity matrix-bound proteins were eluted with 500 μl of Ni-NTA elution buffer with 8M urea. Total 6 such fractions were eluted. The eluted was dialyzed against 1x PBS with 4M urea for 1 hr at 4^{o}C, shifted to 1x PBS with 2M urea for 1 hr at 4^{o}C and later shifted to glycerol-containing PBS buffer for 16 hrs (0.1 mM DTT + 0.2 mM PMSF + 10% glycerol). Dialysed samples were analyzed by protein

concentration measurements and SDS PAGE, to ensure successful and efficient purification prior to freezing in liquid N_2 and storage at -80^{o}C.

2.2.18 GST pull-down assay

In case of GST pull-down assay, bacterial purified recombinant GST and GST-Snu66(HIND) proteins were individually mixed with *6xHis-hub1-DD* and *6xHis-hub1-GG* proteins. The mixture was incubated for 2 hr at 4°C on slow speed rotator. Around 50 μl of pre-washed glutathione agarose beads were added to each tube separately and incubated for 1 hr at 4°C on slow speed rotator. After binding with beads, unbound soluble fraction was washed away by centrifugation at 3000 rpm, 4°C for 2-3 min. Later, beads were washed three times with wash buffer 1 (1X PBS with 15% glycerol, 0.5% triton X-100) and finally with wash buffer 2 (without triton X-100). The supernatant was discarded thoroughly using vaccusip, both inputs (~10%) as well as pull-down proteins were extracted by heating at 65°C for 10 min in the presence of 30 μl HU buffer. After centrifuging it at 14,000 rpm, 5 min, room temperature; 20 μl of eluted proteins was loaded on NU-PAGE, and pull-down proteins were analyzed by staining the NU-PAGE gel. The experiment was performed once only.

Note: In case of *C. elegans* Hub1-Snu66 interaction, the bacterial purified recombinant GST and GST-*Ce* Snu66(HIND) proteins were were individually mixed with 6xHis-*Ce* Hub1 protein (The experiment was performed twice). The experiment was performed similar to 2.2.18.

2.2.19 Co-immunoprecipitation (Co-IP) assays using worm lysates

Worms from plates were washed thrice with M9 buffer, transferred to 1.5 ml MCT, supernatant discarded and cell pellets were dissolved in 0.4 ml *C. elegans* HEPES lysis buffer per plate of worms (PMSF, protease inhibitor, phosphatase inhibitor). Further froze in liquid nitrogen and stored -80^{o}C and lysates were transferred to steel Dounce homogenizer on ice, stroked 30-40 times. Further lysates were subjected to homogenization using mortar and pestle. Around 0.2 ml of lysis buffer (volume of 1.5X lysis buffer with protease inhibitors) to each gram of adult/ embryo beads. The worm lysate was centrifuged for 10 mts at 5000 rpm at 4^{o}C. Protein concentration

was estimated using Bradford reagent. The experiment was as in 2.2.9 (HEPES buffer with 0.5% triton x-100) was used in washing steps.

2.2.20 *C. elegans* strains maintenance

All strains were maintained at 20^{o}C, as described previously (Brenner, 1974). The *E. coli* strain OP50 was used for seeding the *C. elegans* plates. The Bristol N2 strain was used as the wild-type (WT) control strain. The strains were maintained for long term in liquid nitrogen by using freezing media as mentioned in wormbook. The synchronization of strains was performed using sodium hypochlorite treatment after growing the worms till adult stage. Also, to get rid of bacterial and fungal contamination sodium hypochlorite treatment was used.

2.2.21 *C. elegans* RNA isolation

Worms were grown on OP50 containing plates for mixed stage population of worms. The worms were harvested by washing with M9 buffer in a microcentrifuge tube. Centrifuging at 14,000 rpm pelleted the worms down. The worms were washed 2-3 times to get rid of OP50. For stage, specific RNA isolation worm population were first synchronized by bleaching the worms and then the worms are collected at different time points to collect L1, L2, L3, L4, and adult worms. The RNA isolation performed similar to 2.2.13.

2.2.22 *C. elegans* splicing-sensitive microarray

The microarray hybridization and scanning were performed at the Agilent certified microarray facility of genotypic technology, Bengaluru, India. The samples for gene expression were labeled using Agilent Quick-Amp labeling Kit (p/n5190-0442). 500ng each of total RNA were reverse transcribed at 40°C using oligo dT primer tagged to a T7 polymerase promoter and converted to double stranded cDNA. Synthesized double stranded cDNA were used as template for cRNA generation. cRNA was generated by in vitro transcription and the dye Cy3 CTP (Agilent) was incorporated during this step. The cDNA synthesis and in vitro transcription steps were carried out at 40°C. Labeled cRNA was cleaned up using Qiagen RNeasy columns (Qiagen, Cat No: 74106) and quality assessed for yields and specific

activity using the Nanodrop ND- 2000. 600ng of labeled cRNA sample were fragmented at 60°C and hybridized on to an Agilent gene expression Microarray 8X60K. Fragmentation of labeled cRNA and hybridization were done using the gene expression hybridization kit (Agilent Technologies, In situ Hybridization kit, Part Number 5190-0404). Hybridization was carried out in Agilent's surehyb chambers at 65°C for 16 hours. The hybridized slides were washed using Agilent gene expression wash buffers (Agilent Technologies, Part Number 5188-5327) and scanned using the Agilent microarray scanner (Agilent Technologies, Part Number G2600D).

Microarray Data Analysis

Feature extracted raw data was analyzed using Agilent genespring GX software. Normalization of the data was performed in genespring GX using the 75th percentile shift method [Percentile shift normalization is a global normalization, where the locations of all the spot intensities in an array are adjusted. This normalization takes each column in an experiment independently, and computes the n^{th} percentile of the expression values for this array, across all spots where n has a range from 0-100 and n=75 is the median. It subtracts this value from the expression value of each entity] and fold change values were obtained by comparing mutant samples with respect to specific wild-type samples. significant genes up regulated fold> 1 (logbase2) and down regulated <-1 (logbase2) in the mutant samples with respect to wild-type samples were identified. Statistical student T-test p-value among the replicates was calculated based on volcano Plot Algorithm.

Chapter 3

The role of ubiquitin-like protein Hub1 in RNA splicing in *Schizosaccharomyces pombe*

Abstract

The ubiquitin-like protein Hub1 is a conserved regulator of pre-mRNA splicing and alternative splicing. In *Saccharomyces cerevisiae*, Hub1 binds to the spliceosomal protein Snu66 through its Asp-22 surface (surface I), and to the DEAD-box helicase Prp5 through its His-63 surface (surface II). Hub1 is not essential for viability in *S. cerevisiae,* but the protein becomes essential in *Schizosaccharomyces pombe,* possibly because Hub1 is required for splicing of a larger number of introns. By using microarray and RT-PCR, we show that Hub1 is crucial for pre-mRNA splicing in *S. pombe*. The protein specifically modulates the spliceosome. It possibly facililates assembly of acitivated spliceosome by removing a specific set of essential splicing factors. Strikingly, however, mutations in both surfaces resulted in only mild growth and splicing defects, thereby suggesting that Hub1 might work in pre-mRNA splicing thorugh other unidentified surface(s).

By using approaches of NMR, mutagenesis, and complementation of a *hub1*-knockout *S. pombe* strain, we have identified a new Hub1 surface, surface III, containing Arg-9 and Arg-41 residues (referred to as hub1-R941). This surface is functionally conserved in all eukaryotes and alteration of this surface resulted in temperature sensitive growth and splicing defects in *S. pombe*. Further, SPCPB16A4.06C, a Schizosaccharomyces specific gene, was identified as a high-copy suppressor of the temperature sensitive *hub1* surface III mutant. We further show that Hub1 promotes splicing of the transcripts that are synthesized faster. It possibly facilitates coupling of transcription with splicing. Altogether, this part of my study revealed that, despite its small globular size of only 73 amino acids (~8.4 kDa), Hub1 employs multiple surfaces for its function in pre-mRNA splicing. These surfaces facilitate binding of specific (known and certain unknown) factors. My

findings are highly relevant not only for regulatory and tissue-specific gene expression, but also for understanding new mechanisms of alternative splicing.

3.1 Introduction

The proteins of the ubiquitin family (also called ubiquitin-like proteins or UBLs) are central regulators of various cellular processes (Hochstrasser, 2000). Ubiquitin and UBLs contain a globular β-grasp fold and a flexible C-terminal tail. They are synthesized as inactive precursors and processed by UBL-specific proteases after a conserved di-glycine (GG) motif. These proteins get conjugated mostly to proteins through an enzymatic cascade involving activation, conjugation and ligase enzymes (Taherbhoy et al., 2012; Nakai et al., 2008).

Introduction on Hub1 (detailed in section 1.8). In *S. cerevisiae*, two independent surfaces of Hub1 are reported for its function in alternative splicing. Through its Asp-22 surface, it binds to splicing factors Snu66 and/or Prp38, and through its His-63 surface, it interacts and activates the RNA helicase Prp5. Contributions of these surfaces have not been elucidated in intron-rich or multicellular organisms, though the role of Hub1 in pre-mRNA splicing have been studied by RNAi-mediated knockdown approaches in human cells (Ammon et al., 2014; Oka et al., 2014). The surfaces relevant for its splicing function has not been elucidated. In our study, we show that Hub1 is required for pre-mRNA splicing in *S. pombe*. However, mutations in both surfaces reported previously resulted in only mild growth and splicing defects. We have identified a new Hub1 surface (surface III / Hub1-R941 surface). This surface is functionally conserved in eukaryotes. Alteration of this surface results in severe growth and splicing defects in *S. pombe*. A Schizosaccharomyces specific gene was identified as a high-copy suppressor of *hub1* surface III mutant temperature sensitivity. These results suggest that the Hub1 contains multiple surfaces to facilitate the binding of certain other proteins for its splicing function.

3.2 Results

3.2.1 Hub1 is required for pre-mRNA splicing in *S. pombe*

To understand Hub1 role in pre-mRNA splicing, my thesis guide Shravan Kumar Mishra performed genome-wide splicing-sensitive microarrays for the *hub1-1* mutant in collaboration with Jeffrey A. Pleiss at Cornell University. For each intron-containing gene, a minimum of three probes were used; the intronic probes detected pre-mRNA, exon-exon junction probes detected mature mRNA, and exonic probes detected total mRNAs (Inada and Pleiss 2010). Total RNA isolated from wild-type (WT) and the *hub1-1* strain was converted into cDNA using random nine-mer primers. The cDNAs were labeled with Cy3, Cy5, and dye-swapped and hybridized on splicing-sensitive microarrays. After hybridization, the relative abundance of the transcripts was quantified and compared in both the strains. The heat map of microarrays illustrates the $\log_2$ ratio of transcripts in the *hub1-1* strain as compared to the wild-type strain (Figure 3.1A). The splicing of many genes was affected in the *hub1-1* strain as evident by the enhanced accumulation of intron-containing transcripts and a corresponding decrease in mature transcript level in comparison to wild-type cells at 37°C (Figure 3.1A). I confirmed the Hub1-dependent splicing defects for the selected targets using RT-PCR assays. For RT-PCR experiments, we isolated total RNA from logarithmically growing wild-type and *hub1-1* cells grown at 30°C and 37°C. The selected targets showed accumulation of intron-containing transcripts in *hub1-1* cells at 37°C (Figure 3.1B). The above results indicate that Hub1 is critical for pre-mRNA splicing in *S. pombe*.

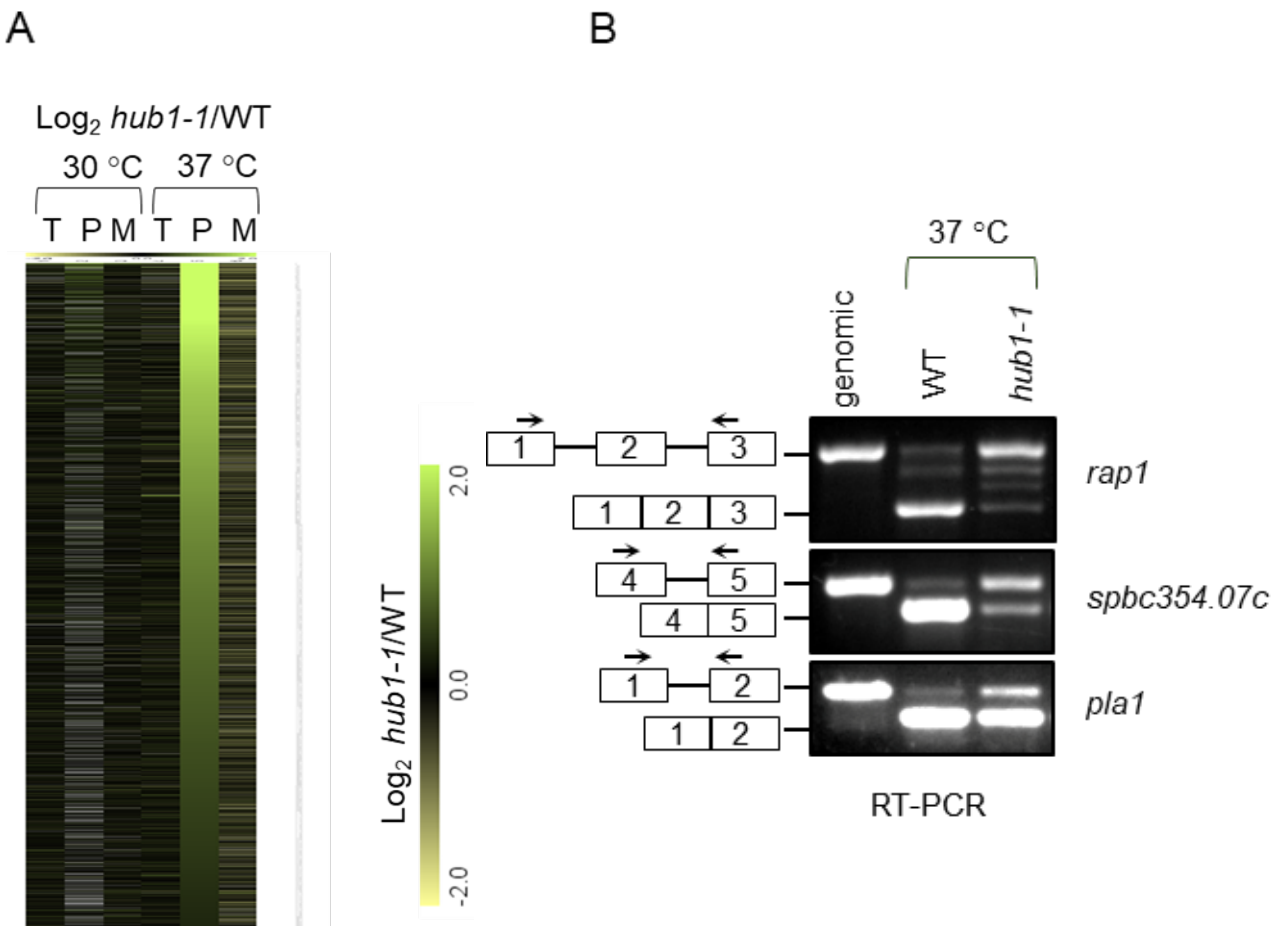


Figure 3.1: Hub1 is required for pre-mRNA splicing in *S. pombe*

A. Analysis of total RNA from *S. pombe* WT (SP9) and *hub1-1* strain (SP10) using splicing-sensitive microarray. Microarray heat-map shows the $\log_2$ *hub1-1*/WT ratio of signals for total transcripts (T), intron-containing transcripts (P), and spliced transcripts (M). 37°C temperature shift was performed for 15 minutes to monitor early splicing defects. The experiment was repeated with the two dyes swapped. Here, light green color represents accumulation; black denotes no change and yellow shows reduction of signal for the transcripts (Data is from Shravan Kumar Mishra).

B. Semi-quantitative RT-PCR shows accumulation of intron-containing transcripts for *rap1, spbc354.07c, and pla1.* 37°C temperature shift was performed for 15 minutes to monitor early splicing defects. Total RNAs were converted into cDNAs using random hexamer primers and reverse transcriptase from *S. pombe* WT and *hub1-1* strains. Following cDNA synthesis, we carried out RT-PCR assays for the selected

targets, which showed an accumulation of pre-mRNA in microarrays. The amplified products were visualized on agarose gel electrophoresis. The block diagrams (not drawn to the scale) represents exons and introns. Primers are depicted with arrows on exons. PCR bands from the genomic DNA corresponds to pre-mRNA.

3.2.2 Hub1 selectively alters the composition of the spliceosome

In *S. pombe*, Hub1 is essential for pre-mRNA splicing, but the molecular mechanism remains elusive. To address this, I immunoprecipitated spliceosomal components in wild-type and *hub1-1* cells with 6HA tagged versions of core-splicing factors Cdc5/ Prp19. The immunoprecipitated complex with anti-HA beads was purified and subjected to mass spectrometry analysis. We found that some of the splicing factors seem to be elevated, and one factor with a diminished level in the spliceosome of the *hub1-1* cells as compared to wild-type cells (Figure 3.2A). The factors elevated in the spliceosome include RNA helicase Prp11 (ortholog of Prp5 in *S. cerevisiae*), Prp24 (RNA binding protein, a part of tri-snRNP complex), Prp1 (a part of tri-snRNP complex), and Cwf16 (complexed with cdc5 protein). On the other hand, Mug161 (post-mRNA release spliceosomal complex) levels was diminished in the spliceosome in *hub1-1* cells as compared to wild-type cells. To verify this result, I performed immunoprecipitation of epitope-tagged Mug161 protein from these strains and confirmed reduced co-immunoprecipitation of Mug161 from *hub1-1* mutant (Figure 3.2B). However, *mug161* pre-mRNA is a splicing target of Hub1 in our splicing-sensitive microarray. This could result in lesser amount of protein being generated in *hub1-1* cells, due to which a lower level of Mug161 might be present in the spliceosome in *hub1-1* cells. Further, a similar set of experiments need to be performed to verify the enriched splicing factors in the spliceosome from *hub1-1* mutant. Nonetheless, most of the components of the spliceosome remained unaltered in both experiments. (The number of unique peptides of spliceosomal proteins identified in mass spectrometry is listed in Table 3.1.) Thus, we infer that Hub1 selectively modulates the composition of the spliceosome.

A

The number of unique peptides

Protein	IP: Cdc5-6HA			IP: Prp19-6HA		
	WT	***hub1-1***	***hub1-1/WT***	**WT**	***hub1-1***	***hub1-1/WT***
Cdc5	72	76	1.05	69	69	1
Prp19	28	28	1	26	24	0.92
Prp11	12	61	5.08	16	58	3.62
Prp24	4	25	6.25	2	26	13
Prp1	3	23	7.66	5	28	5.6
Cwf16	3	6	2	4	10	2.5
Mug161	22	11	0.5	18	9	0.5

B

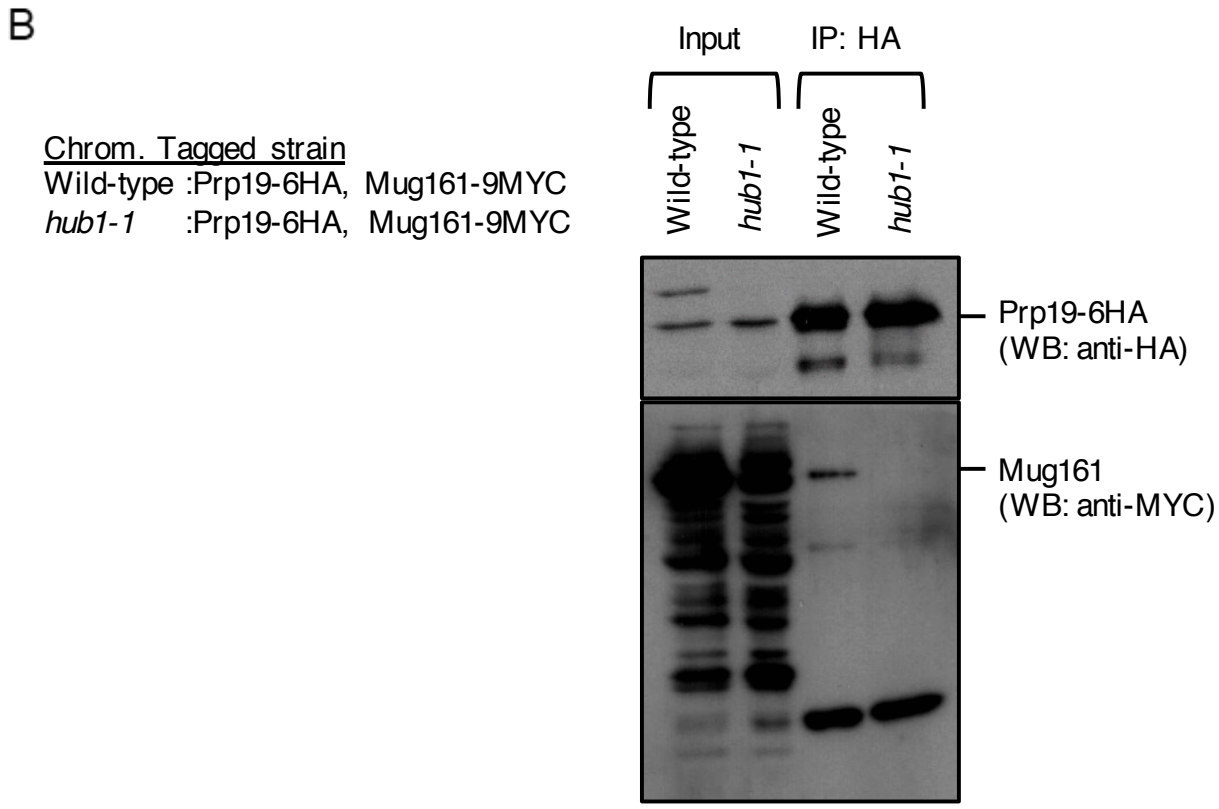


Figure 3.2: Hub1 modifies the spliceosome

A. Cdc5-6HA complexes were immunoprecipitated using anti-HA antibody beads from *S. pombe* lysates wild-type (SP42) and *hub1-1* strains (SP75). Prp19-6HA complexes were immunoprecipitated using anti-HA antibody beads from *S. pombe* lysates wild-type (SP44) and *hub1-1* strains (SP37). Co-IP proteins were analyzed by

mass spectrometry. The table shows number of unique peptides in mass spectrometry for an identified protein. Splicing factors which are overrepresented in the spliceosome in *hub1-1* cells are highlighted with orange color, and underrepresented factor is highlighted with blue color.

B. Hub1 co-immunoprecipitates (Co-IP) splicing factor Mug161 protein *in vivo*. Lysates from *S. pombe prp19-6HA*, *mug161-9MYC* chromosomally tagged genes in wild-type (P8) and *hub1-1* cells (P10) were immunoprecipitated using anti-HA antibody beads. The Co-IP of Mug161-9MYC was analyzed by western blotting (WB) using anti-MYC antibody. The level of Mug161 is diminished in both input and Co-IP samples from *hub1-1* mutant.

	IP: Cdc5-6HA			**IP: Prp19-6HA**		
Protein	**WT**	***hub1-1***	***hub1-1*/WT**	**WT**	***hub1-1***	***hub1-1*/WT**
Spp42	179	178	0.99	180	165	0.91
Brr2	130	135	1.03	118	123	1.04
Cwf10	78	86	1.1	81	77	0.95
Cwf11	86	83	0.96	84	77	0.91
Cdc5	72	76	1.05	69	69	1
Prp22	64	74	1.15	59	70	1.18
Prp16	42	67	1.59	21	43	2.04
Cwf3	68	64	0.94	65	63	0.96
Cwf4	65	63	0.96	54	56	1.03
Prp10	41	61	1.48	46	65	1.41
Prp11	12	61	5.08	16	58	3.62
Prp12	31	59	1.90	46	55	1.19
Prp17	43	48	1.11	44	41	0.93
Cwf22	40	45	1.125	42	44	1.04
Cwf19	52	45	0.86	50	41	0.82
Cdc28	36	43	1.19	48	54	1.12
Dbp2	32	39	1.21	27	24	0.88
Prp5	34	37	1.08	34	35	1.02
Prp39	28	36	1.28	32	36	1.12
Prp40	26	35	1.34	25	35	1.4
Prp45	33	33	1	37	40	1.08
Prp19	28	28	1	26	24	0.92
Sap145	17	28	1.6	24	31	1.29
Cwf25	14	28	2	22	33	1.5
Usp107	24	27	1.12	24	32	1.33
Cwf2	28	26	0.92	27	23	0.85

	IP: Cdc5-6HA			IP: Prp19-6HA		
Protein	**WT**	***hub1-1***	***hub1-1*/WT**	**WT**	***hub1-1***	***hub1-1*/WT**
Prp2	17	26	1.52	15	26	1.73
Sap61	19	26	1.36	22	29	1.31
Syf2	26	25	0.96	26	24	0.92
Srp2	22	25	1.13	17	20	1.17
Prp43	28	25	0.89	17	16	0.94
Prp24	4	25	6.25	2	26	13
Sap114	13	24	1.84	18	23	1.27
Cwf5	28	23	0.82	27	28	1.03
Tcg1	20	31	1.55	21	30	1.42
Cwf21	15	23	1.53	20	15	0.75
Prp1	3	23	7.66	5	28	5.6
Cwf17	25	22	0.88	24	24	1
Dbp3	22	22	1	nd	nd	nd
Cwf15	18	21	1.16	22	18	0.81
Cwf7	20	20	1	23	13	0.56
Lea1	18	19	1.05	13	13	1
Bpb1	12	19	1.58	9	16	1.77
Cay1	15	19	1.26	9	8	0.88
Cwf26	9	19	2.11	15	14	0.93
SPAC20H4.09	19	17	0.89	20	19	0.95
Tif412	11	17	1.54	9	13	1.44
Cwf12	12	16	1.33	9	11	1.22
Bis1	16	16	1	17	18	1.05
Cbf5	19	16	0.84	22	18	0.81
Cwf24	11	16	1.45	20	19	0.95
Dbp9	15	14	0.93	25	20	0.8
Saf4	12	13	1.08	18	9	0.5
Slu7	12	13	1.08	12	8	0.66
Has1	13	13	1	23	18	0.78
Nrl1	8	13	1.62	7	8	1.14
Mug161	22	11	0.5	18	9	0.5
Smd2	8	10	1.25	7	8	1.14
U2af-23	11	10	0.90	6	12	2
Srp1	10	9	0.9	7	5	0.71

	IP: Cdc5-6HA			IP: Prp19-6HA		
Protein	**WT**	***hub1-1***	***hub1-1*/WT**	**WT**	***hub1-1***	***hub1-1*/WT**
Usp109	7	9	1.28	8	9	1.12
Sap62	4	9	2.25	12	8	0.66
Ntr2	12	9	0.75	1	nd	nd
Usp106	3	9	3	7	8	1.14
Smb1	8	8	1	8	9	1.12
Cwf18	9	8	0.88	7	6	0.85
Cwf14	9	8	0.88	9	9	1
Exo2	22	8	0.36	27	5	0.18
Sap14	7	8	1.14	5	8	1.6
Smg1	8	7	0.87	6	7	1.16
Mtr4	8	7	0.87	4	7	1.75
Msl1 (Ru2B)	6	6	1	7	7	1
Smf1	5	6	1.2	6	4	0.66
Usp108	5	6	1.2	6	7	1.16
Rns1	5	6	1.2	5	4	0.8
Usp102	8	6	0.75	8	10	1.25
Usp101	5	6	1.2	7	6	0.85
Cwf16	3	6	2	4	10	2.5
Prp31	nd	6	6	nd	6	6
Prp3	nd	6	6	nd	11	11
Smd1	4	5	1.25	5	4	0.8
Sde2	6	5	0.83	7	7	1
Sap49	4	5	1.25	3	3	1
Cwf28	5	5	1	10	9	0.9
Smd3	6	4	0.66	6	5	0.83
Sme1	5	4	0.8	5	3	0.6
SPAC32A11.02c	15	4	0.26	15	4	0.26
Cwf29	3	4	1.33	3	4	1.33
Ini1	3	4	1.33	2	2	1
Lsm8	3	4	1.33	nd	nd	nd
Cwf23	nd	4	4	nd	nd	nd
Snu66	nd	4	4	nd	2	2
Gar1	3	3	1	4	3	0.75
Nhp2	4	3	0.75	4	3	0.75

	IP: Cdc5-6HA			IP: Prp19-6HA		
Protein	**WT**	***hub1-1***	***hub1-1*/WT**	**WT**	***hub1-1***	***hub1-1*/WT**
Nop10	3	3	1	3	2	0.66
Lsm7	3	3	1	1	2	2
Usp103	1	2	2	2	1	0.5
Lsm2	3	2	0.66	1	2	2
SPAC29A4.06c	3	2	0.66	2	5	2.5
Lsm3	2	2	1	nd	1	1
Rbm8	nd	2	2	nd	1	1
Lsm5	2	2	1	nd	nd	nd
Ubp10	nd	2	2	nd	2	2
Snu13	1	1	1	1	1	1
Lsm4	1	1	1	nd	2	2
Ctnnbl1	4	1	0.25	1	2	2
Cwf20	nd	1	1	nd	nd	nd
Hub1	1	nd	nd	1	nd	nd

Table 3.1: Hub1 selectively modifies the spliceosome. The number of unique peptides in mass spectrometry for an identified protein. Cdc5-6HA and Prp19-6HA complexes were immunoprecipitated using anti-HA antibody from wild-type and *hub1-1* strains cell lysates. Co-IP proteins were subjected and analyzed by mass spectrometry. Nd, not detected.

3.2.3 Hub1 surfaces I and II are not critical for growth and splicing in *S. pombe*

In *S. cerevisiae*, two independent surfaces of Hub1 were reported for its function in splicing. Depicted Hub1 protein alignment in different eukaryotes, showed that D22 and H63 residues are conserved (Figure 3.3A). Through its Asp-22 surface I, it binds to splicing factors Snu66 and/or Prp38, and through its His-63 surface II, it interacts and activates the RNA helicase Prp5. To understand the mechanism of Hub1 in pre-mRNA splicing in *S. pombe*, we generated chromosomal mutants *hub1-D22A* (surface I), *hub1-H63L* (surface II), and *hub1-D22A H63L* (surface I, II). I compared the growth phenotype of *hub1* surface mutants with wild-type cells. The *hub1-D22A* (surface I) and *hub1-D22A H63L* (surface I, II) surface mutants showed moderate growth defects compared to wild-type cells, but *hub1-H63L* (surface II) mutant grew similar to the wild-type strain (Figure 3.3B). Similar to our observation, *hub1* surface I mutant deficient in binding to Snu66 were viable (Mishra et al., 2011). I further

monitored splicing defects in *hub1* surface mutants using RT-PCR assays for the selected targets that showed an accumulation of pre-mRNA in the microarray. *hub1* surface mutants showed mild accumulation of intron-containing transcripts (compared to the wild-type cells) for the selected pre-mRNA at non-permissive temperature 37°C (Figure 3.3C). Similar to these findings, expression of *hub1* surface I mutant construct in human cells complemented the splicing defects in RNAi mediated Hub1 deficient cells (Ammon et al., 2014). These results clearly showed that, as also reported previously (Mishra et al., 2011), Hub1-Snu66 interaction through surface I is not essential for growth and splicing. Further, we infer from our study that the Hub1 surface I and surface II might not be critical in *S. pombe* and likely also in other eukaryotes.

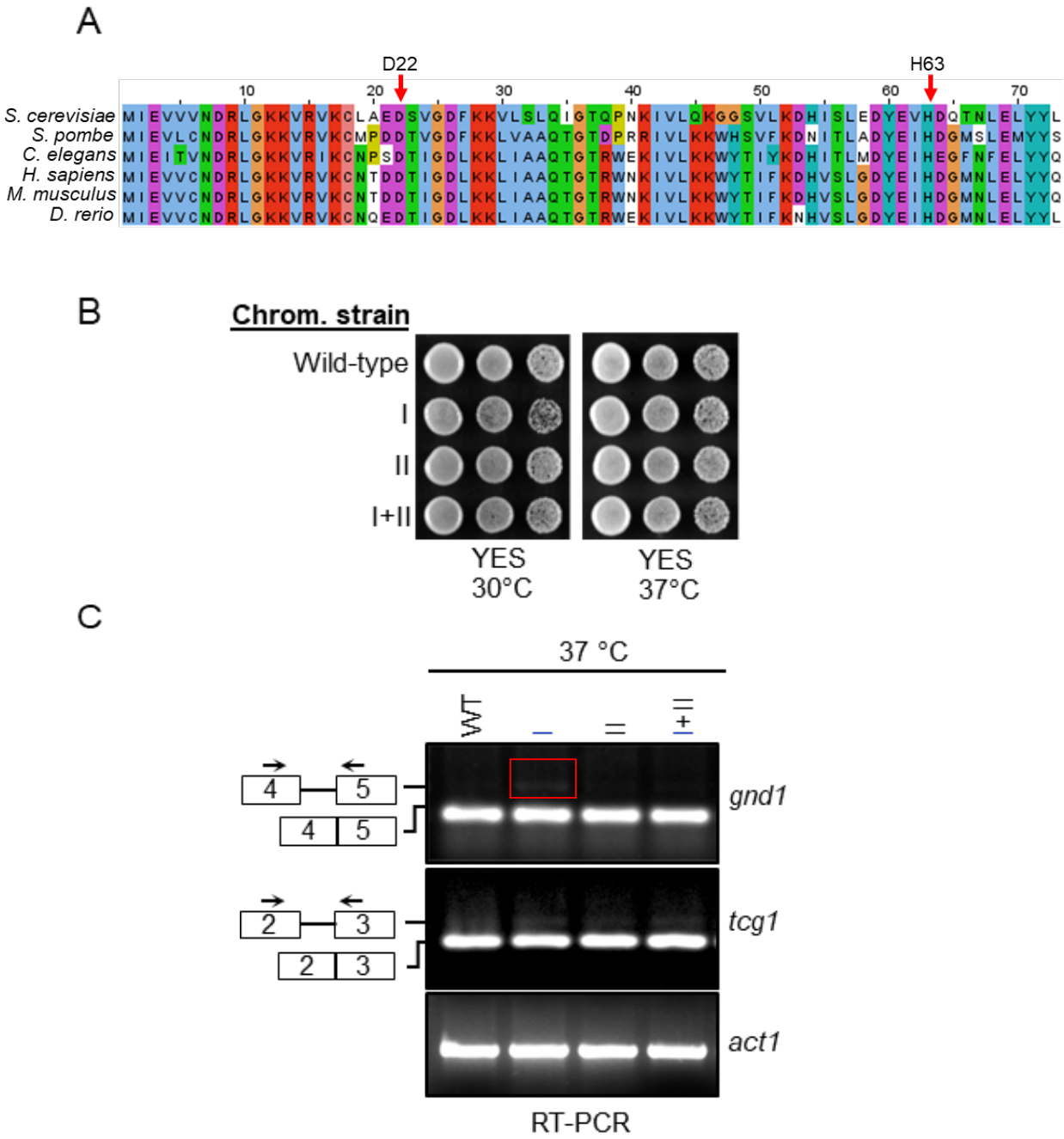


Figure 3.3: *hub1* surface mutants are not critical for splicing in *S. pombe*

A. Multiple sequence alignment of Hub1 protein from different eukaryotes. The red arrow shows the position of the Aspartate residue D22 and Histidine residue H63. The alignment was visualized with JalView. (Waterhouse et al., 2009).

B. Growth phenotypes of wild-type (SP9), *hub1-D22A* (surface I) (P2), *hub1-H63L* (surface II) (P3), *and hub1-D22A H63L* (surface I+II) (K2) mutants were five-fold serial diluted on rich media. Plates were incubated at 30°C and 37°C for 3-4 days.

C. Semi-quantitative RT-PCR assay of strains in (A) shows splicing of transcripts for *gnd1* and *tcg1*. cDNA prepared from total RNA isolated from wild-type and *hub1*

surface mutant strains. The experiment is similar to Figure 3.1B. The red box shows the accumulation of intron-containing transcript.
(Note: Figure 3.3B, C are cropped from Figure 3.5D, E respectively).

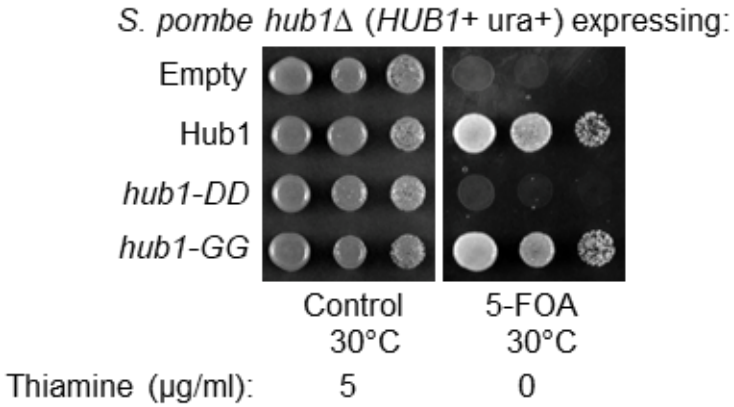
3.2.4 The functional C-terminus surface is essential for Hub1 activity

As the *hub1* surface mutants did not show any growth and splicing defects, we speculated that Hub1 might have unknown surfaces in intron-rich *S. pombe*. Previously, it has been shown that *hub1* mutants with charged C-terminal extensions were unable to support growth in *S. pombe hub1Δ* cells (Mishra et al., 2011); however, the molecular details are lacking. To identify new surfaces relevant for growth and splicing, I performed complementation assays in *S. pombe hub1Δ* strain with the expression of *hub1* C-terminal variants *hub1-GG* and *hub1-DD* extensions. As reported previously (Mishra et al., 2011) *S. pombe hub1Δ* lethality was complemented with *hub1-GG* mutant expression; by contrast, the expression of the *hub1-DD* mutant was unable to rescue the lethality (Figure 3.4A), indicating that a free C-terminus of Hub1 is essential for its function. Similarly, in *hub1-1* strain, the expression of *hub1-GG* variant resulted in functional Hub1, but *hub1-DD* variants led to non-functional Hub1. The expression of *hub1-DD* variant also showed dominant-negative phenotype at 30°C (Figure 3.4B). I performed RT-PCR experiments in *hub1-1* strain with *hub1* variants to assess the efficiency of pre-mRNA splicing for the selected targets from the splicing-sensitive microarray. The expression of the *hub1-DD* variant led to accumulation of intron-containing transcript at 37°C for the selected targets, but the expression of the *hub1-GG* variant did not lead to accumulation of intron containing trasncripts (Figure 3.4C). To understand the molecular details of *hub1* variants, we performed GST pull-down experiments using recombinantly purified *6xHis-hub1-GG*, *6xHis-hub1-DD* variants, and GST-Snu66 (HIND). GST-Snu66 were able to pull-down both the variants of *hub1* (Figure 3.4D). Thus, it indicates that Hub1 binding to Snu66 is not essential for its function, which further testifies our observation that *hub1* surface I mutant does not show any considerable growth and splicing defects.

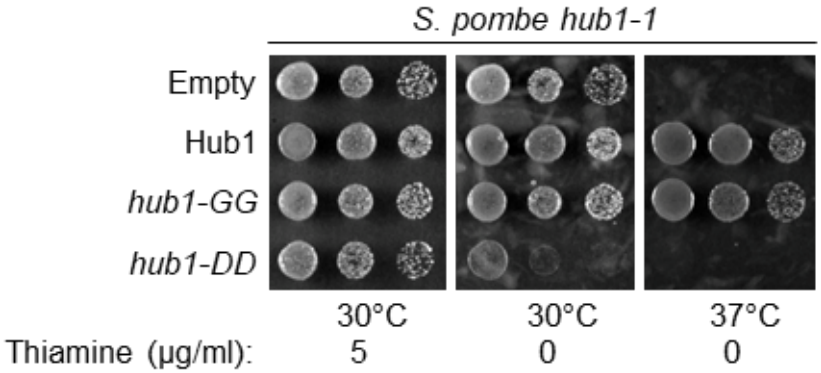
To understand whether any structural rearrangements on *hub1* variant led to the expression of non-functional Hub1, we performed circular dichroism (CD) and 3D

NMR experiments (in collaboration with Ranabir Das at NCBS, Bangalore). CD spectra overlay was almost identical for both the *hub1* variants, which implies that there were no major changes in secondary structure of the mutant protein (Figure 3.4E). We then performed 3D NMR experiments which revealed that *hub1-DD* extensions might form a salt bridge with R9 and R41 residues on itself (data are not shown). Thus, we speculate that non-functionality of the *hub1-DD* variant may be due to a hinderance of R9 R41 residues.

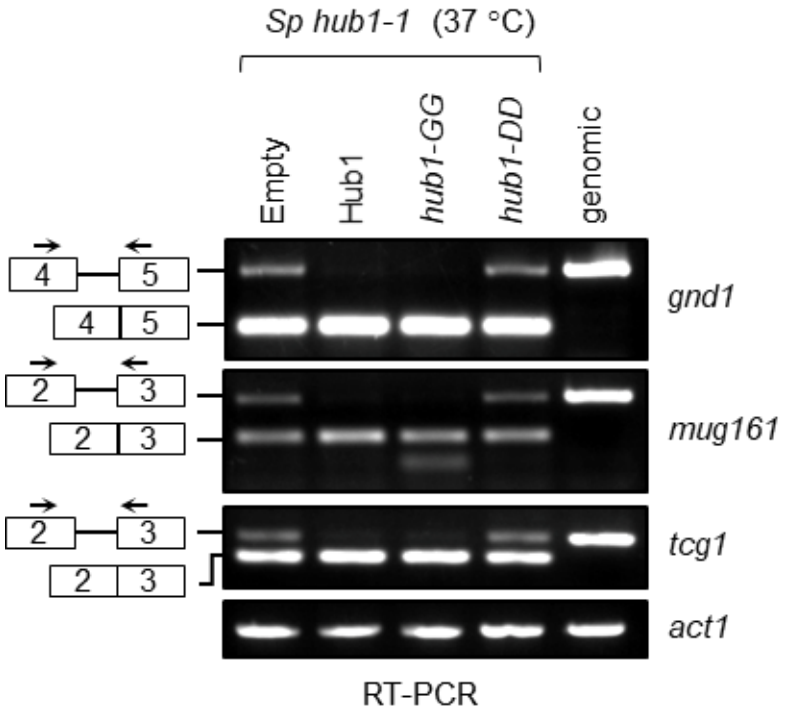
A


S. pombe hub1Δ (HUB1+ ura+) expressing:
Empty
Hub1
hub1-DD
hub1-GG
Control
30°C
5-FOA
30°C
Thiamine (µg/ml):
5
0

B


S. pombe hub1-1
Empty
Hub1
hub1-GG
hub1-DD
30°C
30°C
37°C
Thiamine (µg/ml):
5
0
0

C


Sp hub1-1 (37 °C)
Empty
Hub1
hub1-GG
hub1-DD
genomic
4
5
4
5
gnd1
2
3
2
3
mug161
2
3
2
3
tcg1
act1
RT-PCR

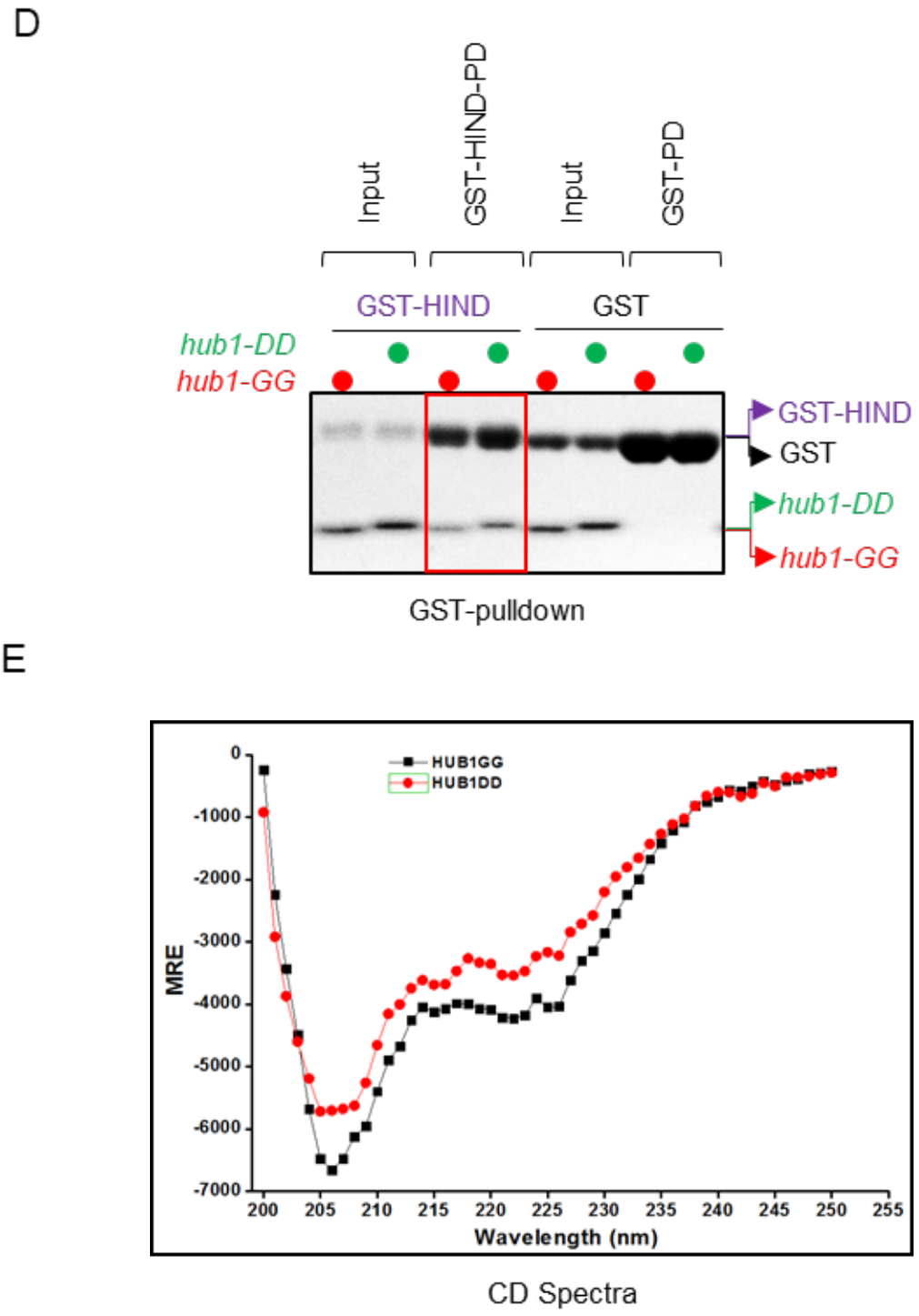


Figure 3.4: A free C-terminus is critical for Hub1's splicing function

A. Complementation assay of *S. pombe hub1Δ*, using the expression of *hub1* variants. *hub1-GG* variant lead to functional Hub1, by contrast, *hub1-DD* gave rise to non-functional Hub1. For complementation in *S. pombe, a URA4*-bearing plasmid expressing wild-type *S. pombe* Hub1 was shuffled-out from the *hub1Δ* strain by counter-selection with 5-FOA. Five-fold serial diluted cells were spotted on control or FOA-containing plates. Plasmids were expressed from the weak nmt81 promoter.

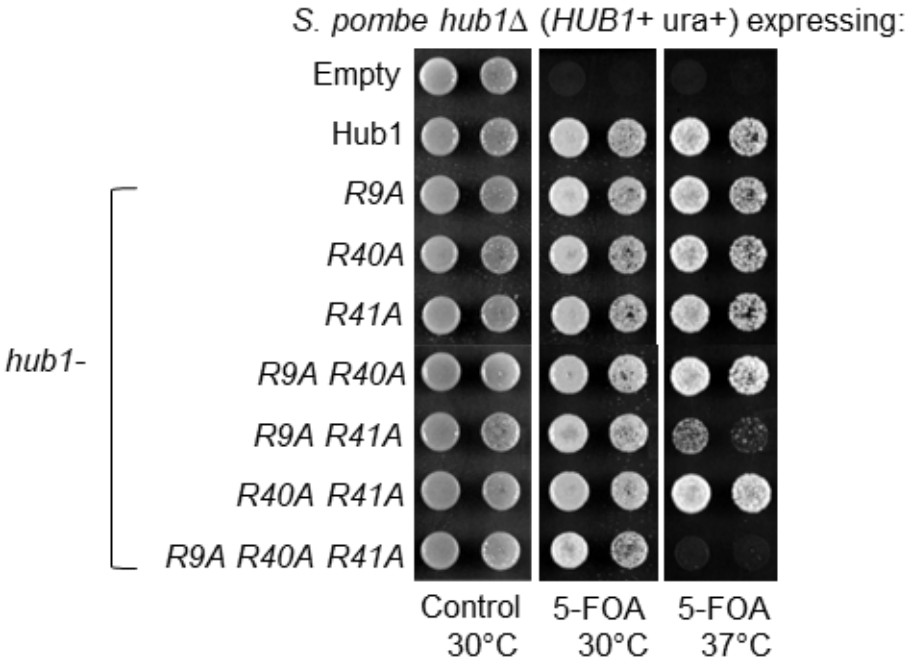
B. Complementation assay of *S. pombe hub1-1* cells using the expression of *hub1* variants. *hub1-GG* variant leads to functional Hub1, by contrast, *hub1-DD* gave rise to non-functional Hub1. Five-fold serial diluted cells were spotted on selective media at 30°C and 37°C for 3-4 days. *hub1-DD* construct shows dominant-negative phenotype at the permissive temperature 30°C.
C. Semi-quantitative RT-PCR assay of strains in (B) was performed. *hub1-DD* variant shows splicing defects for *gnd1, mug161,* and *tcg1 transcripts*. cDNA prepared from total RNA isolated from the *hub1-1* strain. The experiment is similar to Figure 3.1B.
D. GST–pull-down assays were carried using GSH beads. Bacterially purified equimolar concentrations of *6xHis-hub1-GG, 6xHis-hub1-DD* were mixed with GST, GST-Snu66 (only the HIND part of Snu66 was used). *hub1* variants were not defective in binding to Snu66(HIND), GST was used as a negative control. Inputs represent one-tenth amount of total proteins used. The red box shows GST-HIND interaction with *hub1-DD* and *hub1-GG* proteins. GST and GST-Snu66(HIND) were purified according to standard GST-tag affinity purification, whereas, *6xHis-hub1-GG* and *6xHis-hub1-DD* were purified according to standard His-tag affinity purification as described in the methods section (2.2.17). The experiment was performed once only.
E. The far-UV circular dichroism spectra for *hub1-GG* and *hub1-DD* variants. (Obtained in collaboration with Ranabir Das at NCBS, Bangalore). *6xHis-hub1-GG* and *6xHis-hub1-DD* were purified according to standard His-tag affinity purification as described in methods section (2.2.17). Far UV CD measurements were carried out on Jasco J-815 spectropolarimeter. 10 µM protein concentration was used for CD measurements in 0.1 cm path length cuvette. Buffer scans were subtracted from protein's scans. The samples were mixed properly before every measurement.

3.2.5 Hub1 possesses a novel surface III critical for growth and splicing

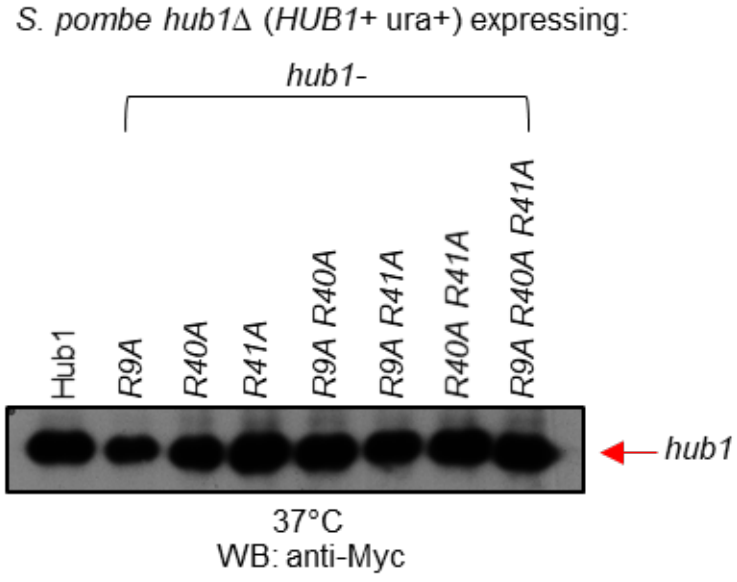
To understand the importance of R9 and R41 residues on Hub1 surface, I performed complementation assay in *S. pombe hub1Δ* cells expressing various *hub1-R* mutants. The expression of all single R-A mutants, but not the double or triple mutants, *hub1-R9A R41A and hub1-R9A R40A R41A*, lead to functional Hub1 (Figure 3.5A). Both *hub1-R9A R41A* and *hub1-R9A R40A R41A* mutants are unlikely to be structural mutants of the protein, as the protein level remained unaltered (Figure 3.5B). We compared Hub1 protein alignment in different eukaryotes, which showed that R9 and R41 residues are conserved (although the R41 residue in *S. pombe* is replaced by K41 in other organisms), but the R40 residue of *S. pombe* Hub1 is not conserved (in other eukaryotes it is replaced either by asparagine or glutamic acid) (Figure 3.5C). Henceforth, I generated *hub1-R9A R41A* chromosomal mutant and performed growth phenotype and splicing assays. The newly identified *hub1-R9A R41A* mutant showed growth defects similar to the *hub1-1* strain at non-

permissive temperature 37°C (Figure 3.5D). But unlike the *hub1-1* mutant which is likely a structural mutant (since its protein level is strongly diminished at all temperatures), *hub1-941* is not a structural mutant (its protein level is not altered). It implies that Hub1 might mediate its critical functions through R9 and R41 residues. I further monitored splicing defects in *hub1-941* strain using RT-PCR assays for the selected targets which had earlier showed accumulation of pre-mRNA in *hub1-1* strain. The *hub1-R9A R41A* mutant also showed accumulation of intron-containing transcripts for the selected pre-mRNA at non-permissive temperature 37°C (Figure 3.5E). Hence, from our growth and splicing assays, we speculate that Hub1 might form an essential novel surface III with R9 and R41 residues (Figure 3.5F).

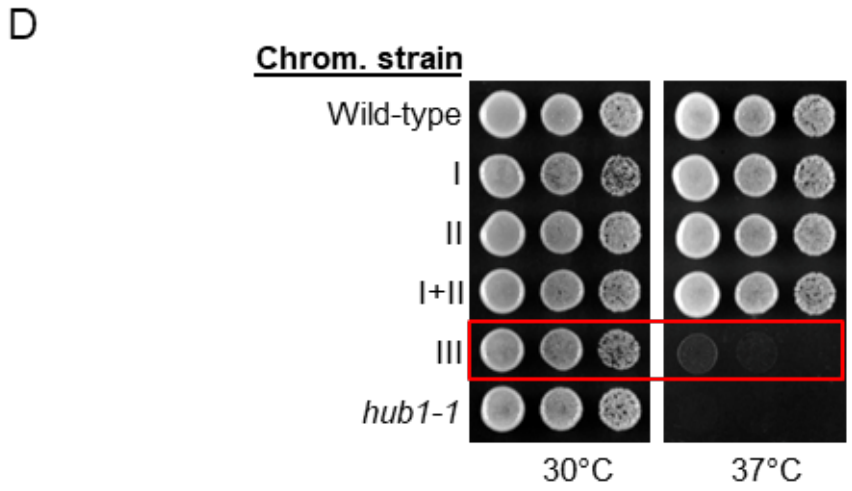
A

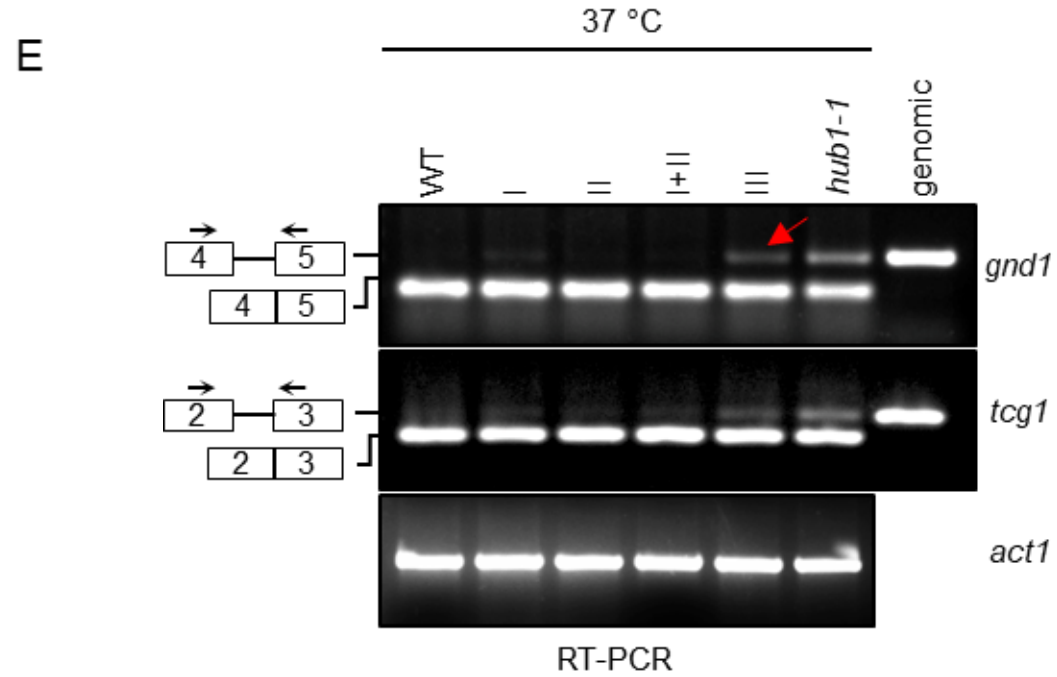

S. pombe hub1Δ (HUB1+ ura+) expressing:
Empty
Hub1
R9A
R40A
R41A
hub1-
R9A R40A
R9A R41A
R40A R41A
R9A R40A R41A
Control 30°C
5-FOA 30°C
5-FOA 37°C

B


S. pombe hub1Δ (HUB1+ ura+) expressing:
hub1-
Hub1
R9A
R40A
R41A
R9A R40A
R9A R41A
R40A R41A
R9A R40A R41A
hub1
37°C
WB: anti-Myc


C
R9
R41/K41
10
20
30
40
50
60
70
S. cerevisiae
S. pombe
C. elegans
H. sapiens
M. musculus
D. rerio


D
Chrom. strain
Wild-type
I
II
I+II
III
hub1-1
30°C
37°C


E
37 °C
WT
I
II
I+II
III
hub1-1
genomic
4
5
4
5
gnd1
2
3
2
3
tcg1
act1
RT-PCR

F

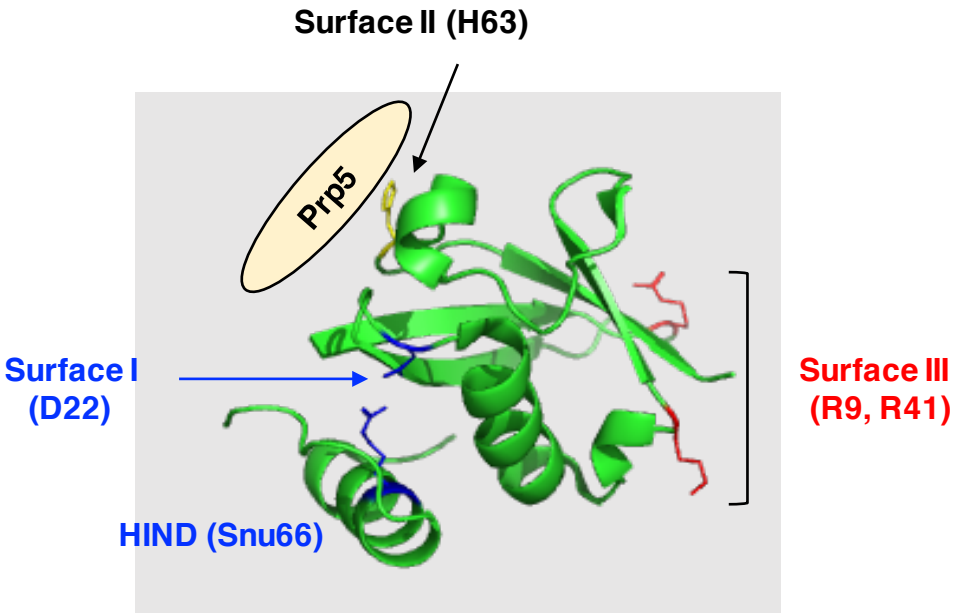


Figure 3.5: Hub1-R9 R41 residues are critical for growth and splicing function

A. Complementation assay of *S. pombe hub1Δ* (SP13) by *hub1* variants. *hub1-R9A, hub1-R40A, hub1-R41A, hub1-R9A R40A, hub1-R9A R41A, hub1-R40A R41A*, and *hub1-R9A R40A R41A* variants. The expression of *hub1-R9A R41A and hub1-R9A R40A R41A* variants gave rise to weakly functional Hub1, and other variants lead to functional Hub1. The experiment is similar to Figure 3.4A. WT, *hub1* mutants were expressed from the Hub1 endogenous promoter.

B. The transformants was as above (A), *S. pombe, a URA4*-bearing plasmid expressing wild-type *S. pombe* Hub1 was shuffled-out from the *hub1Δ* strain by counter-selection with 5-FOA and the level of *hub1* variants were monitored at non-permissive temperature 37°C by western blotting (WB) using anti-MYC antibody.

C. Multiple sequence alignment of Hub1 protein from different eukaryotes. The red arrow shows the position of the arginine residue R9 and the arginine or lysine residues at R41/K41.The alignment was visualized with JalView. (Waterhouse et al., 2009).

D. Growth phenotypes of wild-type, *hub1-D22A* (surface I), *hub1-H63L (*surface II*), hub1-D22A H63L (*surface I+II*), hub1-R9A R41A (*surface III*)* (P46) and *hub1-1* mutant (SP10). The experiment is similar to Figure 3.3B. The red box shows the *hub1-R9A R41A* (surface III) temperature sensitivity.

E. Semi-quantitative RT-PCR assay of strains as in (D) shows splicing of transcripts for *gnd1* and *tcg1.* The experiment is similar to Figure 3.1B. The red arrow indicates the accumulation of intron-containing transcript in *hub1-R9A R41A* (surface III) mutant.

F. Structure of Hub1-Snu66 (HIND) complex (PDB ID 3PLV; modified). Hub1-HIND interaction residues are marked in blue (surface I), Hub1-Prp5 interacting residue is marked in yellow (surface II), and the newly identified residues are in red (surface III).

3.2.6 *hub1* surface III is different from the HIND binding surface

Hub1 interacts with the spliceosomal protein component Snu66 (tri-snRNP complex) in yeast and humans (Mishra et al., 2011; Ammon et al., 2014). To understand whether *hub1* surface III mutant alters HIND-binding surface, we performed a yeast two-hybrid assay. In this assay, I fused Snu66-HIND to the Gal4 binding domain (BD), Hub1, and *hub1* surface III mutant to the Gal4 activation domain (AD). The yeast cells co-transformed with Hub1/ *hub1 III* mutant and Snu66 constructs were viable on histidine lacking plates, which indicates the interaction between the Hub1 and *hub1* surface III mutant with Snu66-HIND (Figure 3.6). Thus, the growth and splicing defects in *hub1* surface III mutant (Figure 3.5D, E) might be due to loss of interaction with certain other factors.

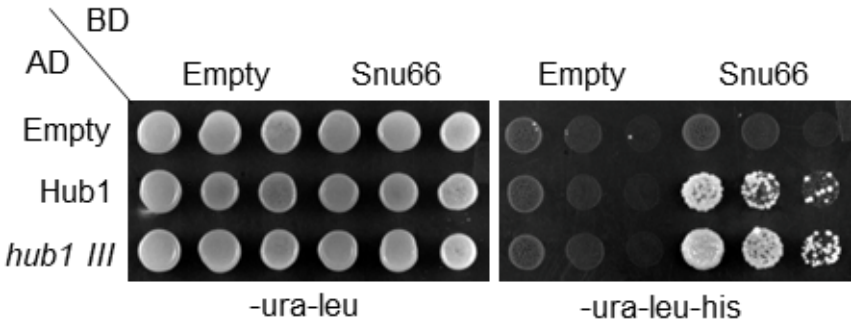


Figure 3.6: *hub1* surface III mutant interacts with Snu66-HIND

In yeast two-hybrid assay pGBDU1-Snu66-HIND, pGADC1-Hub1, and pGADC1-*hub1* surface III mutant constructs were co-transformed, and transformants were monitored for interaction on SC-Ura-Leu-His plates. In the case of physical interaction between the two fusion proteins reporter gene expression (HIS3) was induced, which allowed growth on the selection medium for 3-4 days.

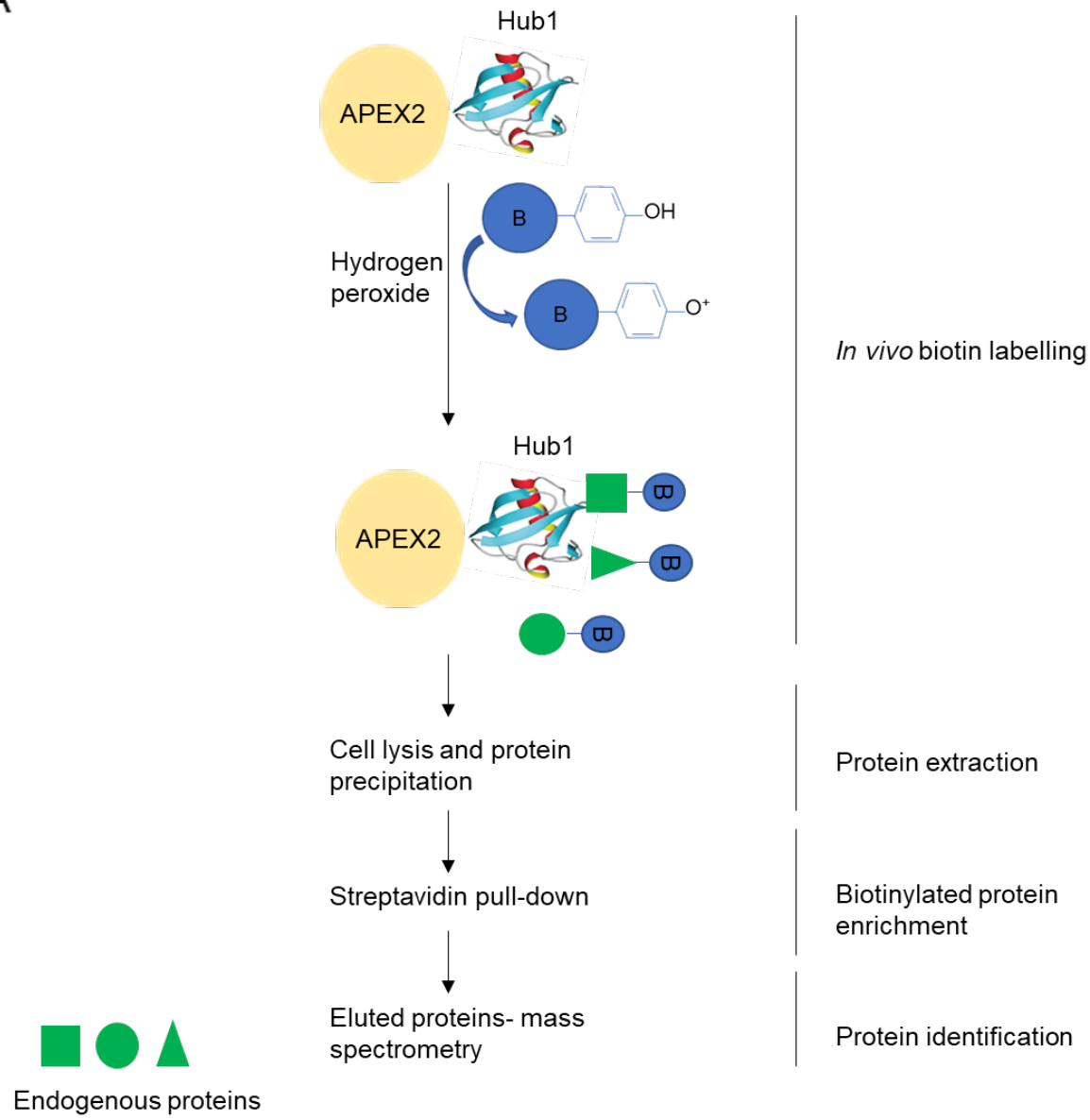
3.2.7 Screen to identify Hub1 interactors

To screen for putative interactors of Hub1, we employed conventional methods such as yeast two-hybrid assay and immunoprecipitation experiments. Neither of the experiments led to a positive new interactor. It is possible that Hub1 might interact with its binding partners only transiently or antibody recognition of the complex is hindered. Therefore, I performed proximity-dependent biotin labeling using the engineered ascorbate peroxidase (APEX2) method coupled to mass spectrometry.

This method enables proteomic mapping of subcellular compartments and further to identify dynamic protein complexes. I fused APEX2 to the N-terminus of Hub1, which was expected to covalently tags protein substrates in close vicinity of Hub1 with biotin-phenol (BP) when hydrogen peroxide is added to the cells (Hwang et al., 2016). We then expressed APEX2-Hub1 and *APEX2-hub1-DD* constructs in *S. pombe* cells treated with biotin-phenol in the presence of hydrogen peroxide to stimulate biotinylation of substrates. Biotinylated proteins were affinity purified using streptavidin beads, and eluted samples were subjected to mass spectrometry. List of proteins identified from the APEX2 screen is shown in Appendix Table 1. From the proximity-dependent labeling, ideally, we would expect Hub1-specific peptides. As we were unable to identify biotin-labelled peptides specific to Hub1 itself, suggesting this experimental approach might not have worked well.

I then carried out a high-copy suppressor screen to identify suppressors of temperature sensitivity of *hub1-941* mutant using a *S. pombe* cDNA library. For this purpose the *hub1* surface III mutant was transformed with a high-copy pREP1 vector-based *S. pombe* cDNA library and plated at 37°C on selective media. Further, 11 plasmids were shuttled out from the transformants which grew at 37°C . Out of 11 plasmids, a Hub1 clone and another clone (DR9) was reproducibly identified as the suppressor of *hub1* surface III mutant (Figure 3.7B). DR9 is a hybrid clone of SPCPB16A4.06C (*Schizosaccharomyces* specific protein) and Tcg1 (single-stranded telomeric binding protein). As Tcg1 sequence lacks 1st 76 nucleotides and in-frame translation start codon, this was unlikely to be the suppressor. We then subcloned SPCPB16A4.06C ORF region into the expression vector, pREP3x. *hub1* surface III mutant transformed with SPCPB16A4.06C ORF rescued the growth defect at 37°C (Figure 3.7C). This dosage rescue could be either due to increased level of protein or RNA. Further tests needs to be carried out to establish the mechanism of SPCPB16A4.06C as a high-copy suppressor of *hub1* surface *III* mutant.

A



B

hub1 III mutants were transformed with *S. pombe* cDNA library

↓

Initially transformed cells were kept at 30°C for 6 hrs

↓

Later cells were shifted to 37°C for 7-8 days

↓

Plasmids were rescued from the transformants that grew at 37°C

↓

We isolated 11 plamids from the transformants that grew at 37°C

C

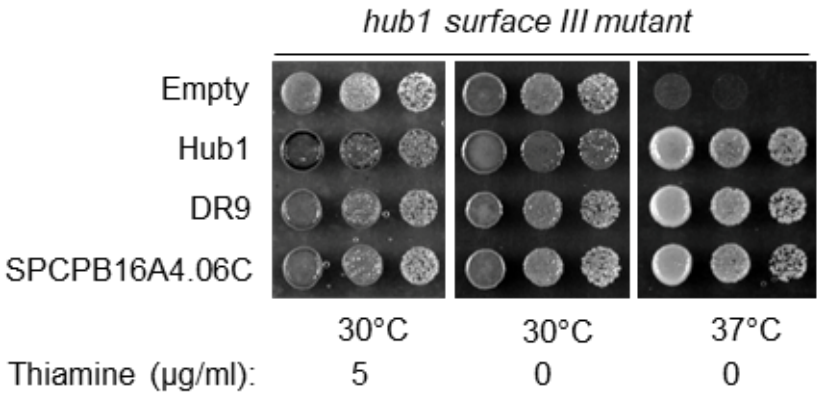


Figure 3.7: High-copy suppressor of *hub1* surface III mutant

A. Schematics depicting APEX2 labeling protein substrates with biotin-tag (adapted from Hwang et al., 2016). APEX2—Hub1 was expected to biotinylate proteins in close proximity to Hub1. APEX2—Hub1 construct was transformed in *S. pombe hub1Δ* strain *(URA4*-bearing plasmid expressing wild-type *S. pombe* Hub1 was shuffled-out from the *hub1Δ* strain by counter-selection with 5-FOA). *APEX2-hub1-DD* construct was transformed in wild-type *S. pombe* strain.
B. Schematic overview of the high-copy suppressor screen for *hub1* surface III mutant. High-copy cDNA library from *S. pombe* was used in the screen.
C. *hub1* III mutant cells transformed with Empty vector, Hub1 clone (expressed from the Hub1 endogenous promoter), multi-copy suppressor pREP1-DR9 and pREP3X-SPCPB16A4.06C clones under thiamine repressible promoter. Hub1, DR9, and SPCPB16A4.06C clones rescued the lethality at 37°C . Transformed cells were five-fold serial diluted on selective media. Plates were incubated at 30°C and 37°C for 3-4 days (pREP3X-SPCPB16A4.06C clone was made by Arundathi).

3.2.8 Hub1 surface III is functionally conserved in eukaryotes

An important question was whether Hub1 surface III is conserved in eukaryotes as that of Hub1. I compared Hub1 protein alignment from different eukaryotes, which showed that R9 and R41 residues are conserved (R41 residues in *S. pombe,* is replaced by K41 in other organisms) (Figure 3.5C). I monitored the expression of *C. elegans* Hub1 in *S. pombe hub1Δ* strain. The expression of *C. elegans* Hub1 complemented lethality in *S. pombe hub1Δ* strain (Figure 3.8), implying that its function is conserved across the organisms. I then complemented with *C. elegans hub1-R9A K41A* to understand the importance of identified Hub1 surface III. The expression of *C. elegans hub1-R9A K41A* was unable to complement *S. pombe*

hub1Δ cells lethality at non-permissive temperature 37°C (Figure 3.8). It implies that R9 K41 residues which are part of the Hub1 surface III are essential for its function. Thus, we infer that Hub1 surface III is a crucial and functionally conserved surface of Hub1 molecule.

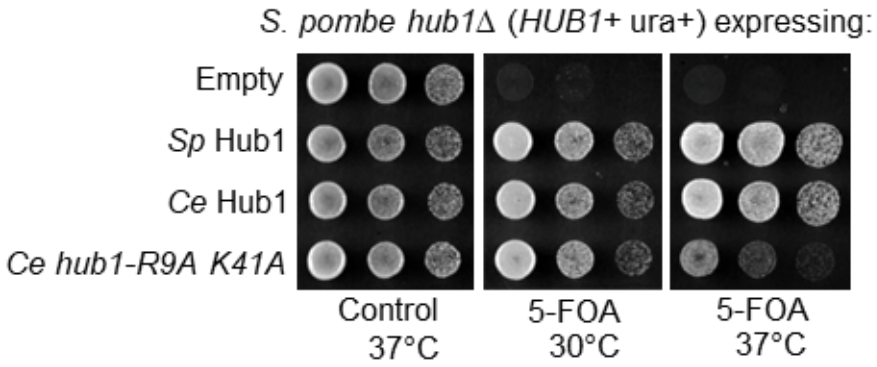


Figure 3.8: Hub1 surface III is functionally conserved in eukaryotes

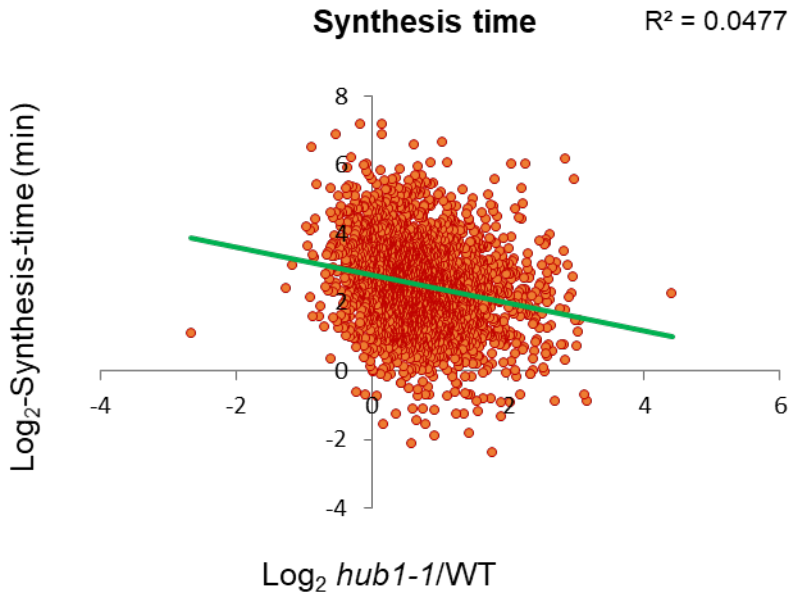
Complementation assay of *S. pombe hub1Δ* by *C. elegans* Hub1 and its *hub1-R9A K41A* variant. *C. elegans hub1-R9A K41A* variant was unable to complement, but *C. elegans* Hub1 complemented *S. pombe hub1Δ* lethality. The experiment is similar to Figure 3.4A. *C. elegans* Hub1 constructs were expressed from the *S. pombe* HUB1 endogenous promoter.

3.2.9 Hub1 promotes splicing of transcripts which are synthesized faster

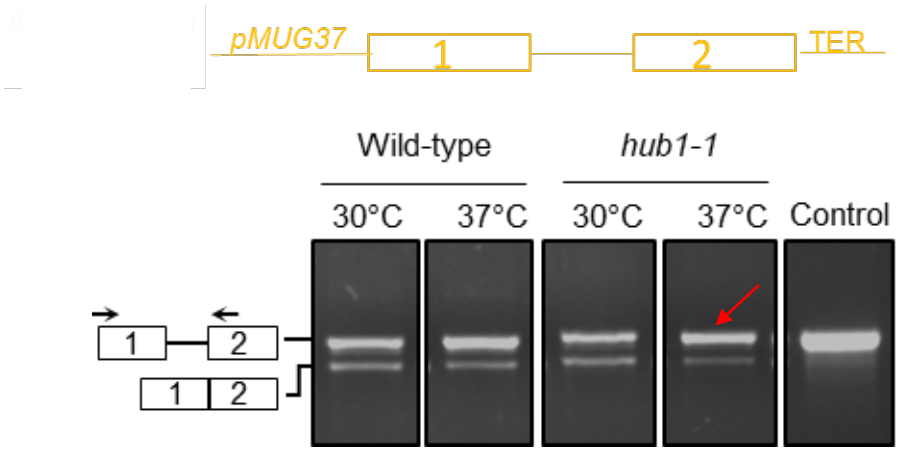
To understand whether Hub1-dependent transcripts show any specific feature with respect to the strengths of their splicing signals, lengths, and other features, we performed bioinformatic analysis in collaboration of Arashdeep Singh from the laboratory of Dr. Kuljeet Sandhu in our department. The Hub1-dependent intron-containing genes showed no obvious characteristic features, except with their rates of RNA synthesis. We found positive correlation between synthesis time and splicing defects in *hub1-1* strain (Figure 3.9A). The synthesis time was derived from a recent study on RNA metabolism in *S. pombe* genome (Eser et al., 2016). Intron-containing genes that are synthesized faster tend to correlate better with Hub1's splicing targets. To validate whether Hub1 promotes splicing of the transcripts that are synthesized faster, we chose two transcripts *sod2* (synthesized faster-1.82 minutes) and *mug37* (synthesized slower-90 minutes) for further analysis. Genomic clones of *sod2* and *mug37* with own promoter and terminator were transformed in *S. pombe*

wild-type and *hub1-1* strains to monitor the splicing defects. In addition, the genomic clone of *sod2* with *mug37* promoter and the genomic clone of *mug37* with *sod2* promoter were transformed to monitor the splicing defects. To monitor the splicing defects, RT-PCR experiments were carried out for the *sod2 and mug37* targets. The *mug37* transcript showed accumulation of intron-containing transcripts in both wild-type and *hub1-1* cells (Figure 3.9B). Whereas, *sod2* transcript showed accumulation of intron-containing transcripts only in *hub1-1* cells (Figure 3.9C), indicating that Hub1 specifically promoted splicing of *sod2* transcripts which are synthesized faster. In addition, upon monitoring splicing defects for the promoter-swapped *sod2* transcripts with mug37 promoter, accumulation of intron-containing transcript were observed in both wild-type and *hub1-1* cells (Figure 3.9D). Whereas promoter swapped *mug37* transcripts with *sod2* promoter showed accumulation of intron-containing transcripts only in *hub1-1* cells (Figure 3.9E). Altogether, above observations showed that *sod2* transcript (synthesized faster) showed accumulation of intron-containing transcripts only in *hub1-1* cells. Additionally, the slowly synthesized transcript, when replaced with the promoter of the gene which is transcribed faster, now showed preferential accumulation of transcript in *hub1-1* cells. Thus, indicating that the Hub1 specifically promotes splicing of transcripts which are synthesized faster.

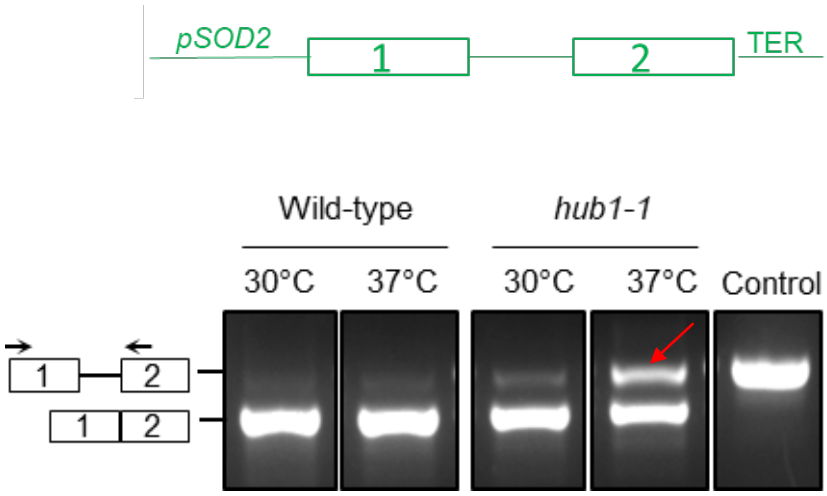
A



B


pMUG37
1
2
TER
Wild-type
hub1-1
30°C
37°C
30°C
37°C
Control
1
2
1
2

C


pSOD2
1
2
TER
Wild-type
hub1-1
30°C
37°C
30°C
37°C
Control
1
2
1
2

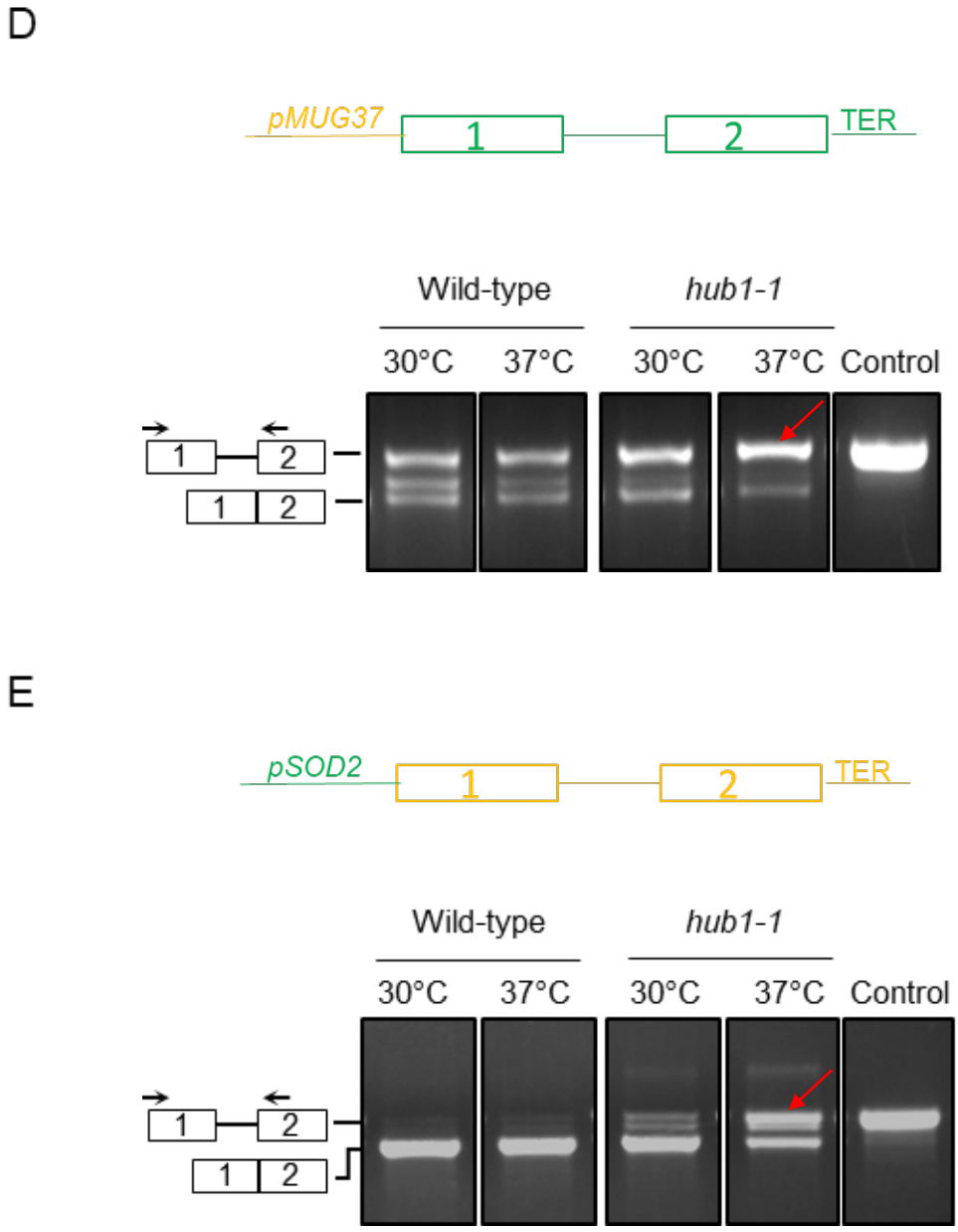


Figure 3.9: Hub1 promotes splicing of transcripts which are synthesized faster

A. Correlation graph depicting the positive correlation between synthesis time and splicing defects in *hub1-1* strain (SP10) (Data is from Arashdeep Singh).

B. Semi-quantitative RT-PCR shows accumulation of intron-containing transcripts for *mug37.* Genomic clone of *mug37* with own promoter and terminator were transformed in *S. pombe* Wild-type and *hub1-1* strains. The experiment is similar to Figure 3.1B. Primers are specific for *mug37* to monitor accumulation of intron-containing transcripts. *Mug37* transcript time of synthesis was reported to be 90

minutes (Eser et al., 2016); splicing defect (log2 *hub1-1*/Wild-type ratio of intron-containing signals from our microarray) was -0.88 neagtive. The red arrow indicates the accumulation of intron-containing transcript.

C. The experiment was as in (B). Semi-quantitative RT-PCR shows accumulation of intron-containing transcripts for *sod2 in hub1-1 strain.* Genomic clone of *sod2* with own promoter and terminator were transformed in *S. pombe* Wild-type and *hub1-1* strains. *Sod2* transcript time of synthesis was reported to be 1.82 minutes (Eser et al., 2016); splicing defect (log2 *hub1-1*/Wild-type ratio of intron-containing signals from our microarray) was 2.4 positive. The red arrow indicates the accumulation of intron-containing transcript.
D. The experiment was as in (B). Semi-quantitative RT-PCR shows accumulation of intron-containing transcripts for *sod2.* Genomic clone of *sod2* with *mug37* promoter and its terminator were transformed in *S. pombe* Wild-type and *hub1-1* strains. Primers are specific for *sod2* to monitor accumulation of intron-containing transcripts. The red arrow indicates the accumulation of intron-containing transcript.
E. The experiment was as in (B). Semi-quantitative RT-PCR shows accumulation of intron-containing transcripts for *mug37 in hub1-1 strain.* Genomic clone of *mug37* with *sod2* promoter and its terminator were transformed in *S. pombe* Wild-type and *hub1-1* strains. Primers are specific for *mug37* to monitor accumulation of intron-containing transcripts. (Note: Figure B, C, D, E were cropped from different gel images and the experiment was performed once only). The red arrow indicates the accumulation of intron-containing transcript. Primers are depicted with black arrows on MYC-F and exon specific-R in Figure 3.9B,C,D and E. Here, green color represents genomic clone of *sod2* with own promoter and terminator and orange represents genomic clone of *mug37* with own promoter and terminator.

3.3 Discussion

UBLs (Ubiquitin-like proteins) covalently conjugate with the protein substrates and post-translationally modify them for various cellular functions (Taherbhoy et al., 2012). Though structurally similar to UBLs, Hub1 functions differently from ubiquitin and other ubiquitin-like proteins through non-covalent interactions with protein substrates (McNally et al., 2003; Ramelot et al., 2003; Luders et al., 2004; Wilkinson et al., 2004; Yashiroda and Tanaka, 2004; Mishra et al., 2011). In the current study, we show that Hub1 specifically modulates the composition of the spliceosome and promotes pre-mRNA splicing in *S. pombe*. We identified a novel surface of Hub1, surface III, and perturbations on this surface causes splicing and growth defects. Further, SPCPB16A4.06C (*Schizosaccharomyces* specific protein) was identified as a high-copy suppressor of *hub1* surface III mutant.

Hub1 is required for pre-mRNA splicing

In *S. cerevisiae*, Hub1 modulates the spliceosome to process the transcripts with non-canonical splice sites to promote alternative splicing (Mishra et al., 2011). As non-canonical SSs are recognized poorly and hinder spliceosomal assembly, Hub1 might potentiate the correct incorporation and spliceosomal assembly to facilitate splicing. It regulates the splice site selection in association with Prp5: low levels of Hub1 is sufficient for splicing of optimal introns, whereas high levels of Hub1 not only activate the spliceosome but also influences splicing accuracy (Karaduman et al., 2017). In *S. cerevisiae*, Hub1 is upregulated in cadmium stress condition to promote stress-induced splicing of introns with non-canonical SSs (Chanarat and Svasti, 2019). Similarly, in perennial ryegrass, Hub1 is induced by drought, and overexpression of Hub1 confers resistance to drought (Patel et al., 2015). Whereas, in *S. pombe* a subset of targets are shown to depend on Hub1 for pre-mRNA splicing (Wilkinson et al., 2004; Mishra et al., 2011). In relatively intron-rich *S. pombe* Hub1 may act as a spicing modulator to regulate the splicing of certain intron-containing pre-mRNAs. In human cell lines, a large number of pre-mRNAs depend on Hub1 for optimal and correct splicing of its introns. Hub1 sensitive introns did not show any obvious sequence similarity (Ammon et al., 2014). In this study by using splicing-sensitive microarray, we show that splicing of a large number genes was affected in

hub1-1 strain at 37°C . Hub1 functions as intron-specific splicing factor and is a unique regulator of spliceosome due to following reasons; (i) it is essential for viability (Wilkinson et al., 2004; Yashiroda and Tanaka, 2004), (ii) expressed at relatively lower level compared to other splicing factors (Karaduman et al., 2017). (iii) required for splicing of large subset of genes, (iv) transiently associated with components of the spliceosome (Wilkinson et al., 2004). Being a unique regulator of the spliceosome, it will be interesting to answer the following; (i) to mechanistically identify how the levels of Hub1 is increased with respect to different stress conditions, (ii) does the increased levels of Hub1 modulate splicing of intron-containing transcripts differently, (iii) does Hub1 sensitive introns show any obvious characteristic features.

Hub1 modifies the composition of the spliceosome

In *S. cerevisiae*, it has been shown that Hub1 influences the spliceosome assembly. In wild-type extracts, pre-spliceosomal H- and A- complexes were fastly converted into completely assembled spliceosome (B, B^{act}, and C). In contrast, in the *hub1Δ* extract, splicing complex formation with the suboptimal substrate was inhibited at pre-spliceosomal H-complex stage. In case of optimal splicing substrates Hub1 was specifically present in B-complex; whereas in case of suboptimal substrates Hub1 was present in E-, A-, and B- complexes (Karaduman et al., 2017). Spliceosomal purifications using Snu66 from wild-type and *hub1Δ* cells did not show any major changes in the components of proteins in the yeast spliceosomal complex B. However, certain proteins of U1 and U2 snRNPs were enriched in *hub1Δ* cells (Mishra et al., 2011). Spliceosomal proteins play critical role in the recognition and pairing of splice sites, and aid in the dynamics of the RNA-RNA, RNA-protein, and protein-protein interaction networks of the spliceosome. These rearrangements ensure proper position of pre-mRNA for catalysis. The composition of the spliceosome is highly dynamic with interchange of proteins during splicing reactions. In yeast around 35 proteins dissociate and 12 other proteins are recruited during the transition from the pre catalytic B complex to the activated B^{act} complex. The dissociated proteins include all U1 and U4/U6 associated proteins. During the transition from B^{act} to C complex, the composition of the spliceosome alters to less

extent. In yeast only 2 proteins are disassociated and nine proteins are recruited to the spliceosome (Will and Luhrmann, 2011). Similarly, from the spliceosomal purifications we show that Hub1 modulates the composition of the spliceosome. In contrast to the canonical regulators of alternative splicing (e.g. SR proteins) that directly bind to *cis*-regulatory elements on pre-mRNAs, Hub1 modulates the spliceosome rather than binding to RNA substrates. The Mug161 (post-mRNA release spliceosomal complex) level was diminished in the spliceosome in *hub1-1* cells as compared to wild-type cells. As *mug161* transcripts depend on Hub1 for splicing, overall the level of protein is lowered, thereby lesser amount of Mug161 is recruited to the spliceosome in *hub1-1* cells. In our study, we observed that in *hub1-1* cells, few components of the spliceosome such as Prp11 (U2 snRNP component associated with A-complex), Prp24 (RNA-binding protein, a U6 associated component), Prp1(U5 snRNP component part of B-complex), and Cwf16 (recruited to the B^{act} complex) show enhanced recruitment to the spliceosome. Prp1 (ortholog of Prp6 in *S. cerevisiae*) is part of B complex but not part of B^{act} complex, is released during the activation of the spliceosome. Prp11 (ortholog of Prp5 in *S. cerevisiae*), is required for prespliceosome formation, facilitates stable association between U2 snRNP to the branch site. Prp11 is present in B complex but not in B^{act} complex, is released during the activation of the spliceosome similar to Prp1. Cwf16 (ortholog of Yju2 in *S. cerevisiae*) is recruited to the B^{act} complex and is shown to promote first splicing step. Transition from the B complex to the B^{act} complex involves unwinding of 24 base pairs between U4 and U6 snRNAs. Prp24 is essential for reannealing these snRNAs. The process of unwinding and renenealing of snRNAs are essential for new rounds of splicing (Fabrizio et al., 2009). Hub1 promotes eviction of splicing factors such as Prp11, Prp24, Prp1, and Cwf16 from the spliceosome. Hub1 might dynamically associate with the components of spliceosome at different stages of splicing cycle to regulate splicing. In the absence of Hub1, it might be possible that splicesome stalls from proceeding further and thereby certain splicing factors accumulate in the spliceosome. This might also be possible because Hub1 binds to the spliceosome for dynamic rearrangements of the spliceosome, thereby helping in release of few factors for stable spliceosomal assembly. However, further

experiments need to be carried out to establish definite role of Hub1 in releasing these splicing factors from the spliceosome to modulate splicing.

Hub1 uses a novel R941 surface III for its splicing function

A recent report on the molecular mechanism of Hub1 shows that Hub1 binds to Prp5 in *S. cerevisiae*, which in turn helps in assembly of the spliceosomes on pre-mRNA with non-canonical 5' SS (Karaduman et al., 2017). In our study, we show that *hub1* mutants defective in binding to Snu66 showed only mild splicing defects for the selected targets and growth defects. Whereas, *hub1* mutants defective in binding to Prp5 did not show any splicing defects for the selected targets and growth defects. This partial growth and splicing defects might be due to: (i) Hub1 interaction with Snu66 and Prp5 might affect only a subset of Hub1 dependent targets for their splicing. Deciphering such subset of targets will help us in better understanding the mode of Hub1 functioning in the cell. (ii) It is possible that Snu66 and Prp5 interaction with Hub1 might not be completely hindered in *hub1* mutants. (iii) Additionally, it is possible that Snu66 and Prp5 might interact through other surfaces of Hub1 in the spliceosome. These observations strengthen the hypothesis that in intron-rich *S. pombe* Hub1 might have additional contacts with other proteins for pre-mRNA splicing. *hub1* mutants with charged extensions at C-terminus failed to complement *S. pombe hub1Δ* lethality, although it could efficiently recruit Snu66-HIND protein (Mishra et al., 2011). We showed that *hub1-GG* and *hub1-DD* mutants bound to spliceosomal protein Snu66, but *hub1-DD* mutant failed to complement growth and splicing defects in *hub1-1* cells. From the 3D NMR analysis, we speculated that in *hub1-DD* mutant, DD extensions might form salt-bridge with R9R41 residues on Hub1. Hub1-R9 resides in linking region between sheets β1β2, Hub1-R41 positioned in linking region between helix α1 and sheet β3 and C-terminus of Hub1 is positioned between R9 and R41 residues. Therefore, we speculated that R9 and R41, along with C-terminus of the Hub1, might contribute to a new crucial surface III. The identified *hub1-R9A R41A* surface III mutant shows growth and splicing defects at 37°C . It is also reported that human Hub1 makes additional contacts with SR-protein kinases Cdc2/Cdc28-like kinases (Kantham et al., 2003). It is likely that Hub1 makes additional contacts with other splicing proteins

through the newly identified R941 surface III, as hindering these residues led to non-functional Hub1 at 37°C .

Snu66 and Rpb10 (RNA polymerase subunit) were identified as suppressors of *hub1-1* temperature sensitivity (Yashiroda and Tanaka, 2004). As the *hub1-1* allele is a structural mutant, its rescue could also occur due to stabilization of the protein by above proteins. Our new *hub1-R9A R41A* mutation did not alter its protein structure as the protein level remains unaltered. We further identified SPCPB16A4.06C (a *Schizosaccharomyces* specific gene) as a high copy suppressor of *hub1-R9A R41A* mutant temperature sensitivity. At this juncture, further experiments need to be carried out to establish role of *Schizosaccharomyces* specific gene in suppressing *hub1-R9A R41A* temperature sensitivity. It is possible that Hub1 contains multiple surfaces to facilitate binding with certain other proteins. Such interactions act like regulatory switches to dynamically modulate complex spliceosome machinery. Identifying these interactors would decipher the molecular mechanism of Hub1 in pre-mRNA splicing.

Hub1 promotes splicing of the transcripts that are synthesized faster

In mouse, embryonic stem cells (ESCs) the rate of RNA polymerase II (RNAPII) elongation has a crucial role in the regulation of alternative splicing (AS) (Maslon et al., 2019). Slow elongation rate affects gene expression and AS, consistent with the coupling of transcription with splicing. It is possible that the slow elongation rate cannot sustain the elevated levels of mRNA production at early stages of development. Optimal transcriptional elongation rates are required for proper gene expression and to regulate AS during development (Maslon et al., 2019). In yeast to humans, the pre-mRNA splicing occurs mostly through co-transcriptional splicing, in which transcription and splicing are coupled. Spt5 a membrane of core transcription elongation machinery in *S. cerevisiae* is required for pre-mRNA splicing. Spt5 promotes co-transcriptional splicing by strengthening the association between U5 snRNPs with the spliceosome complexes as they assemble on the nascent transcript (Maudlin et al., 2019). It has been shown that some of the core components of transcription elongation complex associate with the splicing factors (Moore et al.,

2006; Li et al., 2016). In our study, we showed by using bio-informatic analysis that Hub1 dependent targets showed positive correlation with respect to their synthesis time and splicing defects. We further showed that Hub1 promotes the splicing of the transcripts that are synthesized faster. It is possible that transcripts which are synthesized faster depend on factors like Hub1 to couple transcription with splicing. In the absence of factors like Hub1 faster-synthesized transcripts might undergo post-transcriptional splicing, thereby leading to missplicing. Also, it might be that factors like Hub1 may help in assembly of the spliceosome in the introns which are difficult for recognition by the spliceosome. We cannot rule out the possibility that Hub1 might also associate with transcription machinery in bridging transcription with splicing. It will be exciting to uncover how Hub1 recognizes the transcripts that are synthesized faster.

3.4 Conclusion

This is a first study of the conserved role of Hub1 in pre-mRNA splicing in intron-rich organism *S. pombe* at a genome-wide level. This splicing regulator selectively alters the protein composition of the spliceosome and preferentially promotes splicing of pre-mRNAs that are synthesized faster. Unlike its surface I and II, which have milder role in splicing in *S. pombe*, a novel surface III was discovered to be critical for its splicing function in *S. pombe*. This surface centred at two positively charged residues is functionally conserved in eukaryotes. It is likely that evolutionary pressure conserved the structure of Hub1 with its surfaces which forms multiple networks to fine tune gene expression.

Chapter 4

The Kreb's cycle enzyme fumarase regulates pre-mRNA splicing through the ubiquitin-like protein Hub1

Abstract

The ubiquitin-like protein Hub1 functions in pre-mRNA splicing by binding non-covalently to the spliceosomal proteins Snu66 and Prp5. We have found that *Schizosaccharomyces pombe* Hub1 also binds to a mitochondrial enzyme of the citric acid cycle, fumarase (Fum1). The enzyme binds to a conserved surface of Hub1 centred at a solvent-exposed tryptophan residue. This surface is absent in *S. cerevisiae* Hub1; however, an introduction of tryptophan at analogous position restores its affinity with Fum1. Hub1-Fum1 complex precipitates *in vitro* indicating mutually inhibitory activities of the two partners. In support of potential inhibitory activities, elevated levels of both Fum1 and *fum1ΔSS* (cytosolic fumarase) are more toxic in *hub1-W47G* mutant cells compared to the wild-type *S. pombe* cells. *hub1-W47G* mutant cells show genetic interaction with splicing factor mutants and also exhibit inefficient excision of introns from selected pre-mRNAs. *fum1Δ* mutant also shows genetic interaction with certain splicing factor mutants. Higher levels of *fum1ΔSS* is more toxic in splicing factor mutants, compared to the wild-type cells. Thus, Fum1 protein which is not imported to mitochondria could regulate pre-mRNA splicing through Hub1. Since fumarase is frequently mutated in multiple diseases in humans, phenotypes of some of the mutants may be due to their influence on Hub1-dependent pre-mRNA splicing.

4.1 Introduction

Introduction on Hub1 (detailed in section 1.8). In *S. cerevisiae,* Hub1-HIND complex promotes alternative splicing of *SRC1/HEH1* (Mishra et al., 2011). High levels of Hub1 in association with Prp5 promote excision of introns with suboptimal 5' SSs, 3' SSs, branch point sequences and also combinations of these sequences

(Karaduman et al., 2017). The excessive use of suboptimal SSs by Hub1 modified spliceosomes makes splicing process error-prone (Karaduman et al., 2017; Chanarat and Mishra, 2018). Hub1 activity, therefore, should be kept under check to avoid unwanted splicing or missplicing. Overexpression of Hub1 activates missplicing across several cryptic introns in *S. cerevisiae*, including one in the upstream sequence of *PRP5.* This cryptic intron is cut to reduce levels of functional Prp5 upon Hub1 overexpression (Karaduman et al., 2017; Chanarat and Mishra, 2018). In other eukaryotes, missplicing could be minimized by similar Hub1-dependent negative feedback control of Prp5, or inhibitory molecules could control Hub1 activity by direct binding. Hub1 is also reported to form SDS-resistant adducts with many unknown proteins suggesting its role in multiple processes in the cell (Luders et al., 2003).

In a yeast two-hybrid screen, we found fumarase as an interactor of Hub1 in *S. pombe*. The enzyme fumarase (classII, fumarate hydratase in higher eukaryotes) is a component of the tricarboxylic acid cycle (TCA), which is conserved across organisms from bacteria to humans with respect to its sequence, structure, and enzymatic activity (Singer et al., 2017). The fumarase is targeted between two compartments of the eukaryotic cell. In mitochondria, the enzyme catalyzes the reversible conversion of fumaric acid to L-malic acid and in the cytosol/nucleus as part of the DNA damage response (DDR) (Yogev et al., 2010). The bacterial fumarase (*Bacillus subtilis*, Fum-bc) is induced upon DNA damage and participates in DDR (Singer et al., 2017). In *S. cerevisiae,* cytosolic fumarase physically interacts with Sae2, which is involved in meiotic and mitotic double-strand breaks (Leshets et al., 2018; Baroni et al., 2004). In cytosolic fumarase depleted cells, protein level of Sae2 is reduced, possibly because fumarase regulates its abundance (Leshets et al., 2018). Human fumarase functions as a tumor suppressor and mutations in fumarase gene are associated with hereditary leiomyomatosis and renal cell cancer (HLRCC) syndrome. Many studies indicate that in the absence of fumarase enzyme, increased fumarate concentration inhibits prolyl hydroxylase enzymes (PHD 1, 2, and 3). Thereby fumarate activates hypoxia-inducible transcription factor (HIF) to promote angiogenesis and glucose metabolism which are essential for

tumorigenesis (Isaacs et al., 2005; Pollard et al., 2005; Selak et al., 2005; Hanahan and Weinberg, 2011). Although, studies in the past showed physical interaction of cytosolic fumarase to Sae2 involved in DDR, but other functions of cytosolic fumarase remained obscure. Here we report that Hub1 interacts with Fum1 through conserved solvent-exposed tryptophan (W47) surface. The Hub1-Fum1 complex precipitates *in vitro*. Enhanced cytosolic fumarase is inhibitory for the cell growth and more toxic in *hub1-W47G* mutant cells as compared to the wild-type *S. pombe* cells. *hub1-W47G* mutant cells show genetic interaction with splicing factor mutants and display inefficient excision of introns from a selected subset of pre-mRNAs. *fum1Δ* mutant also shows genetic interaction with splicing factor mutants. The elevated levels of *fum1ΔSS* is more toxic in splicing factor mutants compared to the wild-type cells. Thus, cytosolic Fum1 possibly regulates pre-mRNA splicing through Hub1.

4.2 Results

4.2.1 *S. pombe* Hub1 and Fumarase interaction

To study other functions of *S. pombe* Hub1, my thesis supervisor Shravan Kumar Mishra carried out an unbiased approach to identify novel interacting partners of Hub1. In this screen Hub1 BD-fusion was used as a bait, and *S. pombe* cDNA libraries were expressed as AD-fusion for screening. Six potential candidate clones were identified from the screen. Positive clones were revalidated its growth on (-His) plates after FOA shuffle (Fig 4.1A). Further, DNA sequencing of six independent positive AD-constructs revealed the identity of Hub1's Y2H interactor as fumarase. In order to identify how the interaction is mediated, protein sequence alignment of Hub1 orthologs was compared (Fig 4.1B). From Hub1 alignment, we identified a unique solvent-exposed tryptophan at 47th position containing hydrophobic patch which is replaced by glycine residue in *S. cerevisiae* (Fig 4.1B). To understand the mode of Hub1-Fum1 interaction, Y2H assay was performed using identified fumarase candidate with *S. cerevisiae* HUB1*, S. cerevisiae hub1-G47W* mutant*, S. pombe* HUB1, and *S. pombe hub1-W47G* mutant. Hub1 clones with W at 47th position was able to interact with Fum1, whereas Hub1 clones with G at 47th position was unable to interact (Fig 4.1C). This finding indicates that Hub1-Fum1 interaction is mediated through the tryptophan-47 surface.

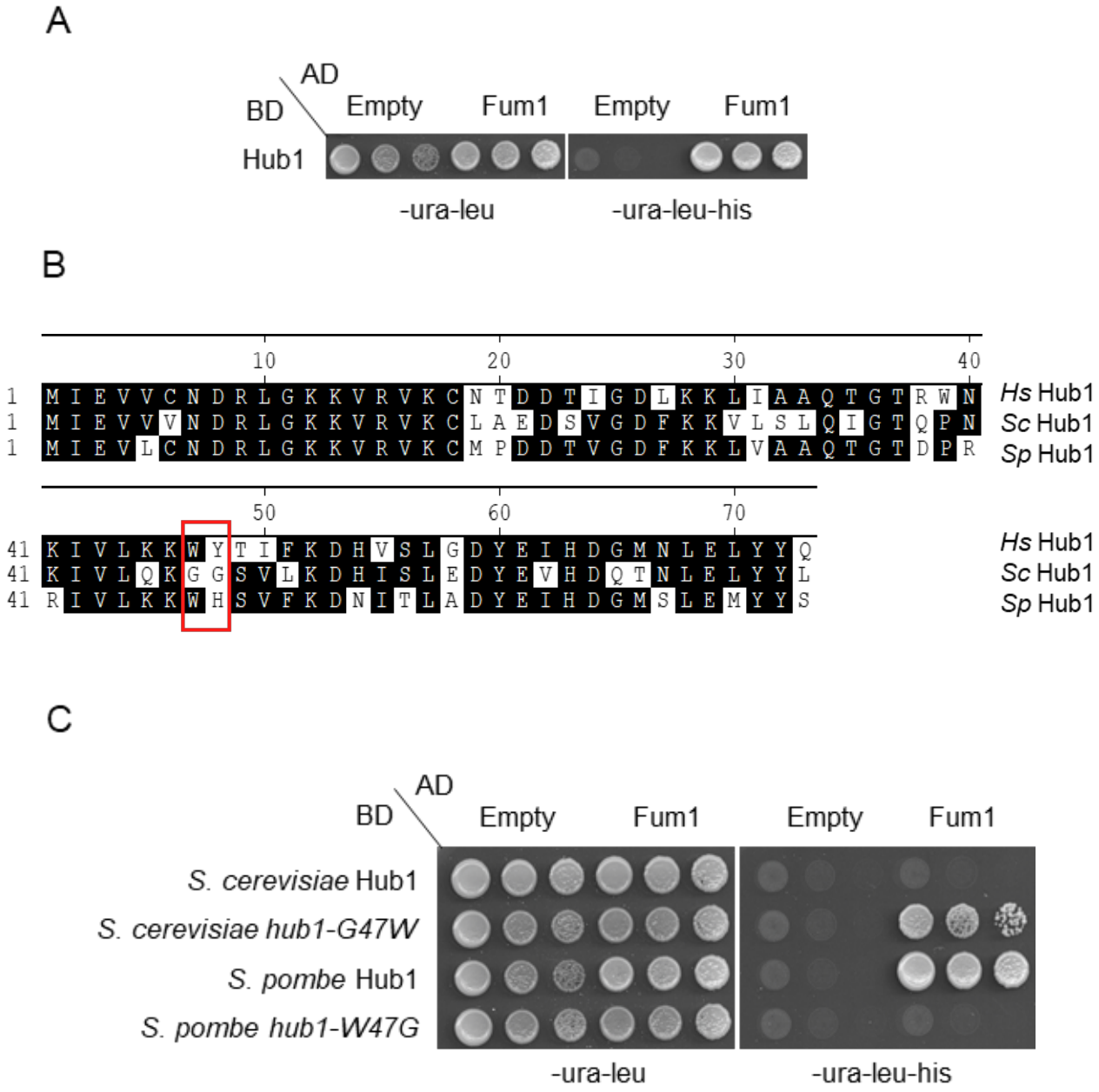


Figure 4.1: Hub1-Fum1 interaction by yeast two-hybrid assay

A. Confirmation of Hub1-Fum1 interaction. In a yeast two-hybrid assay pGBDU1-Hub1 and pGADC1-Fum1 constructs (identified in the cDNA library screen) were co-transformed, transformants were five-fold serial diluted on control and selective plates. The experiment is similar to Figure 3.6. (Note: Figure 4.1A cropped from Figure 4.1C).

B. Protein sequence alignment of Hub1 orthologs. Identical amino acids are shaded black and fumarase binding tryptophan surface site highlighted with red color box.
C. Yeast two-hybrid interaction studies of identified Fum1 with various HUB1 orthologs. BD-fusions of Hub1 and *hub1* mutants from *S. cerevisiae and S. pombe* were co-transformed with AD-Fum1. The transformants were five-fold serial diluted on control and selective plates. The experiment is similar to Figure 3.6 (Data is from Shravan Kumar Mishra).

4.2.2 Hub1-Fum1 *in vitro* interaction

To check if Hub1 binds Fum1 directly, Shravan Kumar Mishra performed *in vitro* interaction assay. Equimolar concentrations of purified Hub1-W47 and Hub1-G47 were mixed with Fum1 separately and incubated on ice for 30 minutes. The protein mixture was separated into soluble and pellet fractions by centrifuging at 15000xg. A fraction of Hub1-W47-Fum1 mixture co-precipitated and thus both the proteins were seen also in the pellet fraction (Fig 4.2A). In contrast, Hub1-G47-Fum1 mixture did not co-precipitate and thus Hub1 was completely in the soluble fraction and lesser amount of Fum1 was seen in the pellet (Fig 4.2A). Addition of HIND elements (Hub1 INteracting Domain) to the Hub1-Fum1 mixture possibly titrated Hub1 from complexing with Fum1, thereby preventing co-precipitation and fractionation of both the proteins into pellet (Fig 4.2A). We further added a disaccharide trehalose to the Hub1-Fum1 protein mixture and subjected to fractionation. Similar to the effect after HIND addition, trehalose inhibited Hub1-Fum1 co-precipitation (Fig 4.2B). From these findings, we infer that Hub1-Fum1 binds directly and co-precipitates through the conserved tryptophan-47 surface.

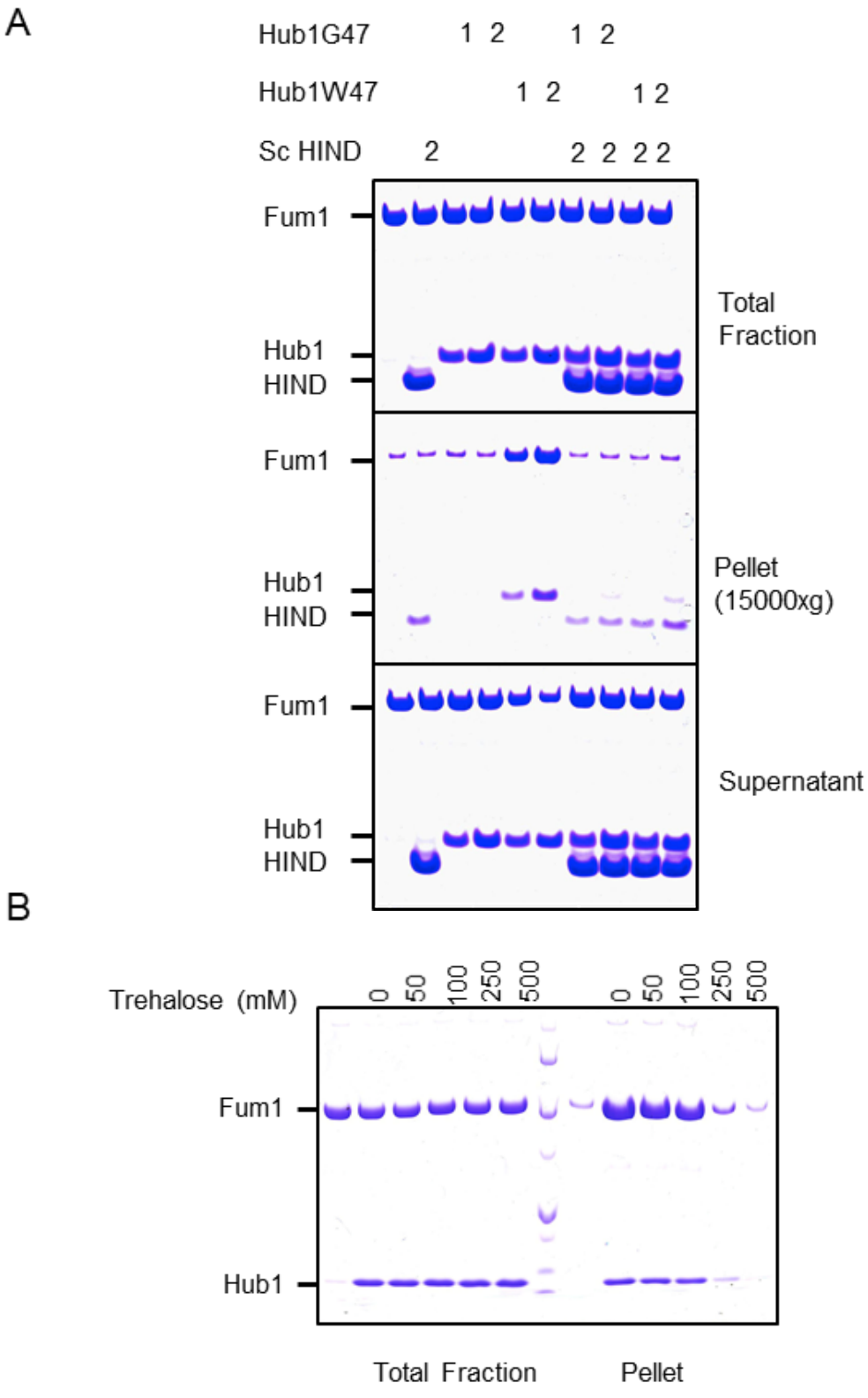


Figure 4.2: Hub1-Fum1 *in vitro* interaction

A. Hub1 interacts with Fum1 *in vitro.* Equimolar concentration of bacterially purified *S. pombe* Hub1-W47 and *S. cerevisiae* Hub1-G47 were mixed with *S. pombe* Fum1 separately. The mixture was incubated on ice, then separated into soluble and pellet

fractions and samples were loaded on SDS-PAGE. After electrophoresis, the proteins were detected by staining the gels with coomassie brilliant blue. Hub1-Fum1 interaction co-precipitates and addition of HIND element inhibits Hub1-Fum1 co-precipitation (Data is from Shravan Kumar Mishra). The proteins were purified according to standard His-tag affinity purification as described in the methods section (2.2.17). 1, 2 represent 1X, 2X molar concentrations of the proteins.

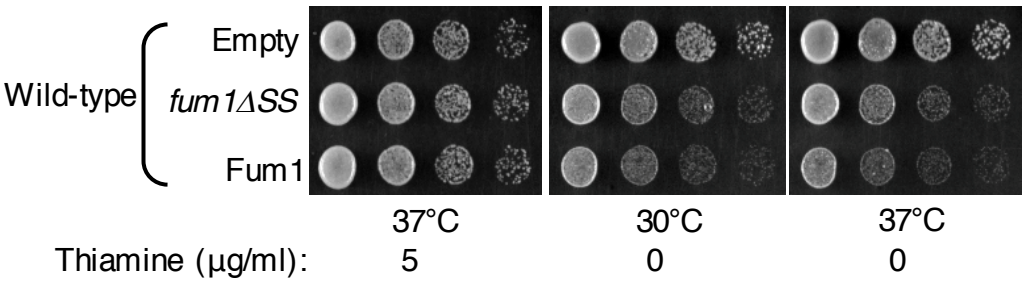
B. Trehalose inhibits Hub1-Fum1 interaction, the experiment was as in (A) (Data is from Shravan Kumar Mishra). The central lane represents molecular weight markers of the protein. The first lane in the total fraction and the first lane in the pellet fraction represent Fum1 protein alone.

4.2.3 Fumarase overexpression inhibits *S. pombe* growth

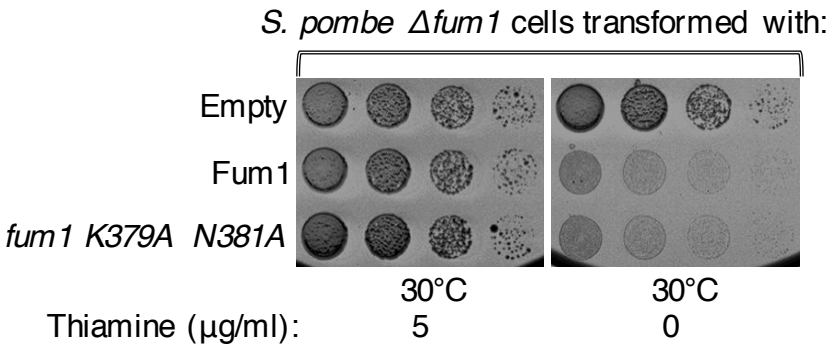
Hub1 shows dual localization both to the nucleus and cytosol, whereas Fum1 under normal conditions it is localized to mitochondria and cytosol. Fum1 is also translocated from the cytosol to the nucleus upon DNA damage induction. In Yeast and human cells, Fum1 levels were increased in response to DNA damaging agents (Yogev et al., 2010). To understand the physiological relevance of Hub1-Fum1 interaction, we overexpressed both Fum1 and *fum1ΔSS* (lacking the N-terminal mitochondrial signal sequence (MTS)) in wild-type *S. pombe* cells. Higher levels of both Fum1 and *fum1ΔSS* inhibited cell growth (Figure 4.3A). From this we conjecture that it is the cytosolic form that is causing growth defects. To understand whether Fum1 toxicity is associated with its enzymatic activity, we overexpressed both Fum1 and enzyme inactive fumarase mutant (*fum1 K379A, N381A*) in *S. pombe Δfum1* cells. As before, both wild-type fumarase and fumarase mutant lacking enzymatic activity showed similar toxicity (Figure 4.3B). To assess the enzymatic activity, Fum1 and *fum1 K379A, N381A* transformed cells were lysed and cell lysates were subjected to fumarase activity assay. In this assay conversion of malate to fumarate leads to an increase in the absorbance at 240_{nm}. The cells transformed with Fum1 showed increase in absorbance with time at 240_{nm}, whereas *fum1 K379A, N381A* transformed cells did not show any increase in absorbance (Figure 4.3C). Thus, our findings suggest that higher levels of cytosolic Fum1 inhibited cell growth and the associated toxicity is independent of its enzymatic activity. To understand the role of Hub1 in Fum1-induced toxicity, I induced higher levels of Fum1 in wild-type and *hub1-W47G* mutant cells. The high levels of fumarase led to more toxicity in *hub1-W47G* mutant cells compared to wild-type cells

(Figure 4.3D). It is possible that in *hub1-W47G* mutant cells, the Hub1 is unable to titrate the Fum1 aggregates and thereby it does not lower the toxicity. It cannot be ruled out that cellular toxicity is additive: because of independent effects of *hub1-W47G* mutant and higher level of cytosolic Fum1.

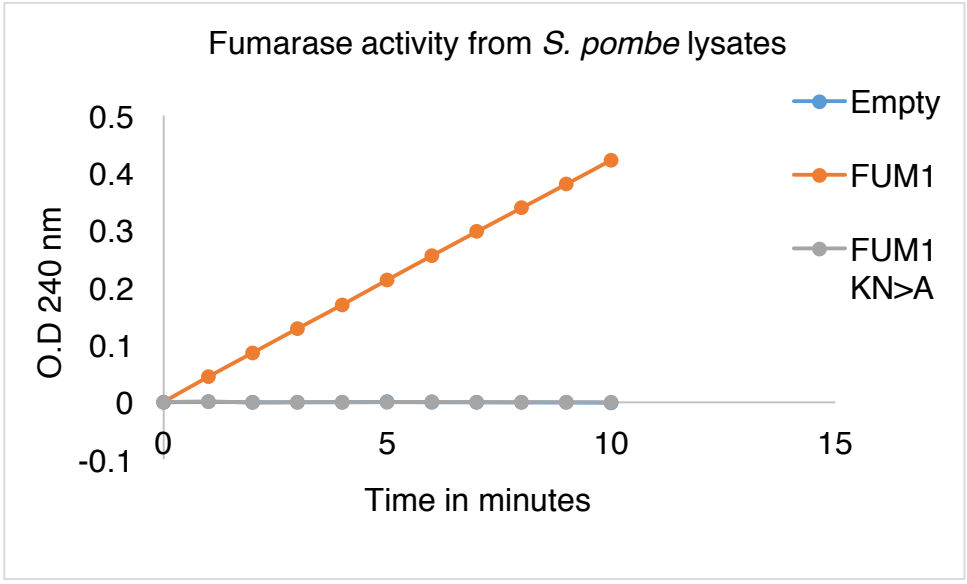
A



B



C



D

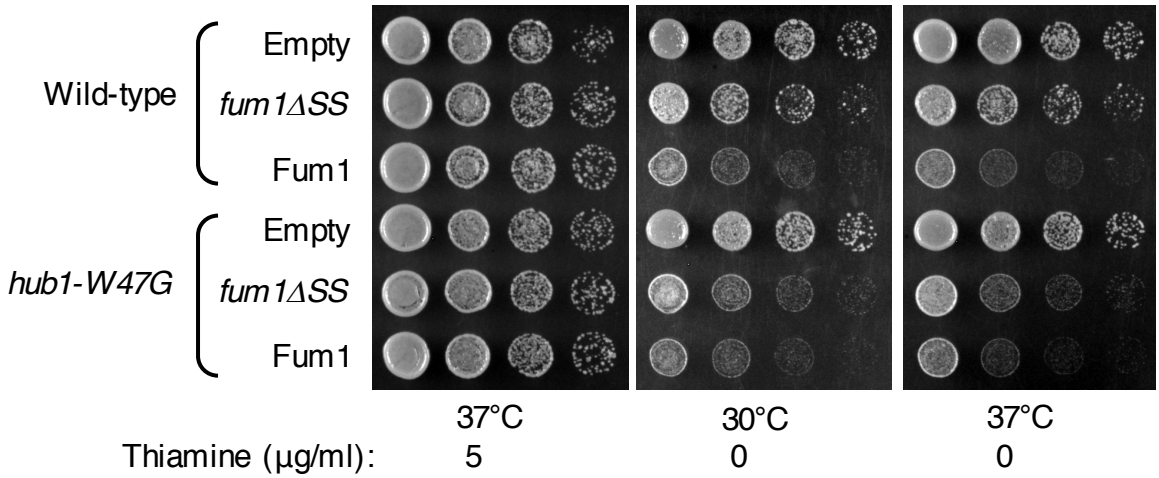


Figure 4.3: Higher levels of Fum1 inhibits cell growth

A. *S. pombe* wild-type cells transformed with Fum1-6HA, *fum1ΔSS-6HA* constructs were expressed from the strong thiamine-repressible nmt3 promoter. Five-fold serial dilution spotting was done on indicated agar plates. Plates were incubated at 30°C and 37°C for 3-4 days.

B. *S. pombe Δfum1* cells transformed with Fum1-6HA, *fum1K379A N381A-6HA* constructs were expressed from thiamine repressible nmt3 promoter. The experiment is similar to Figure 4.3A (Data is from Shravan Kumar Mishra).

C. The experiment was as in (B), the cell lysates were subjected to fumarase activity assay (Data is from Shravan Kumar Mishra).

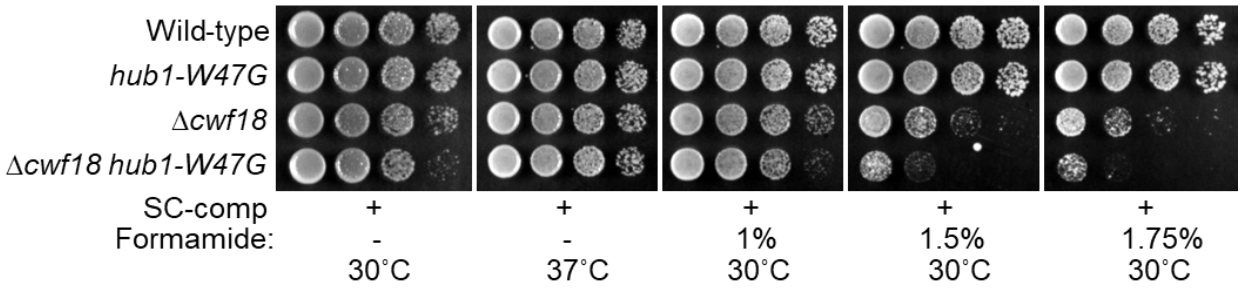
D. *S. pombe* wild-type (SP1), *hub1-W47G* cells (P50) transformed with Fum1-6HA, *fum1ΔSS-6HA* constructs were expressed from thiamine repressible nmt3 promoter. The experiment is similar to Figure 4.3A.

4.2.4 *hub1-W47* surface is required for pre-mRNA splicing

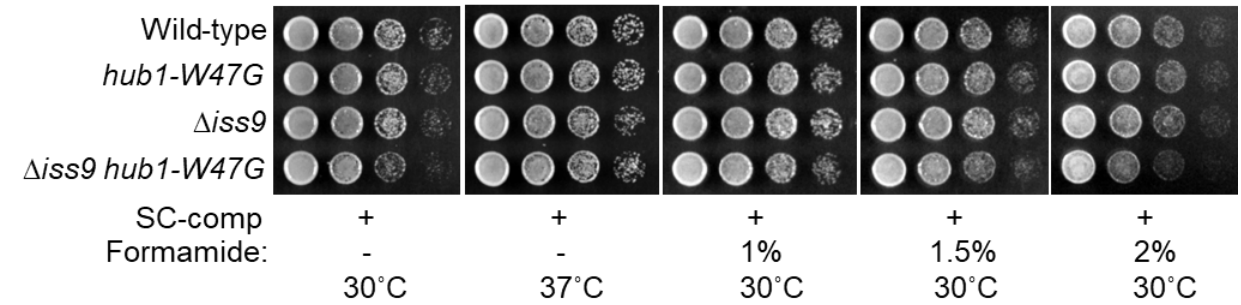
Hub1 is required for pre-mRNA splicing (Wilkinson et al., 2004; Mishra et al., 2011; Ammon et al., 2014). It contains a unique solvent-exposed tryptophan surface at 47th position, which is required for its interaction with fumarase and to promote association with cellular proteins. To understand the function of Hub1-W47 surface, we performed genetic interaction assays with splicing factor mutants. This assay was performed by a colleague Nivedha Balaji. We also treated *S. pombe* cells with formamide, as it is known to reduce splicing efficiency (Manchado et al., 2017). For this assay, we generated double mutant strain of *hub1-W47G* and *Δcwf18*

(complexed with Cdc5 protein) in *S. pombe. Δcwf18 hub1-W47G* double mutant cells show slight growth defect compared to single mutants at 30°C (Figure 4.4A). *Δcwf18 hub1-W47G* double mutant cells treated with formamide showed severe growth defect compared to single mutants indicating negative genetic interaction (Figure 4.4A). Similarly, we generated double mutant strain of *hub1-W47G* and *Δiss9* (iss9 is predicted to function as intron-specific pre-mRNA splicing regulator). *Δiss9 hub1-W47G* double mutant also showed moderate growth defect compared to single mutants at 30°C and in the presence of formamide, indicating negative genetic interaction (Figure 4.4B). As these results provided evidence for splicing function of Hub1-W47 surface, we attempted to understand its splicing function. The total RNA was isolated from logarithmically growing wild-type and *hub1-W47G* cells. RNAs were converted into cDNAs using random hexamer primers and reverse transcriptase. Following cDNA synthesis, we carried out RT-PCR assays for the selected targets, which showed an accumulation of pre-mRNA in microarrays. The amplified products were visualized on agarose gel electrophoresis. The selected targets showed accumulation of intron-containing transcripts in *hub1-W47G* cells at 37°C (Figure 4.4C), indicating that Hub1-W47 surface promotes pre-mRNA splicing.

A


Wild-type
hub1-W47G
Δcwf18
Δcwf18 hub1-W47G
SC-comp + + + + +
Formamide: - - 1% 1.5% 1.75%
30˚C 37˚C 30˚C 30˚C 30˚C

B


Wild-type
hub1-W47G
Δiss9
Δiss9 hub1-W47G
SC-comp + + + + +
Formamide: - - 1% 1.5% 2%
30˚C 37˚C 30˚C 30˚C 30˚C

C

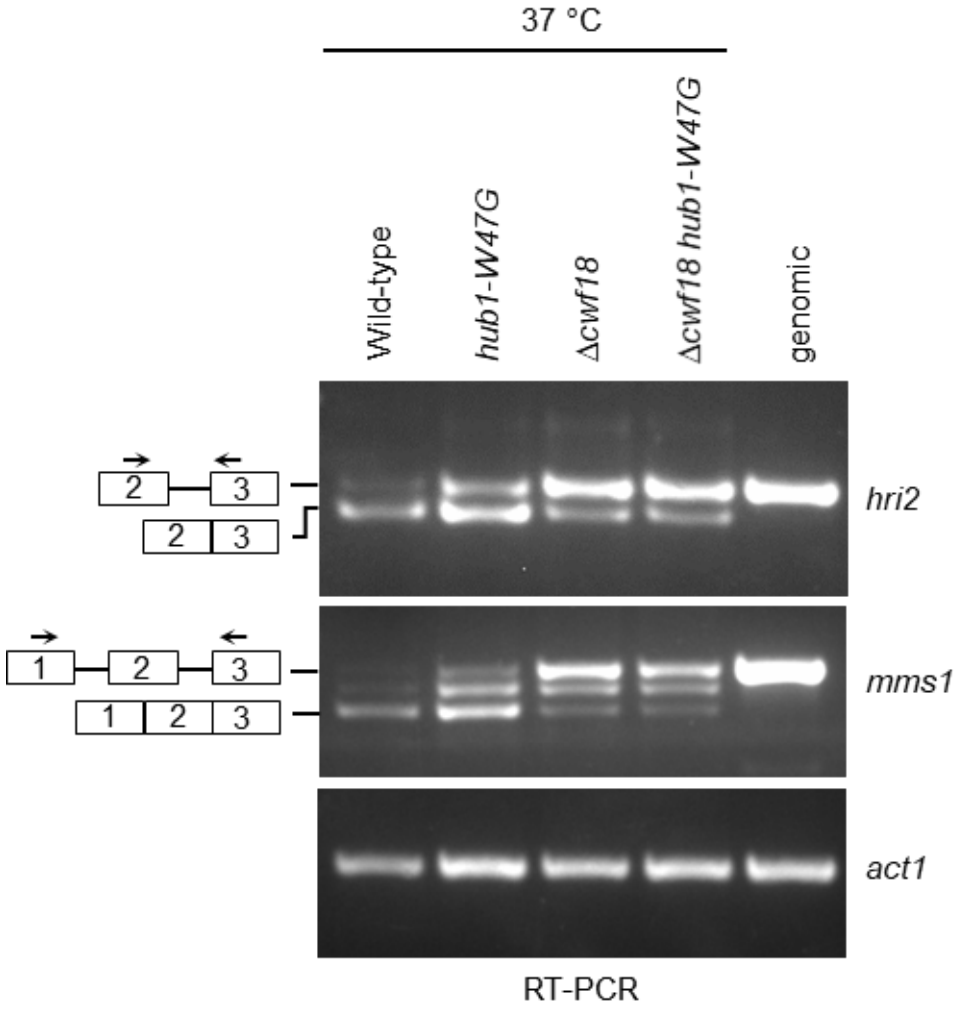


Figure 4.4: Hub1-W47 surface promotes RNA splicing

A. *Δcwf18 hub1-W47G* (PN4) double mutant shows synthetic growth defect in the presence of 1.5% and 2% formamide compared to *Δcwf18* strain (PN3). The experiment is similar to Figure 4.3A (the experiment was peformed by Nivedha Balaji).

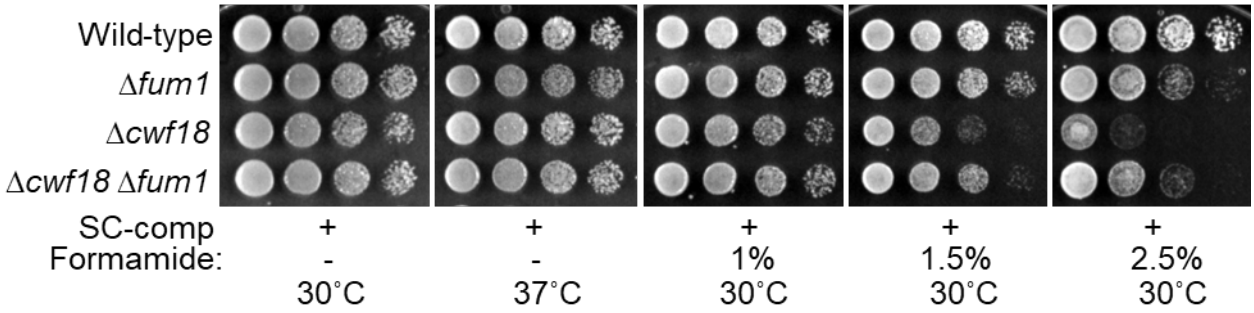
B. *Δiss9 hub1-W47G* (PN8) double mutant shows synthetic growth defect in the presence of 2% formamide compared to *Δiss9* strain (PN7). The experiment is similar to Figure 4.3A (the experiment was peformed by Nivedha Balaji).

C. Semi-quantitative RT-PCR shows accumulation of intron-containing transcripts for *hri2 and mms1.* cDNA was prepared from total RNA isolated from *S. pombe* wild-type (PN1), *hub1-W47G* (PN2), *Δcwf18* (PN3) and *Δcwf18 hub1-W47G* (PN4) strains. The experiment is similar to Figure 3.1B.

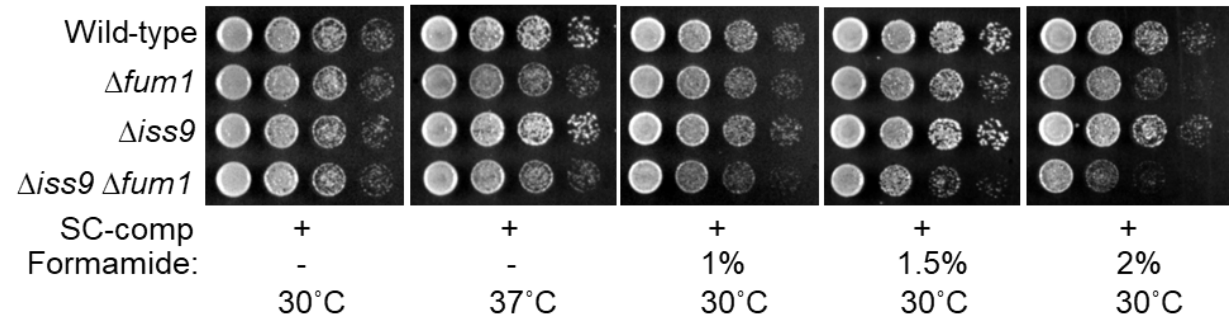
4.2.5 *S. pombe Δfum1* mutant shows genetic interaction with splicing factors

Overexpression of Hub1 is known to activate missplicing across several cryptic introns in *S. cerevisiae*, including one in the upstream sequence of *PRP5.* This cryptic intron is cut to reduce levels of functional Prp5 upon Hub1 overexpression (Karaduman et al., 2017). In other eukaryotes, missplicing could be minimized either by similar Hub1-dependent negative feedback controls of Prp5, or by inhibitory molecules could control Hub1 activity by direct binding. We speculated that fumarase might regulate splicing function through Hub1 as the *hub1-W47G* mutant showed genetic interactions with splicing factors. To understand the role of Fum1 in splicing, we generated *Δcwf18 Δfum1* and *Δiss9 Δfum1* double mutants to monitor the genetic interaction. This assay was performed by colleague Nivedha Balaji. In presence of formamide, *Δcwf18 Δfum1* double mutant cells showed improved growth compared to *Δcwf18 single* mutant indicating positive genetic interaction (Figure 4.5A). In presence of formamide *Δiss9 Δfum1* double mutant cells showed growth defect compared to single mutants indicating negative genetic interaction (Figure 4.5B). To further understand the role of Fum1 in regulating splicing function, we induced high levels of *fum1ΔSS* with pREP4X-nmt based promoter in various splicing factor mutants. Higher levels of *fum1ΔSS* led to more toxicity compared to wild-type cells (Figure 4.5C). Thus, our data suggest that Fum1 might regulate splicing function through Hub1.

A


Wild-type
Δfum1
Δcwf18
Δcwf18 Δfum1
SC-comp + + + + +
Formamide: - - 1% 1.5% 2.5%
30˚C 37˚C 30˚C 30˚C 30˚C

B


Wild-type
Δfum1
Δiss9
Δiss9 Δfum1
SC-comp + + + + +
Formamide: - - 1% 1.5% 2%
30˚C 37˚C 30˚C 30˚C 30˚C

C

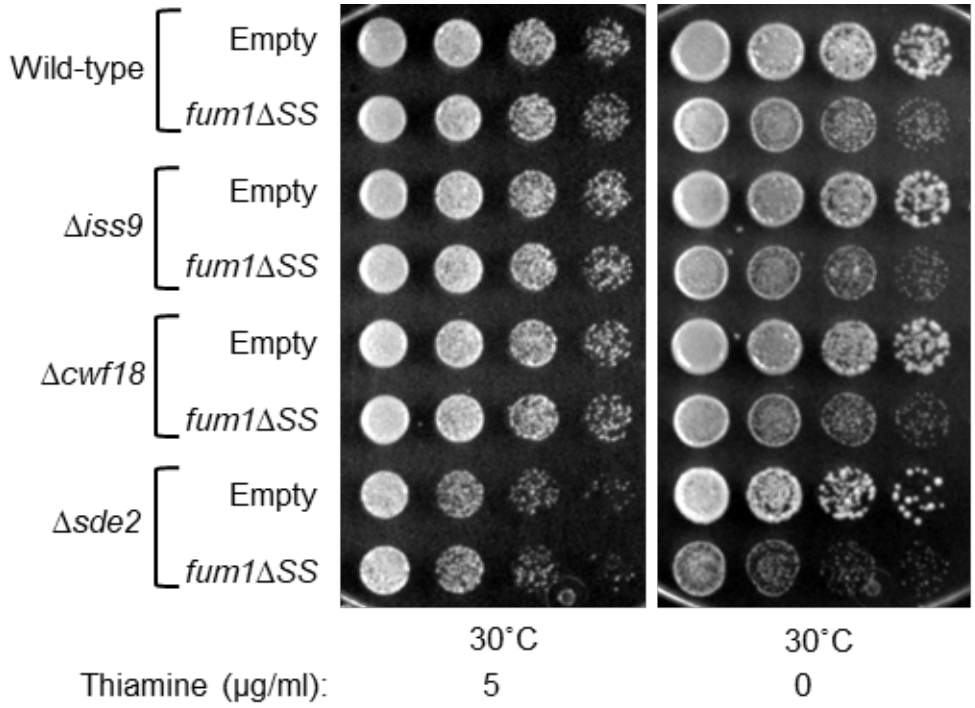

Wild-type
Empty
fum1ΔSS
Δiss9
Empty
fum1ΔSS
Δcwf18
Empty
fum1ΔSS
Δsde2
Empty
fum1ΔSS
30˚C 30˚C
Thiamine (µg/ml): 5 0

Figure 4.5: *Δfum1* mutant shows genetic interaction with splicing factor mutants

A. *Δcwf18 Δfum1* (PN50) double mutant grows better than *Δcwf18* (PN49) mutant cells in the presence of 1.5% and 2% formamide. The experiment is similar to Figure 4.3A (the experiment was peformed along with Nivedha Balaji).
B. *Δiss9 Δfum1* (PN26) double mutant shows synthetic growth defect in the presence of 1%, 1.5%, and 2% formamide compared to *Δiss9* strain (PN25). The experiment is similar to Figure 4.3A (the experiment was peformed along with Nivedha Balaji).
C. Higher levels of *fum1ΔSS* were more toxic in splicing factor mutant cells compared to the wild-type cells. The experiment is similar to Figure 4.3A (the experiment was peformed along with Nivedha Balaji).

4.3 Discussion

Hub1 function is mostly involved in pre-mRNA splicing. Previous work on the ubiquitin-like protein Hub1 was mainly focused on intron-poor organism, *S. cerevisiae* (Luders et al., 2003; Mishra et al., 2011). This is the first study in intron-rich *S. pombe*, addressing the role of mitochondrial enzyme, fumarase in regulating the splicing function through Hub1. We show that Hub1 binds to Fum1 and this interaction is mediated through its tryptophan-47 surface. The purified Hub1-Fum1 mixture co-precipitates *in vitro*. *hub1-W47G* mutant shows shows accumulation of intron-containing transcripts. *fum1Δ* mutant showed genetic interaction with certain splicing factor mutants. Overexpression of fumarase lacking mitochondrial targeting signal (*fum1ΔSS)* showed higher toxicity in mutants of various splicing factors compared to wild-type cells. Thus, our data suggest that Fum1 might regulate splicing function through Hub1.

Hub1-Fum1 interaction

Hub1 protein is conserved from yeast to humans. In *S. pombe,* Hub1 and Snu66 are essential, although the mutants deficient in Hub1-Snu66 interaction are viable. This finding, along with detection of significant cytosolic pools of Hub1 and the levels of Hub1 is more abundant in cells than snu66, indicating that Hub1 function may not be just restricted to splicing. The *S. cerevisiae Δhub1 Δsnu66* double mutant is temperature-sensitive but not the corresponding single mutant. This non-epistatic behavior suggests the separate functions and further points other functions of Hub1 in addition to splicing (Mishra et al., 2011). In *C. elegans*, Hub1 is shown to play a

role in mitochondrial unfolded protein response (UPR_{mt}). The present study also shows that Hub1 binds to mitochondrial enzyme Fum1 in *S. pombe*. Here HUB1 is essential for viability and required for pre-mRNA splicing, whereas FUM1 is not essential. Fum1 is implicated in various functions such as TCA metabolic cycle, DNA double-strand breaks, homologous recombination (HR) repair pathway and also functions as a tumor suppressor metabolite.

From protein sequence alignment of Hub1 orthologs, we observed a unique solvent-exposed tryptophan at 47^{th} position containing hydrophobic patch which is replaced by glycine residue in *S. cerevisiae*. Using Y2H assay, we show that Hub1 clones with W at 47^{th} position associates with Fum1, whereas Hub1 clones with G at 47^{th} position was unable to bind. Indicating that Hub1-Fum1 interaction is mediated through tryptophan surface.

Porcine fumarase (96% sequence homology to a human enzyme) leads to fibril aggregation and catalytic inactivation in presence of H_2O_2 and $OH^{\bullet}$, as it gets modified both in secondary and tertiary protein structure (Barteri et al., 2007). This clearly indicates that Fum1 has an intrinsic tendency to form aggregates when stimulated. Corroborating the above findings, we show that Hub1-Fum1 co-precipitates. Further HIND elements titrate Hub1 from complexing with Fum1 and prevent aggregation of Hub1-Fum1 complex. Trehalose, a naturally occurring disaccharide, is known to act as a chemical chaperone to stabilize protein structure, thereby decreases misfolding and protein aggregation (Singer et al., 1998). Trehalose also inhibits aggregation of Hub1-Fum1 complex. Hub1 protein is conserved from *S. cerevisiae* to humans. It possesses a unique solvent-exposed tryptophan at 47^{th} position containing hydrophobic patch which is replaced by diglycine residues in *S. cerevisiae*. Using yeast two-hybrid assays and *in vitro* interaction assays, we show that Hub1-W47 surface mediates its interaction with Fum1.

In *S. cerevisiae*, Hub1 is not essential for its viability (Dittmar et al., 2002). It moderately affects the splicing of a few transcripts only. In contrast, Hub1 is

essential for viability in *S. pombe* where it affects the splicing of several pre-mRNAs (Mishra et al., 2011). Hub1 might associate with multiple binding partners in intron-rich organisms to perform its splicing function. Other than binding with Fum1, Hub1-W47 surface may promote such associations to perform its splicing function.

Hub1-W47 surface promotes its interaction with Fum1

The fumarate hydration reaction is the seventh step of the tricarboxylic acid (TCA) cycle, and this biochemical pathway is responsible for synthesis of ATP in aerobic respiration (Mescam et al., 2011). Fumarase enzyme catalyzes the reversible conversion of fumaric acid to L-malic acid in mitochondria. Fum1 in the cytosol functions in DNA damage response (DDR), and is translocated from the cytosol to the nucleus upon DNA damage induction (Yogev et al., 2010). In *S. cerevisiae,* cytosolic fumarase physically interacts with Sae2, which is involved in meiotic and mitotic double-strand breaks (Leshets et al., 2018; Baroni et al., 2004). In cytosolic fumarase depleted cells, protein level of Sae2 is reduced, possibly fumarase regulates the abundance of Sae2 (Leshets et al., 2018). Human and yeast cells show increased level of fumarase (2-2.3 fold) after 24 hours of exposure to hydroxyurea (HU) indicating that fumarase levels being regulated by DNA damage. We show that in *S. pombe* cells, higher levels of Fum1 and *fum1ΔSS* (lacking mitochondrial signal sequence at the N-terminus) are toxic. Fumarase overexpression induced toxicity is not due its enzymatic activity, but might be due to presence of physical aggregates in the cell. The higher levels of Fum1, *fum1ΔSS* inhibits cell growth severely in *hub1-W47G* mutant cells than wild-type cells. In case of wild-type cells, it is possible that Hub1 co-aggregates with Fum1 and clears the toxic aggregates. In contrast, in *hub1-W47G* cells, Hub1 unable to co-aggregate with Fum1, thereby accumulates pronounced toxicity. Elevated toxicity in *hub1-W47G* cells might also be attributed collectively to independent functions of *hub1-W47G* mutant and high level of Fum1. The toxicity associated with Fum1 was more severe compared to cytosolic fraction of fumarase (*fum1ΔSS)*, which might be due to dual compartmentalization and different functions of Fum1.

Hub1-W47 surface promotes splicing

In yeast and human cell cultures, Hub1 promotes pre-mRNA splicing (Karaduman et al., 2017). Hub1 shows synthetic sickness with various splicing factor mutants (Mishra et al., 2017). In our study, both *Δcwf18 hub1-W47G* and *Δiss9 hub1-W47G* double mutants treated with formamide showed negative genetic interaction. This provides further evidence of Hub1-W47 surface functioning in pre-mRNA splicing. The *hub1-W47G* mutant showed an accumulation of intron-containing transcripts from a selected subset of pre-mRNAs. *hub1-W47G* mutant showing splicing defects might be due to its abolished interaction with Fum1. *fum1Δ* mutant showed genetic interaction with splicing factor mutants. In the presence of formamide, *Δcwf18 Δfum1* double mutant cells showed better growth compared to *Δcwf18* single mutant indicating positive genetic interaction. In this case, Hub1 protein in the absence of fumarase might improve the splicing, and thereby double mutant might show better growth. *Δiss9 Δfum1* double mutant cells in the presence of formamide showed growth defect compared to single mutants indicating negative genetic interaction. The higher levels of *fum1ΔSS* led to more toxicity in splicing mutants compared to wild-type cells. *hub1-1* cells were synthetically sick with deletion mutants of multiple splicing factors (Thakran et al., 2018). It is possible that elevated levels of cytosolic Fum1 show dominant-negative phenotype by titrating Hub1 levels in the cells thereby shows growth sickness with deletion mutants of multiple splicing factors. Fum1 may regulate splicing function through Hub1. Further experiments need to be carried out to establish Fum1 role in regulating splicing through Hub1. Many splicing factors contain intrinsically disordered regions (IDRs) and low amino acid complexity. These regions can undergo liquid-liquid phase separation, to form membrane-less organelles that can regulate alternative splicing (Zlotorynski, 2017). From our study, we speculate that Hub1 also binds with other splicing factors and undergo similar phase separation to regulate splicing. Hub1-W47 surface might promote association with Fum1 and various other factors, which further helps to regulate its splicing function.

4.4 Conclusion

We demonstrated that the Kreb's cycle enzyme, fumarase regulates pre-mRNA splicing through ubiquitin-like protein, Hub1. This is the first study to show that Hub1 directly binds to Fum1 through its W47 surface. The purified Hub1-Fum1 mixture co-precipitates *in vitro*. Hub1-W47 surface might promote association with various other cellular proteins, however, the functional relevance needs to be addressed. Hub1-W47 surface promotes efficient excision of introns from certain pre-mRNAs. Thus, Fum1 might regulate the pre-mRNA splicing through Hub1. Since fumarase is frequently mutated in multiple diseases in humans, phenotypes of some of the mutants may be due to loss of Hub1-dependent pre-mRNA splicing.

Chapter 5

The ubiquitin-like protein Hub1 required for trans-splicing in *Caenorhabiditis elegans*

This chapter contains work performed in collaboration with Dr. Kavita babu (Associate professor IISc, Banglore).

K. Kiran Kumar performed all the experiments. Pallavi Sharma (Dr. Kavita babu Lab) made all the *C. elegans* constructs. Pallavi Sharma and Nagesh Kadam helped in designing the probes for microarray analysis. They provided HUB1 knockout worms for splicing-sensitive microarray experiment. Shravan Kumar Mishra and Kavita Babu supervised the experiments, helped with experimental design and data interpretation.

Abstract:

The ubiquitin-like protein Hub1 is a conserved member of the UBL family but functions distinctly from ubiquitin. The protein modifies spliceosomes non-covalently by binding to HIND (Hub1 INteraction Domain) containing pre-mRNA splicing factor Snu66 and promotes alternative splicing. However, the splicing function of Hub1 remains unexplored in multicellular eukaryotes. Hub1 has been reported to play a role in mitochondrial unfolded protein response (UPRmt) in *Caenorhabditis elegans* and worms depleted of HUB1 did not show any defects in pre-mRNA splicing. In the present study, we show that *C. elegans* Hub1 rescues *S. pombe hub1Δ* lethality. It also complements *S. pombe* Hub1 splicing function. We further show that *C. elegans* Hub1 interacts with HIND-containing splicing factors Snu66 and Prp38. The mode of Hub1 interaction through salt bridge formation with Snu66 is conserved in *C. elegans*. Hub1 also associates with the components of spliceosome. It is further shown that HUB1 knockout worms were lethal at Larval 3 stage, which suggests an essential role of HUB1 in development. HUB1 transcript levels are stage-specific and at L3-L4 stage levels are higher than other stages of *C. elegans* life cycle. By using splicing-sensitive microarray, we show that HUB1 knockout led to accumulation of

outron-containing transcripts. These results suggest that Hub1 might play role in RNA trans-splicing.

5.1 Introduction

Most eukaryotic genes are interrupted by introns which are precisely excised out by a process called RNA splicing. In this process, non-coding sequences (introns) are removed, and the coding sequences (exons) are ligated together. Whereas, some exons are constitutively spliced, and many are alternatively spliced to generate multiple mRNAs from a single pre-mRNA (Will and Luhrmann, 2011). In *C. elegans*, around 25% of genes are reported to have alternative splicing events. In contrast, more than 90% of genes undergo alternative splicing (AS) in humans. The average size of an intron in *C. elegans* is around 65 bases, whereas, in humans, the average size of the intron is around 1334 bases. Larger intron size has been correlated with alternative splicing in several organisms (Soler et al., 2013).

In *C. elegans,* around 70% of genes undergo trans-splicing (Allen et al., 2011). Trans-splicing reaction is closely related to cis-splicing; in this process, splice leader (SL) RNA replaces the 5' end of a transcript by spliceosomal splicing. The pre-mRNA undergoing trans-splicing contains 3' splice site, but lack 5' splice site. However, spliced leader RNAs (SL RNA) which are similar to small nuclear RNAs (snRNAs) provide the 5' splice site (Thomas Blumenthal, worm book). In trans-splicing reaction, the 22-nucleotide (nt) SL is donated by a 100-nt SL snRNP (small nuclear ribonucleoprotein) to a pre-mRNA with an outron (the intron-like region between 5' cap and trans-splice site). SL RNAs exists as snRNPs. They have a discrete secondary structure similar to snRNA. They are bound to the sm proteins and have a trimethylguanosine (TMG) cap similar to U snRNAs. In the SL snRNPs the 5'ss is base-paired to the upstream part of SL sequence identical to U1-5'ss base pairing (Blumenthal 2005). In *C. elegans* many genes are transcribed as polycistronic units (operons). In polycistronic units, from single promoter several genes are transcribed. The polycistronic pre-mRNAs are separated into individual cistrons with the help of 3' end formation and trans-splicing (Allen et al., 2011). Trans-splicing occurs in several lower eukaryotes, such as trypanosomatids,

nematodes, trematodes, and euglena (Agabian, 1990; Rajkovic et al., 1990; Tessier et al., 1991). In nematodes, SL RNA interacts with Sm antigens as well specific proteins to form spliced leader ribonucleoprotein particles (SL RNP). SLRNP together with other factors including U2, U4/U6, and U5 snRNPs involve in trans-splicing (Maroney et al., 1990; Denker et al., 1996). The outron at the 5' end of the pre-mRNA undergoes trans-splicing (Conard et al., 1995; Conrad et al., 1993; Conrad et al., 1991). Therefore, the presence of outron distinguishes the genes which undergo trans-splicing from the genes which do not. It is very difficult to isolate outron-containing pre-mRNAs as trans-splicing is very efficient (Conrad et al., 1993). *C. elegans* uses two types of SL RNP (SL1 and SL2) for trans-splicing. More than 50% of C. elegans pre-mRNAs are subjected to SL1 trans-splicing, and about 30% are not trans-spliced. The remaining genes are subjected to SL2 trans-splicing. SL2 trans-spliced genes are downstream genes in polycistronic units that are similar to bacterial operons. The poly-cistronic pre-mRNA is co-transcriptionally processed by cleavage and polyadenylation at the 3' end of each gene. This event is coupled with SL2 trans-splicing event that occurs only 100 nt further downstream. SR (ser/arg) proteins are a family of related phosphoproteins plays various roles in both constitutive and regulated pre-mRNA splicing. (Valcarcel and Green, 1996). SR proteins contain two major domains, an N-terminal region containing one or two RNA recognition motif(s) (RRM) and a C-terminal region rich in RS dipeptides (Caceres and Krainer, 1993; Zuo and Manley, 1993). In nematode *Ascaris lumbricoides*, SR proteins participate in both cis-splicing and trans-splicing events (Sanford and Bruzik, 1999).

In nematode *C. elegans,* Hub1 (referred to as UBL-5) is reported to play a role in mitochondrial unfolded-protein response (UPRmt). Hub1 forms a complex with transcription factor DVE-1 for activating UPRmt pathway (Haynes et al., 2007). In this process, Hub1 increases the expression of mitochondrial chaperones HSP-60 and HSP-6 (Benedetti et al., 2006). Knockdown of HUB1 is reported to inhibit UPRmt (Benedetti et al., 2006); however, the mechanism of its direct involvement in this pathway is not completely clear. In the above study, worms depleted of HUB1 did not show any defects in pre-mRNA splicing. Importantly, Hub1 and HINDs are

conserved; thus, the complex is expected to play a role in RNA splicing. Recent reports suggest that the observed effects of Hub1 in UPRmt in worms could be a consequence of its expected role in pre-mRNA splicing (Bennett and Kaeberlein, 2014).

In most eukaryotes, HIND elements are located in the homologs of RNA splicing factor Snu66/SART1. In plants, HIND is not observed in Snu66 homolog; its absence is compensated by its presence in another splicing factor Prp38 (Mishra et al., 2011). Interestingly, in certain organisms including *C. elegans* HINDs are observed in homologs of both Snu66 and Prp38. At present, implications of more than one splicing factor associating with Hub1 are not clear, but we postulate it to be linked to the higher prevalence of trans-splicing in *C. elegans* (Blumenthal, 2005). Therefore, the present study was aimed to investigate the role of Hub1 in trans-splicing.

5.2 Results

5.2.1 *C. elegans* Hub1 rescues *S. pombe hub1Δ* lethality

In *C. elegans,* Hub1 splicing function has not been addressed. To understand the function of *C. elegans* Hub1, I performed complementation assay in *S. pombe hub1Δ* cells. The expression of *C. elegans* Hub1 rescued the lethality similar to *S. pombe* Hub1 expression at 30°C and 37°C (Figure 5.1). It indicates that *C. elegans* Hub1 able to functionally complement *S. pombe hub1Δ* lethality.

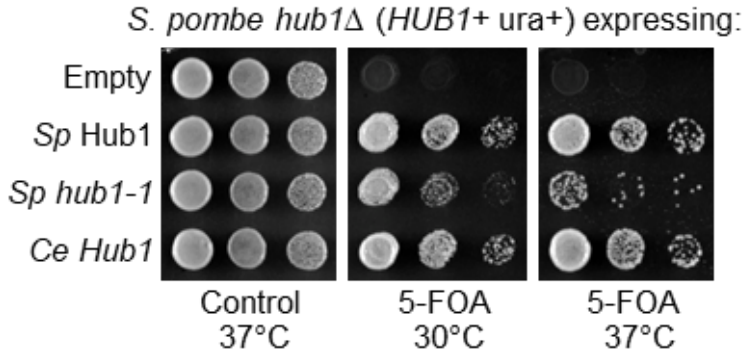


Figure 5.1: *C. elegans* Hub1 complements *S. pombe hub1Δ* mutant

In *S. pombe hub1Δ* cells, the expression of *S. pombe* Hub1 and *C. elegans* Hub1 rescues the lethality as see in plates with FOA. The experiment is similar to Figure 3.4A. However, *S. pombe hub1-1* (temperature-sensitive mutant allele) were able to rescue the lethality of *hub1Δ* cells partially at 30°C but not at non-permissive temperatures 37°C. Transformants were five-fold serial diluted on selective media. Plasmids were expressed from the weak nmt81 promoter. Figure 5.1A is part of Pallavi's thesis.

5.2.2 *C. elegans* Hub1 rescues *S. pombe hub1-1* splicing defects

To understand whether *C. elegans* Hub1 is involved in splicing, I performed complementation assay in *S. pombe hub1-1* cells. *hub1-1* (*hub1-I42S*) cells grow at 30°C and become lethal at 37°C . The *hub1-1* cells show splicing defects at 37°C (Yashiroda and Tanaka, 2004). The expression of *S. pombe* Hub1 and *C. elegans* Hub1 complemented the lethality of *S. pombe hub1-1 cells* at non-permissive temperature 37°C (Figure 5.2A), which indicates that *C. elegans* Hub1 performs a function similar to the wild-type protein. Further to monitor the splicing function, I

performed RT-PCR assays. We isolated total RNA from *hub1-1* cells expressing *S. pombe* Hub1 and *C. elegans* Hub1 constructs. RNAs were converted into cDNAs using random hexamer primers and reverse transcriptase. Following cDNA synthesis, I carried out RT-PCR assays for selected targets, which showed an accumulation of pre-mRNA in splicing-sensitive microarray (shown in chapter 3). The amplified products were visualized on agarose gel electrophoresis. The expression of *C. elegans* Hub1 rescued the accumulation of intron-containing transcripts like *S. pombe* Hub1 expression in *hub1-1* cells at 37°C (Figure 5.2B). We infer from the above findings that Hub1 possesses conserved splicing function across the eukaryotes.

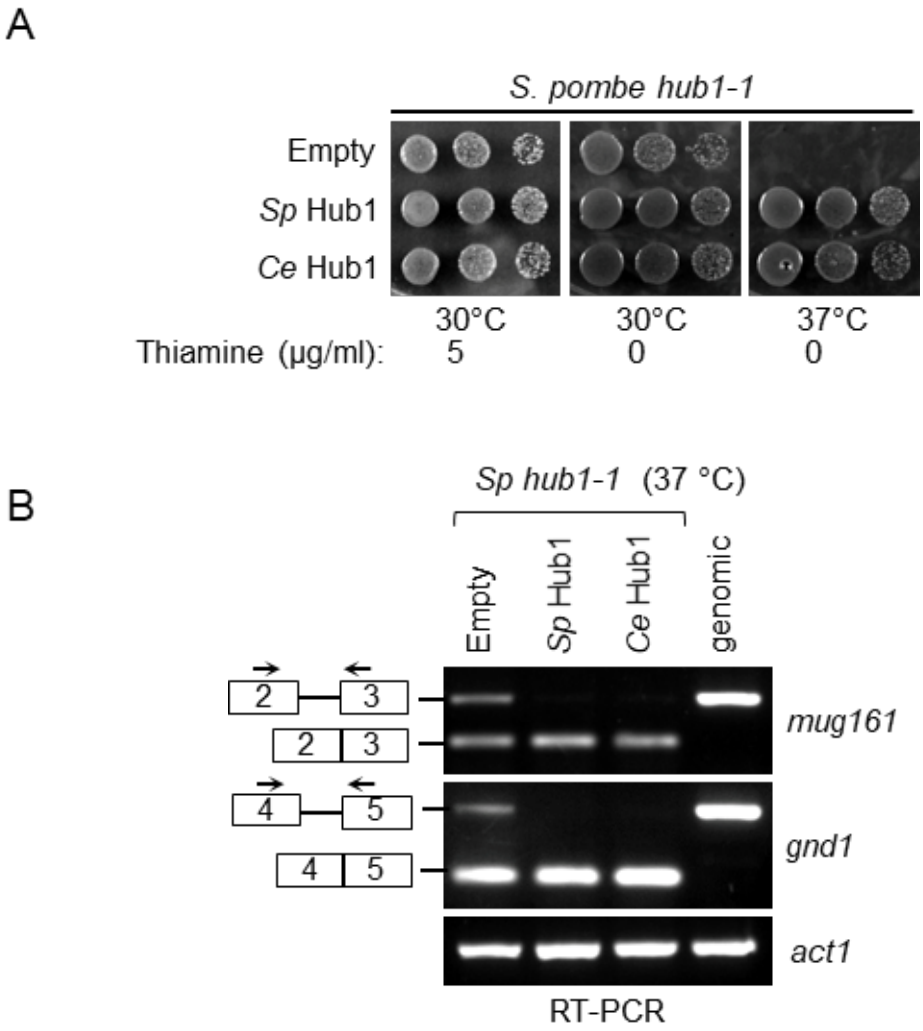


Figure 5.2: *C. elegans* Hub1 rescues *S. pombe hub1Δ* splicing defects

A. In *S. pombe hub1-1* cells, the expression of *S. pombe* Hub1 and *C. elegans* Hub1 rescues the temperature-sensitive at 37°C. Plasmids were expressed from the weak nmt81 promoter. The experiment is similar to Figure 3.4B.

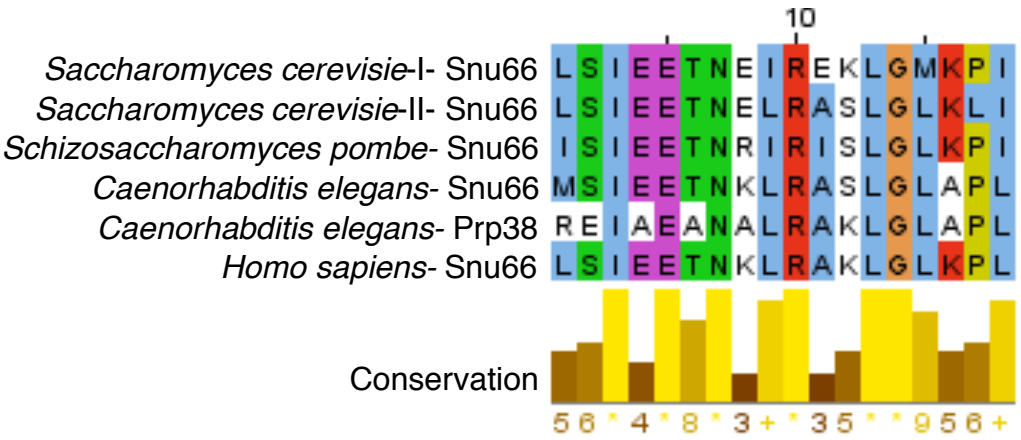
B. Semi-quantitative RT-PCR shows rescue of accumulated intron-containing transcripts for *mug161 and gnd1.* Plasmids were expressed from the weak nmt81 promoter. The experiment is similar to Figure 3.4C. Figure 5.2B is part of Pallavi's thesis.

5.2.3 *C. elegans* Hub1-Snu66 interaction

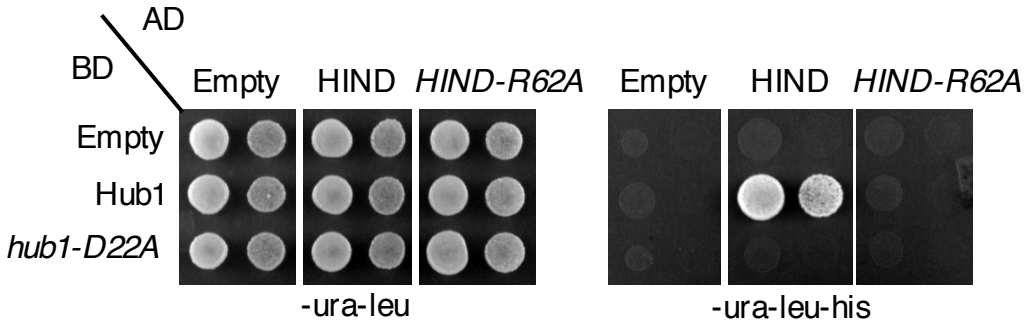
In yeast and human cell lines, Hub1 binds to HIND containing pre-mRNA splicing factor Snu66 to promote alternative splicing (Mishra et al., 2011; Ammon et al., 2014). *S. cerevisiae* contains two HIND elements at the N-terminus of Snu66.

Whereas in *S. pombe* and humans, a single HIND element is present at the N-terminus. In plants and amoebozoa, Snu66 lacks HIND element. However, the absence of HIND in Snu66 is counterbalanced by the presence of HIND in another spliceosomal protein Prp38 at the C-terminus. In case of plasmodium, both Snu66 and Prp38 homologs contain HIND elements (Mishra et al., 2011). We observed two putative HIND elements in *C. elegans*; at the N-terminus of Snu66 and the C-terminus of Prp38 (Figure 5.3A). To verify whether *C. elegans* Hub1 binds to the putative elements, we carried out yeast two-hybrid assay. In this assay, we fused *C. elegans* Hub1 construct to the Gal4 binding domain (BD) and *C. elegans* Snu66-HIND domain fused to the Gal4 activation domain (AD). The AD and BD-fused constructs were transformed in *S. cerevisiae* cells, and selected cells were spotted on histidine lacking plates (-His) to monitor the growth. The yeast cells transformed with Hub1 and Snu66 constructs were viable on histidine lacking plates which indicates the interaction (Figure 5.3B). It was reported that Hub1-Snu66 interaction is through formation of salt bridge and hydrophobic interactions (Mishra et al., 2011). To confirm whether the mode interaction is conserved. We replaced the salt bridge forming residues on Hub1-D22 and Snu66-R62 with amino acid alanine (A) by quick-change site-directed mutagenesis. Using the mutagenized clones yeast two-hybrid assay was performed to test the interaction between the two proteins. We observed that altering salt bridge forming residues on either Hub1 or Snu66 abolishes the interaction between Hub1-Snu66 (Figure 5.3B). Thus, the mode of interaction seems to be conserved. To further understand whether Hub1-Snu66 interact directly, we purified Hub1 and HIND containing proteins from bacterial cells (clones were made by Pallavi Sharma, Dr. Kavita lab at IISER, Mohali) and GST-pull down assay was carried out. In this assay, equimolar concentrations of Hub1 and GST-HIND containing proteins were mixed and incubated on roller at 4°C for 1 hr. GSH beads were added to the mixture and incubated on roller at 4°C for 1 hr. Later, the mixture was washed, and pull-down GSH fraction was eluted using HU-buffer and analyzed on SDS-PAGE. GST-HIND were able to pull-down Hub1, which indicates a direct interaction between *C. elegans* Hub1 and Snu66 (Figure 5.3C).

A


10
Saccharomyces cerevisie-I- Snu66
Saccharomyces cerevisie-II- Snu66
Schizosaccharomyces pombe- Snu66
Caenorhabditis elegans- Snu66
Caenorhabditis elegans- Prp38
Homo sapiens- Snu66
Conservation

B


AD
BD
Empty
HIND
HIND-R62A
Empty
HIND
HIND-R62A
Empty
Hub1
hub1-D22A
-ura-leu
-ura-leu-his

C


Input
GST-PD
GST-Snu66-HIND
GST
6His-Hub1
GST-Snu66-HIND
GST
6His-Hub1
Coomassie

Figure 5.3: *C. elegans* Hub1-Snu66 binding

A. HIND protein sequence alignment across the organisms.
B. Yeast two-hybrid interaction studies with the cells expressing *C. elegans* Hub1, *hub1-D22A* mutants fused to Gal4 binding domain (BD) and with Snu66-HIND and *Snu66-HIND (R62A)* mutant fused to activation domain (AD). The experiment is similar to Figure 4.1A. Figure 5.3B is part of Pallavi's thesis (clones were made by Pallavi Sharma).
C. GST–pull-down assays were carried using GSH beads. Bacterially purified equimolar concentrations of 6xHis-Hub1 were mixed with GST, GST-Snu66(HIND) proteins. GST-Snu66(HIND) was able to pulldown Hub1, GST was used as a control. Input represents the one-tenth amount of the total proteins. GST and GST-Snu66(HIND) were purified according to standard GST-tag affinity purification, whereas, *6xHis*-Hub1 was purified according to His-tag affinity purification under denaturing conditions as described in the methods section (2.2.17). The experiment was performed twice. Figure 5.3C is part of Pallavi's thesis (clones were made by Pallavi Sharma).

5.2.4 *C. elegans* Hub1-Prp38 interaction

To verify whether *C. elegans* Hub1 binds to the putative HIND elements of Prp38, I carried out yeast two-hybrid assay. I did not observe any interaction between Hub1 and Prp38-HIND element. I then carried out the GST pull-down assay. In this assay, Hub1, GST, and GST-Prp38-HIND containing proteins were induced with IPTG in bacterial cells. Lysates with soluble proteins were mixed and incubated on roller at 4°C for 1 hr. Further, GSH beads were added to the mixture and incubated on roller at 4°C for 1 hr. Later, the mixture was washed, and pull-down GSH fraction was eluted using HU-buffer and analyzed on SDS-PAGE. GST-Prp38 (HIND) containing protein was able to pull down Hub1, whereas GST control could not pull down Hub1 (Figure 5.4A). I further verified the interaction with Hub1 specific antibody (Figure 5.4B). Thus, we infer that *C. elegans* Hub1 interacts with HIND-containing splicing factors Prp38.

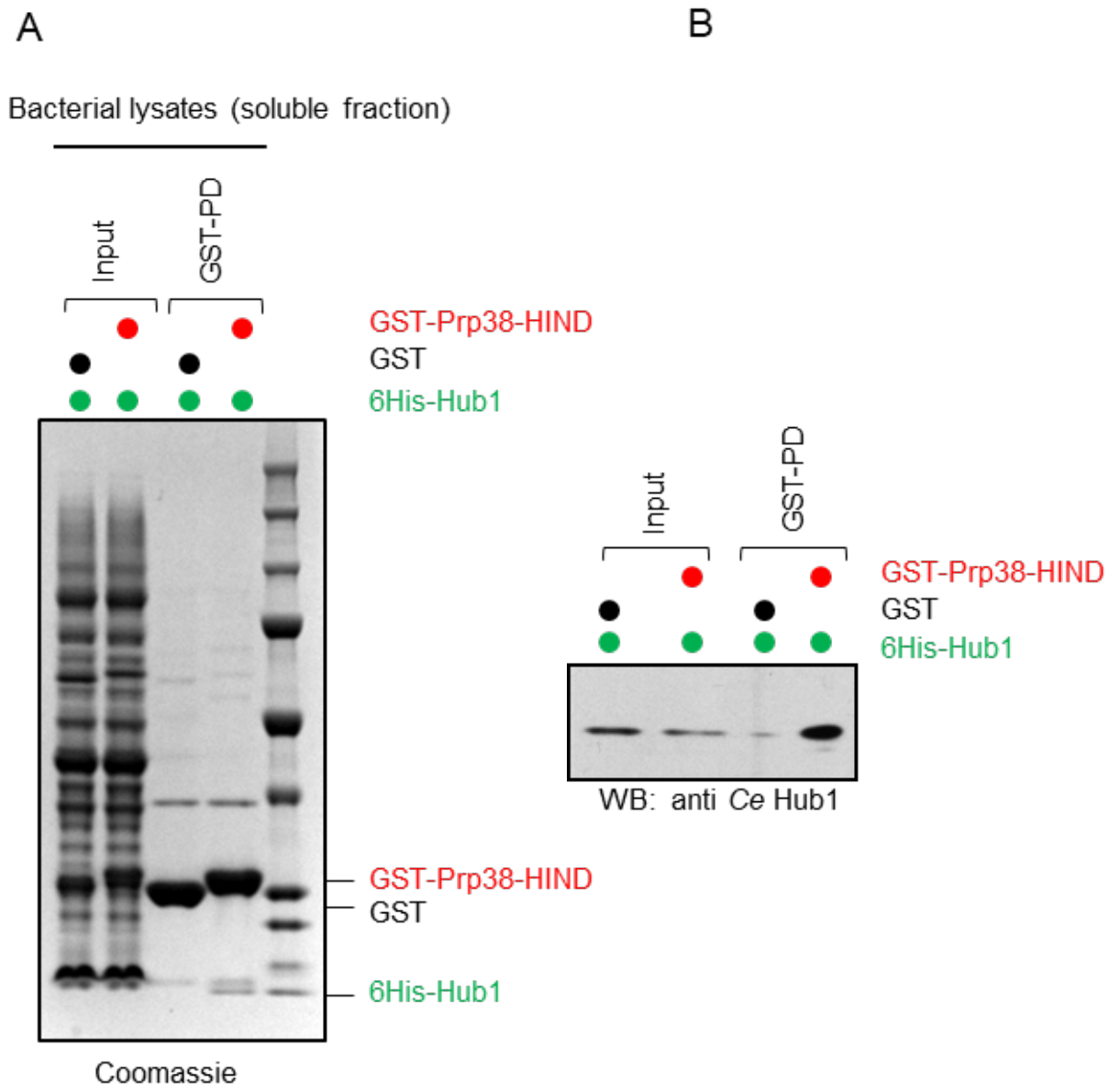


Figure 5.4: Hub1-Prp38 complex

A. GST–pull-down assays were carried using GSH beads. Bacterial cells expressing GST, GST-Prp38 (HIND), and 6xHis-Hub1 were induced with IPTG. The bacterial soluble fraction containing GST-Prp38 (HIND) and 6xHis-Hub1 were mixed together and GST-pull down assay performed. GST-Prp38 (HIND) was able to pulldown Hub1, GST was used as a control. Input represents the one-tenth amount of the total proteins. The experiment was performed twice (Note: Last lane represents protein marker).

B. GST pull-down samples in the above experiment were loaded on SDS-PAGE. After electrophoresis, the proteins were detected by immunoblot using *C. elegans* Hub1 antibody raised in rabbit.

5.2.4 *C. elegans* Hub1 associates with splicing factors

The expression of *C. elegans* Hub1 rescued the accumulation of intron-containing transcripts in *S. pombe hub1-1* cells. To understand the molecular mechanism of Hub1 in splicing, we immunoprecipitated Hub1 complex with anti-FLAG beads. *C. elegans* cells expressing 3xFlag-Hub1 constructs were used for the immunoprecipitation experiments (the 3xFlag-Hub1 plasmid DNA was micro-injected in *C. elegans* gonads). The immunoprecipitated complex was eluted and analyzed by mass spectrometry. The lysates from wild type worms were used as a control. The splicing factors enriched in Hub1 purified complexes are listed below (Table 5.1). Only a few splicing factors were co-purified with Hub1 protein; possibly Hub1 might interact with the components of the spliceosome transiently. Thus, we infer that *C. elegans* Hub1 associates with some of the components of the spliceosome.

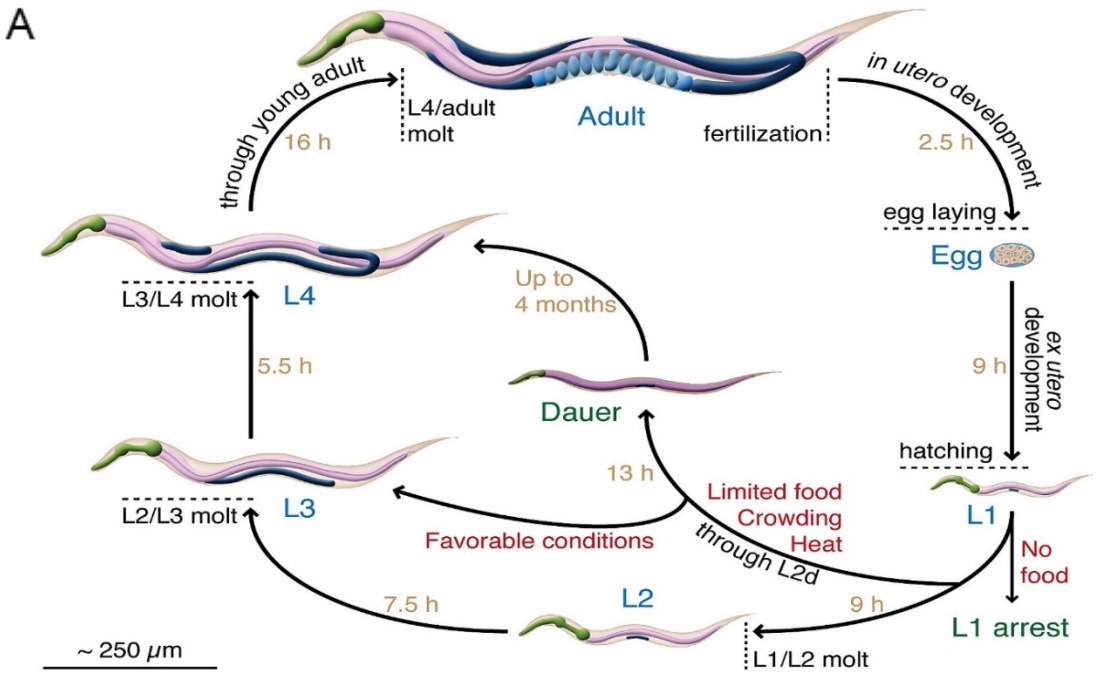
Gene symbol	Annotation	Intensity (Wild-type)	Intensity (FLAG-Hub1)
Hub1/ ubl-5	Ubiquitin-like protein 5	Nd	5100000
hel-1	Spliceosome RNA helicase DDX39B homolog	54000	930000
hrp-2	HnRNP A1 homolog	50000	700000
rsp-3	Probable splicing factor, arginine/serine-rich 3	140000	650000
ddx-17	DEAD boX helicase homolog	Nd	380000
rsp-2	Probable splicing factor, arginine/serine-rich 2	Nd	210000
prp-19	Pre-mRNA-processing factor 19	Nd	180000
rsp-1	Probable splicing factor, arginine/serine-rich 1	Nd	150000
Y46G5A.4	Putative U5 small nuclear ribonucleoprotein 200 kDa helicase	Nd	140000
prp-8	Pre-mRNA-splicing factor 8 homolog	Nd	130000
rsp-6	Probable splicing factor, arginine/serine-rich 6	Nd	65000
eftu-2	Elongation Factor TU family	Nd	57000

Gene symbol	Annotation	Intensity (Wild-type)	Intensity (FLAG-Hub1)
prp-1	Putative RNA-binding protein PRP-1	Nd	18000
hrpf-1	HnRNP F homolog	Nd	17000

Table 5.1: *C. elegans* Hub1 interacts with splicing factors: A intensity of unique peptides for the identified proteins in mass spectrometry. 3xFlag-Hub1 expressing *C. elegans* cells were immunoprecipitated using anti-FLAG antibody, Co-IP proteins were subjected and analyzed by mass spectrometry. Nd, not detected (the experiment was peformed along with Pallavi Sharma).

5.2.5. HUB1 expression is crucial after larval stage 3

In *C. elegans*, HUB1 is essential for viability. HUB1 knockout worms do not survive after the L3 stage; hence, HUB1 knock out strains was maintained in the heterozygous condition. HUB1 knockout worms lethality was rescued with expression of HUB1 gene (Figure 5.5B). We performed quantitative RT-PCR assays to monitor the HUB1 transcript level at different stages in wild-type worms. We found that the expression of HUB1 varies in stage specific manner and at developmental L3 stage HUB1 transcript levels were higher than other stages of development (Figure 5.5C). Thus, Hub1 is essential for the development of *C. elegans*.



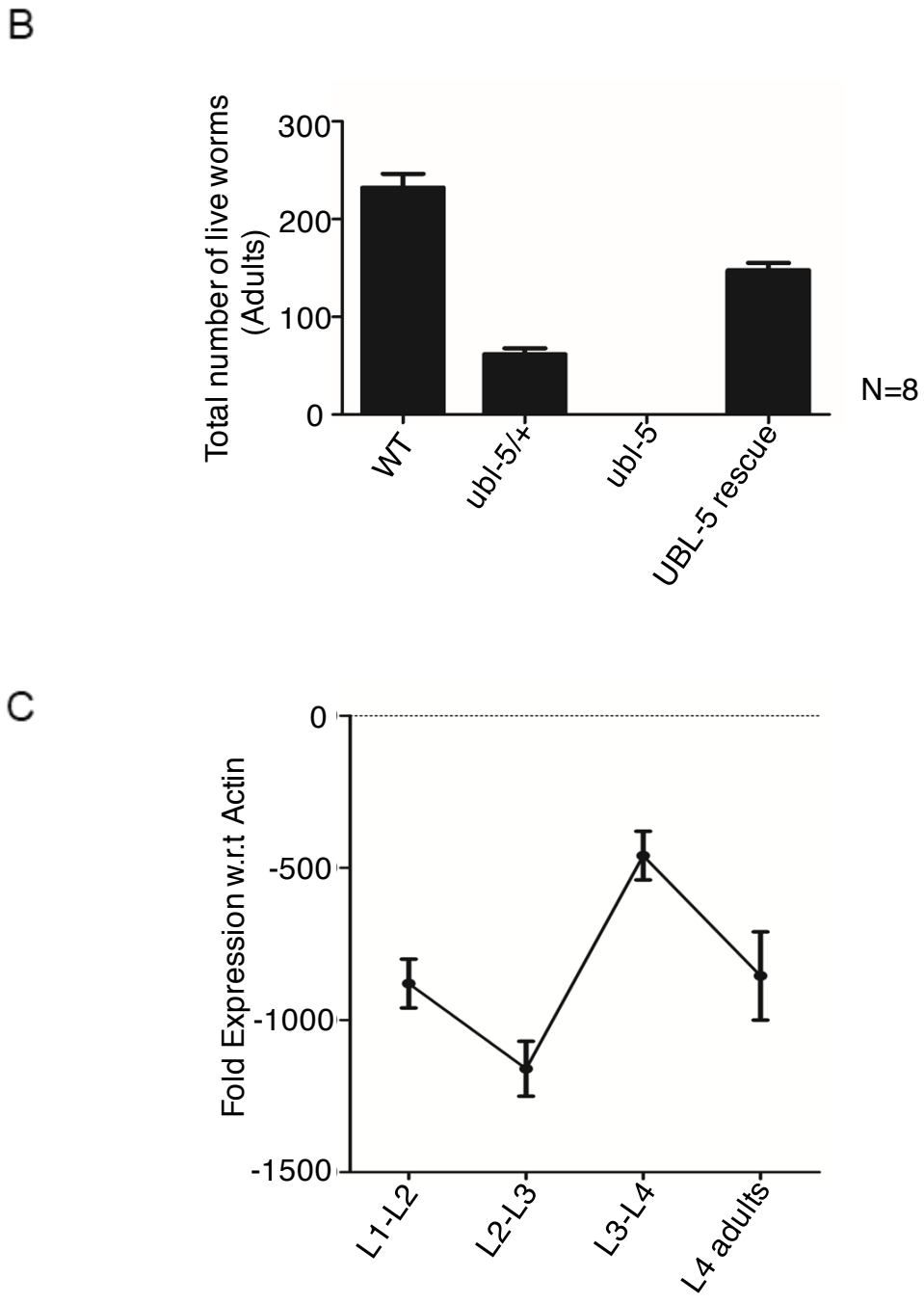


Figure 5.5: Hub1 expression in *C. elegans*

A. *C. elegans* life cycle consists of four larval stages L1, L2, L3, L4, and adult stage. In the absence of food after the L1 stage they get transformed into dauer stage (Image adapted from Byerly et al., 1976).
B. The Graph represents the rescue of lethality in *HUB1* knockout worms. The expression of genomic HUB1 in *HUB1 knockout* worms rescued the lethality. The total number of adult's worms surviving from a single parent worm is counted for WT, *ubl-5*/+, *ubl-5* and UBL-5 rescue (Data is obtained from Pallavi thesis).
C. The HUB1 expression in L1-L2, L2-L3, L3-L4, and L4-adult stages using real-time PCR. L3-L4 shows the highest expression when compared to all other stages (Data is obtained from Pallavi thesis).

5.2.6 *C. elegans* HUB1 mutants show defects in trans-splicing

In yeast and humans, Hub1 plays a role in pre-mRNA splicing (Yashiroda and Tanaka, 2004; Mishra et al., 2011; Ammon et al., 2014). In the nematode *C. elegans*, Hub1 is implicated in unfolding protein response in mitochondria (UPR^{mt}) (Benedetti et al., 2006), where it forms a complex with transcription factor DVE-1 (Haynes et al., 2007). In the above study, RNAi mediated *HUB1* knockdown worms did not show any splicing defects (Benedetti et al., 2006). To understand the role of *C. elegans* Hub1 in RNA splicing, we performed splicing sensitive microarray. In *C. elegans* 25% of genes are alternatively spliced, and 75% of genes undergo trans-splicing (Ramani et al., 2011; Allen et al., 2011). As Hub1 binds to both Snu66 and Prp38 proteins in *C. elegans*, we speculated that Hub1 might play a role in both alternative-splicing and trans-splicing. To study trans-splicing events, we customized splicing-sensitive microarray. For this study, we chose neuronal genes which were reported to be alternatively spliced. Out of this genes, we chose the genes which undergo either SL1/ SL2 mediated trans-splicing. *HUB1* knockout is lethal in *C. elegans*, as the worms survive up to L3 stage only and are maintained in heterozygous state with the help of a balancer chromosome. To perform splicing-sensitive microarray, for each gene we designed a minimum of three probes; the outronic probes detected pre-mRNA, trans-spliced junction probes (SL1) detected mature SL1-mRNA, and trans-spliced junction probes (SL2) detected mature SL2-mRNA. The total RNA was isolated from wild-type and *HUB1* knock out worms at L3 stage. 500ng each of total RNA was reverse transcribed at 40°C using oligo dT primer tagged to a T7 polymerase promoter and converted to double-stranded cDNA. Synthesized double-stranded cDNA was used as a template for cRNA generation. cRNA was generated by *in vitro* transcription, and the dye Cy3 CTP (Agilent) was incorporated during this step. The cDNA synthesis and *in vitro* transcription steps were carried out at 40°C. 600ng of labeled cRNA sample was fragmented at 60°C and hybridized on to an Agilent Gene Expression Microarray 8X60K. Microarray heat map represents fold change values obtained by comparing mutant samples with respect to wild-type samples (Figure 5.6). From splicing-sensitive microarray data, we found that *HUB1* knockout worms showed accumulation of outron containing transcripts compared to wild-type worms. We thus infer that Hub1 is required for trans-splicing in *C. elegans*.

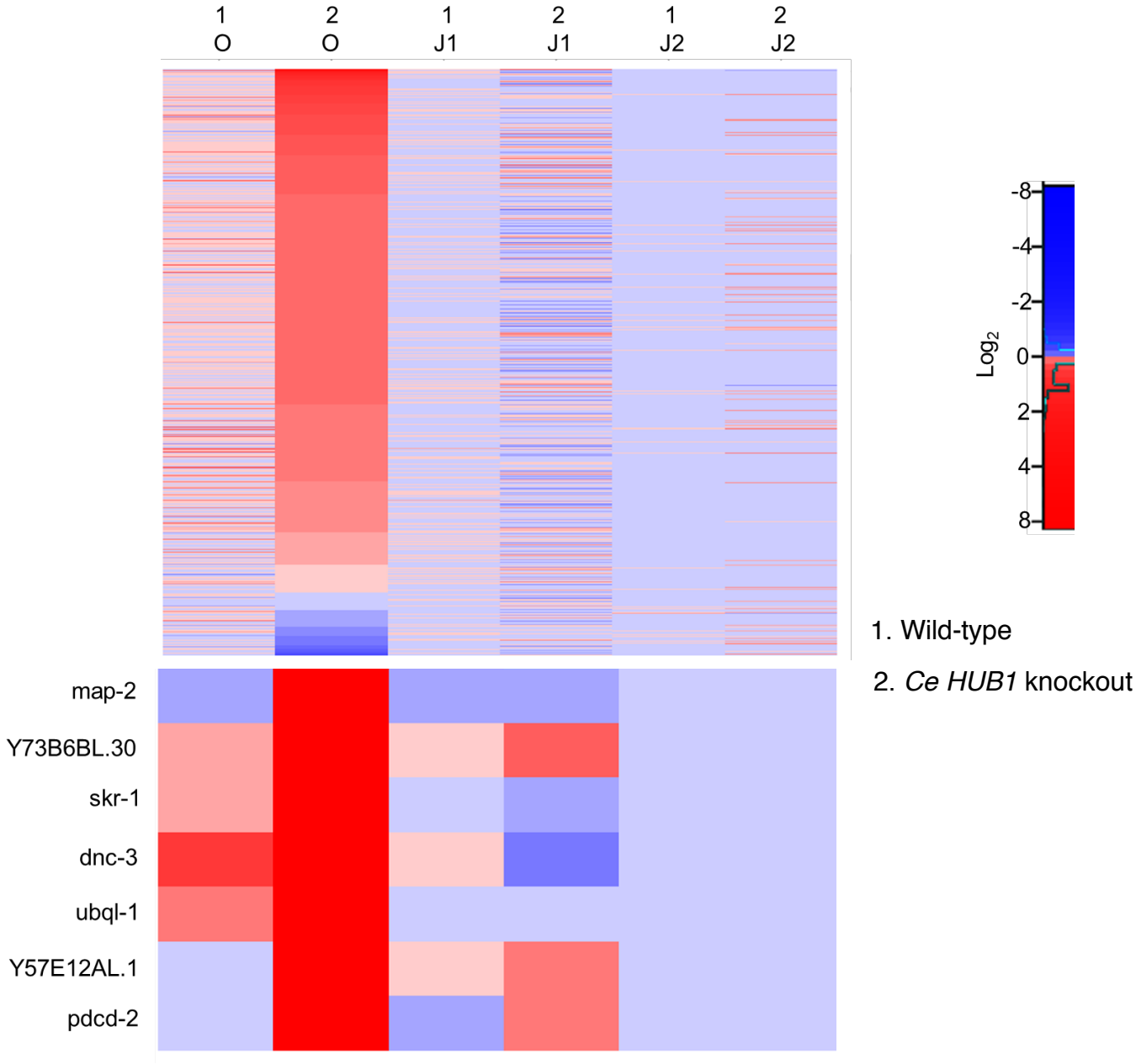


Figure 5.6. Hub1 is required for trans-splicing in *C. elegans*

Analysis of total RNA from wild-type and *HUB1* knockout worms using splicing-sensitive microarray. Microarray heat-map shows the $\log_2$ mutant samples with respect to wild-type samples for outron-containing transcripts (O), SL1 trans-spliced transcripts (J1), and SL1 trans-spliced transcripts (J2). Here, red color represents accumulation; light blue denotes no change and dark blue shows reduction of signal for the transcripts. The expanded view of particular transcripts that are affected in *HUB1* knockout worms is shown. (The experiment was performed along with Pallavi Sharma and Nagesh Kadam).

5.3 Discussion

Here we show that *C. elegans* Hub1 splicing function is conserved and it interacts with two HIND containing spliceosomal proteins Snu66 and Prp38. The mode of interaction with the HIND element is conserved. The expression of *C. elegans* Hub1 complemented *S. pombe hub1Δ* lethality, and also rescued the accumulation of intron-containing transcripts in *S. pombe hub1-1 cells.* We also show that *C. elegans* Hub1 associates with the splicing factors *in vivo*. *HUB1* knock out worms are lethal as the worms do not survive after the L3 stage. We found that the expression of HUB1 varies in stage-specific manner and at developmental L3 stage HUB1 transcript levels were higher than other stages of development. Thus, Hub1 is essential for the development of worms. Splicing microarray with RNA from *HUB1* knockout worms showed the accumulation of outron containing transcripts compared to wild-type worms. We thus infer that Hub1 is required for trans-splicing in *C. elegans*.

C .elegans Hub1 complements *S. pombe hub1-1* cells lethality and splicing defects

The Hub1 functions in pre-mRNA splicing through non-covalent interaction with the protein substrates. In *S. cerevisiae*, Hub1 binds to Snu66, a protein of the U4/U6.U5 small nuclear-ribonucleic particle (tri-snRNP) to promote alternative splicing. In intron-poor *S. cerevisiae*, Hub1 is not essential for viability; probably its function is restricted to promote alternative splicing of *SRC1* pre-mRNA. Importantly, in *S. cerevisiae*, Hub1 becomes essential if these cells are partially defective in certain spliceosomal proteins like Prp8 (Mishra et al., 2011). In intron-rich *S. pombe* and human cells, Hub1 is required for optimal splicing of a large number of pre-mRNAs. In these cells Hub1 becomes essential for viability, most likely due to its splicing functions (Wilkinson et al., 2004; Mishra et al., 2011; Ammon et al., 2014).

In nematode *C. elegans,* Hub1 (referred to as UBL-5) is reported to play a role in mitochondrial unfolded-protein response (UPRmt), a stress signaling pathway that senses increase in unfolded proteins in the mitochondria. Where, Hub1 forms complex with transcriptional factor DVE-1 and signals to the nucleus for expression of mitochondrial chaperones HSP-60 and HSP-6. It has been shown that HUB1

knockdown worms did not show any splicing defects (Benedetti et al., 2006). We show that the expression of *C. elegans* Hub1 expression compensates for the loss of *S. pombe* Hub1 and performs the functions similar to the wild-type protein. Most likely *C. elegans* Hub1 rescuing the *S. pombe* lethality could be due to its compensatory splicing function. *C. elegans* Hub1 also rescues the accumulation of intron-containing transcripts in *S. pombe hub1-1* cells at 37°C , indicating that *C. elegans* Hub1 splicing function remains conserved.

C .elegans Hub1-HIND complex

Hub1 binds to HIND (Hub1 Interaction Domain) region on Snu66 and/or Prp38. HIND element is present on Snu66 in yeast and vertebrates, but in plants, another tri-snRNP protein Prp38 contains the HIND element. In *plasmodium,* HIND elements are present on both Snu66 and Prp38 (Mishra et al., 2011). Hub1 also binds to Prp5, a DEAD-box helicase to promote alternative splicing (Karaduman et al., 2017). From our yeast two-hybrid and pull-down assays, we show that Hub1 binds to both the spliceosomal proteins Snu66 and Prp38 in *C. elegans*. The mode of interaction is also conserved from yeast to vertebrates. Hub1 might have evolved to bind with additional HIND contain proteins to perform wider functions in RNA-splicing.

C .elegans Hub1 associates with splicing factors

From yeast two-hybrid screens, it has been shown that *C. elegans* Hub1 interacts with Npp-9, a nuclear pore complex protein, and Hsp-17, a heat shock protein (Simonis et al., 2009). From this study, we show that *C. elegans* Hub1 associates with splicing factors, of which notably with arginine/serine-rich Rsp family proteins such as Rsp-1, Rsp-2, Rsp-3, and Rsp-6. Hub1 also associates other splicing factors such as Hrpf-1 (HnRNP F homolog) and Hrp-2 (HnRNP A1 homolog). Trans-acting splicing factors such as SR proteins bind with the regulatory sequences on pre-mRNA for splicing control (Gontarek and Derse, 1996; Kanopka et al., 1996). Heterogeneous nuclear proteins (hnRNP) A/B family members bind to regulatory sequences on exons and inhibit splicing through preventing SR proteins from binding to exons (Caputi et al., 1999; Zhu et al., 2001). Interestingly, in our study Hub1

associated with both types of trans-acting splicing factors. Further studies need to be carried out to understand the relevance of such interactions.

C .elegans Hub1 required for trans-splicing

In human cell lines, Hub1 has been shown to modulate the spliceosome and facilitate alternative splicing. Additionally, it also is shown that Hub1 is required for splicing of certain introns (Ammon et al., 2014). *C. elegans HUB1* knock out worms are lethal as the worms do not survive after the L3 stage. From splicing-sensitive microarrays, we show that Hub1 is also required for alternative splicing in *C. elegans* (This collaborative work includes as part of Pallavi Sharma thesis). Hub1 splicing function seems to be evolutionarily conserved in different organisms.

In *C. elegans*, the 5' ends of pre-mRNAs of many genes undergo trans-splicing with either SL1 or SL2 two short leader RNAs and around 70% of genes undergo trans-splicing events (Allen et al., 2011). In this process, SR recruits U2 snRNP to the branch point of natural trans-splicing substrates (Sanford and Bruzik, 1999; Romfo et al., 2001). In trans-splicing reaction, SL snRNPs and U4/U6.U5 snRNPs form tetra-snRNP or separate SL snRNP, and the U4/U6.U5 snRNPs potentiates the transspliceosome assembly. SR proteins also promote the entry of the U4/U6.U5 snRNP into the cis-spliceosome (Roscigno and Blanco, 1995). SR proteins appear essential for the formation of complete, catalytically active trans spliceosome for trans-splicing (Furuyama and Bruzik, 2002). From our splicing-sensitive microarray assay, we show that Hub1 is required for trans-splicing. As Hub1 containing complexes associated with SR proteins (Rsp-1, Rsp-2, Rsp-3, and Rsp-6), it is possible that Hub1 together with SR proteins might facilitate the trans spliceosome assembly and potentiates the trans-splicing reaction in *C. elegans*.

5.4 Conclusion

We demonstrated a possible role of Hub1 in trans-splicing in an intron-rich organism *C. elegans*. We show that Hub1 associates with components of the spliceosome, besides the HIND containing splicing factors, Snu66 and Prp38. Hub1 splicing function is conserved in *C. elegans.* Hub1 is required for larval to adult development

of the worms as the *HUB1* knock out worms are lethal at L3 stage. In addition to its role in alternative splicing (independent study by a colleague, Pallavi Sharma), we found that Hub1 is required for trans-splicing in *C. elegans*.

Appendix

Gene names	*APEX2-hub1-DD*		APEX2-Hub1	
atp2	10	7	9	10
fba1	9	9	12	9
hsp60	8	10	11	11
ilv5	6	5	6	7
Adenosylhomocysteinase	6	6	4	6
ssa2	5	7	7	6
gpm1	5	2	7	7
adh1	4	1	8	8
Putative D-3-phosphoglycerate dehydrogenase	4	3	7	6
tif1	4	2	3	5
SPBC24C6.04	3	4	5	5
tpi1	3	3	3	3
SPCC1827.06c	3	3	3	3
kap123	3	0	3	3
SPAC25B8.12c	2	2	6	5
SPAC56E4.03	2	1	5	5
tpx1	2	2	4	2
eca39	2	2	3	2
ppa1	2	2	2	2
hxk1	2	1	2	2
tef3	2	2	1	2
rps0b	1	1	2	1
ade8	1	1	2	2
SPAC1F12.07	1	nd	2	2
SPAC3G9.11c	1	3	2	3
SPCC1827.03c	1	1	2	2
dld1	1	1	1	1
btf3	1	nd	1	1
cct4	nd	1	3	3

gst2;gst1	nd	1	3	3
cct2	nd	nd	3	3
SPAC1F5.02	nd	nd	2	2
SPAC19G12.09	nd	nd	2	2
rps10b	nd	nd	1	nd

Table 1: Hub1 specific biotinylated proteins. The number of unique peptides in mass spectrometry for an identified protein. Biotinylated complexes were pulled-down by streptavidin beads from cells expressing APEX2-Hub1, *APEX2-hub1-DD* constructs. Biotinylated proteins were subjected and analyzed by mass spectrometry. nd, not detected.

Table 2: Collaborators and contributions

Collaborators and contributions	Experiments performed	Displayed in section
Dr. Jeffrey A. Pleiss	Splicing-sensitive microarray	Chapter 3-Figure 3.1A
Dr. Ranabir Das	Circular dichroism (CD)	Chapter 3-Figure 3.4E
Dr. Kuljeet Sandhu and Dr. Arashdeep Singh	Bioinformatic analysis	Chapter 3-Figure 3.9A
Dr. Kavita Babu and Pallavi Sharma	Rescue of lethality in *HUB1* knockout worms, quantitative RT-PCR and splicing sensitive microarray	Chapter 5-Figure 5.5B, C & Figure 5.6

References

1. Agabian, N. Trans Splicing of Nuclear Pre-mRNAs Minireview. Cell 61, 1157–1160 (1990).
2. Ajuh, P. Functional analysis of the human CDC5L complex and identification of its components by mass spectrometry. EMBO J. 19, 6569–6581 (2002).
3. Allen, M. A., Hillier, L. W., Waterston, R. H. & Blumenthal, T. A global analysis of C. elegans trans-splicing. Genome Res. 21, 255–264 (2011).
4. Ammon, T. et al. The conserved ubiquitin-like protein Hub1 plays a critical role in splicing in human cells. J. Mol. Cell Biol. 6, 312–323 (2014).
5. Anderson, J. S. J. & Parker, R. The 3' to 5' degradation of yeast mRNAs is a general mechanism for mRNA turnover that requires the SK12 DEVH box protein and 3' to 5' exonucleases of the exosome complex. EMBO J. 17, 1497–1506 (1998).
6. Ast, G. How did alternative splicing evolve? Nat. Rev. Genet. 5, 773–782 (2004).
7. Barberán-Soler, S. & Ragle, J. M. Alternative splicing regulation of cancer-related pathways in caenorhabditis elegans: An in vivo model system with a powerful reverse genetics toolbox. Int. J. Cell Biol. 2013, (2013).
8. Baroni, E., Viscardi, V., Cartagena-Lirola, H., Lucchini, G. & Longhese, M. P. The Functions of Budding Yeast Sae2 in the DNA Damage Response Require Mec1- and Tel1-Dependent Phosphorylation. Mol. Cell. Biol. 24, 4151–4165 (2004).
9. Barrass, J. D. & Beggs, J. D. Splicing goes global. Trends Genet. 19, 295–298 (2003).
10. Barteri, M., Coluzza, C., & Rotella, S. Fractal aggregation of porcine fumarase induced by free radicals. Biochimica et Biophysica Acta - Proteins and Proteomics, 1774(2), 192–199 (2007).
11. Benedetti, C., Haynes, C. M., Yang, Y., Harding, H. P. & Ron, D. Ubiquitin-like protein 5 positively regulates chaperone gene expression in the mitochondrial unfolded protein response. Genetics 174, 229–239 (2006).
12. Bennett, C. F. & Kaeberlein, M. The mitochondrial unfolded protein response and increased longevity: Cause, consequence, or correlation? Exp. Gerontol. 56, 142–146 (2014).
13. Bentley, D. Coupling RNA polymerase II transcription with pre-mRNA processing. Curr. Opin. Cell Biol. 11, 347–351 (1999).
14. Berget.S. Exon Recognition in Vertebrate Splicing. J. Biol. Chem. 270, 2411–2414 (1995).
15. Blumenthal, T. Trans-splicing and operons. WormBook 1, 1–9 (2005).
16. Byerly, L., Scherer, S. & Russell, R. L. The life cycle of the nematode Caenorhabditis elegans. *Dev. Biol.* **51**, 34–48 (1976).
17. Cáceres, J. F. & Krainer, A. R. Functional analysis of pre-mRNA splicing factor SF2/ASF structural domains. EMBO J. 12, 4715–4726 (1993).
18. Caputi, M. & Zahler, A. M. SR proteins and hnRNP H regulate the splicing of the HIV-1 tev-specific exon 6D. EMBO J. 21, 845–855 (2002).

19. Caputi, M., Mayeda, A., Krainer, A. R. & Zahler, A. M. hnRNP A/B proteins are required for inhibition of HIV-1 pre-mRNA splicing. EMBO J. 18, 4060–4067 (1999).

20. Chanarat, S. & Mishra, S. K. Emerging Roles of Ubiquitin-like Proteins in Pre-mRNA Splicing. Trends Biochem. Sci. 43, 896–907 (2018).

21. Chang, T. H., Tung, L., Yeh, F. L., Chen, J. H. & Chang, S. L. Functions of the DExD/H-box proteins in nuclear pre-mRNA splicing. Biochim. Biophys. Acta - Gene Regul. Mech. 1829, 764–774 (2013).

22. Conrad, R., Lea, K., and Blumenthal, T. SL1 trans-splicing specified by AU-rich synthetic RNA inserted at the 5' end of Caenorhabditis elegans pre-mRNA. Rna 1, 164–170 (1995).

23. Conrad, R., Liou, R. F. & Blumenthal, T. Conversion of a trans-spliced C. elegans gene into a conventional gene by introduction of a splice donor site. Trends Genet. 9, 234–234 (1993).

24. Conrad, R., Thomas, J., Spieth, J. & Blumenthal, T. Insertion of part of an intron into the 5' untranslated region of a Caenorhabditis elegans gene converts it into a trans-spliced gene. Mol. Cell. Biol. 11, 1921–1926 (1991).

25. Cooper, T. A, & Mattox, W. GENE REGULATION '97 The Regulation of Splice-Site Selection, and Its Role in Human Disease Intronic Splicing Elements and Splicing Regulators Although modulation of the nuclear concentrations of constitutive RNA processing factors causes some al. Am. J. Hum. Genet 61, 259–266 (1997).

26. Cope, G. A. & Deshaies, R. J. COP9 signalosome: A multifunctional regulator of SCF and other cullin-based ubiquitin ligases. Cell 114, 663–671 (2003).

27. Cramer, P. et al. Coupling of Transcription with Alternative Splicing. Mol. Cell 4, 251–258 (1999).

28. Cramer, P., Pesce, C. G., Baralle, F. E. & Kornblihtt, A. R. Functional association between promoter structure and transcript alternative splicing. *Proc. Natl. Acad. Sci. U. S. A.* **94**, 11456–11460 (1997).

29. Cvitkovic, I. & Jurica, M. S. Spliceosome database: A tool for tracking components of the spliceosome. Nucleic Acids Res. 41, 132–141 (2013).

30. Dahmus, M. E. Reversible phosphorylation of the C-terminal domain of RNA polymerase II. J. Biol. Chem. 271, 19009–19012 (1996).

31. Denker, J. A., Maroney, P. A., YU, Y. T., Kanost, R. A. & Nilsen, T.W. Multiple requirements for nematode spliced leader RNP function in trans-splicing. Rna 2, 746–755 (1996).

32. Dittmar, G. A. G., Wilkinson, C. R. M., Jedrzejewski, P. T. & Finley, D. Role of a Ubiquitin-Like Modification in Polarized Morphogenesis. Science 295, 2442–2446 (2002).

33. Dvinge, H., Kim, E., Abdel-Wahab, O. & Bradley, R. K. RNA splicing factors as oncoproteins and tumour suppressors. Nat. Rev. Cancer 16, 413–430 (2016).

34. Elkon, R., Ugalde, A. P. & Agami, R. Alternative cleavage and polyadenylation: Extent, regulation and function. Nat. Rev. Genet. 14, 496–506 (2013).

35. Eser, P. et al. Determinants of RNA metabolism in the Schizosaccharomyces pombe genome. Mol. Syst. Biol. 12, 857 (2016).

36. Fabrizio, P. *et al.* The Evolutionarily Conserved Core Design of the Catalytic Activation Step of the Yeast Spliceosome. *Mol. Cell* **36**, 593–608 (2009).

37. Flotho, A. & Melchior, F. Sumoylation: A Regulatory Protein Modification in Health and Disease. Annu. Rev. Biochem. 82, 357–385 (2013).

38. Fox-Walsh, K. L. et al. The architecture of pre-mRNAs affects mechanisms of splice-site pairing. Proc. Natl. Acad. Sci. 102, 16176–16181 (2005).

39. Furuyama, S. & Bruzik, J. P. Multiple Roles for SR Proteins in. Society 22, 5337–5346 (2002).

40. Gontarek, R. R. & Derse, D. Interactions among SR proteins, an exonic splicing enhancer, and a lentivirus Rev protein regulate alternative splicing. Mol. Cell. Biol. 16, 2325–2331 (1996).

41. Gottschalk, A. et al. Identification by mass spectrometry and functional analysis of novel proteins of the yeast [U4/U6•U5] tri-snRNP. EMBO J. 18, 4535–4548 (1999).

42. Grabowski, P. J. & Black, D. L. Alternative RNA splicing in the nervous system. Prog. Neurobiol. 65, 289–308 (2001).

43. Grosso, A. R. et al. Tissue-specific splicing factor gene expression signatures. Nucleic Acids Res. 36, 4823–4832 (2008).

44. Häcker, I. et al. Localization of Prp8, Brr2, Snu114 and U4/U6 proteins in the yeast tri-snRNP by electron microscopy. Nat. Struct. Mol. Biol. 15, 1206–1212 (2008).

45. Hanahan, D. & Weinberg, R. A. Hallmarks of cancer: The next generation. Cell 144, 646–674 (2011).

46. Haynes, C. M., Petrova, K., Benedetti, C., Yang, Y. & Ron, D. ClpP Mediates Activation of a Mitochondrial Unfolded Protein Response in C. elegans. Dev. Cell 13, 467–480 (2007).

47. Hirose, Y. & Manley, J. L. RNA polymerase II and the integration of nuclear events. Genes Dev. 14, 1415–1429 (2000).

48. Ho, L. & Crabtree, G. R. Chromatin remodelling during development. Nature 463, 474–484 (2010).

49. Hochstrasser, M. Evolution and function of ubiquitin-like protein-conjugation systems. 2, 153–157 (2000).

50. Hoyos-Manchado, R. et al. RNA metabolism is the primary target of formamide in vivo. Sci. Rep. 7, 1–16 (2017).

51. Hwang, J. & Espenshade, P. J. Proximity-dependent biotin labelling in yeast using the engineered ascorbate peroxidase APEX2. Biochem. J. 473, 2463–2469 (2016).

52. Inada, M. & Pleiss, J. A. Genome-wide approaches to monitor pre-mrna splicing. Methods in Enzymology 470, (Elsevier Inc., 2010).

53. Isaacs, J. S. et al. HIF overexpression correlates with biallelic loss of fumarate hydratase in renal cancer: Novel role of fumarate in regulation of HIF stability. Cancer Cell 8, 143–153 (2005).

54. Janke, C. *et al.* A versatile toolbox for PCR-based tagging of yeast genes: New fluorescent proteins, more markers and promoter substitution cassettes. *Yeast* **21**, 947–962 (2004).

55. Jentsch, S. & Pyrowolakis, G. Ubiquitin and its kin: How close are the family ties? Trends Cell Biol. 10, 335–342 (2000).

56. Kanopka, A., Muhlemann, O. & Akusjarvi, G. Inhibition by SR proteins splicing of a regulated adenovirus pre-mRNA. Nature 381, 535–538 (1996).

57. Kantham, L. et al. Beacon interacts with cdc2/cdc28-like kinases. Biochem. Biophys. Res. Commun. 304, 125–129 (2003).

58. Karaduman, R., Chanarat, S., Pfander, B. & Jentsch, S. Error-Prone Splicing Controlled by the Ubiquitin Relative Hub1. Mol. Cell 67, 423-432.e4 (2017).

59. Karijolich. J & Yu. Y.T Spliceosomal snRNA modifications and their function. RNA Biol. 7, 192–204 (2010).

60. Keren, H., Lev-Maor, G. & Ast, G. Alternative splicing and evolution: Diversification, exon definition and function. Nat. Rev. Genet. 11, 345–355 (2010).

61. Kooi, C. W. Vander et al. Interaction Region Essential for its Function. structure 18, 584–593 (2011).

62. Kornblihtt, A. R. et al. Alternative splicing: A pivotal step between eukaryotic transcription and translation. Nat. Rev. Mol. Cell Biol. 14, 153–165 (2013).

63. Krecic, A. M. & Swanson, M. S. hnRNP complexes: Composition, structure, and function. Curr. Opin. Cell Biol. 11, 363–371 (1999).

64. Leshets, M., Ramamurthy, D., Lisby, M., Lehming, N. & Pines, O. Fumarase is involved in DNA double-strand break resection through a functional interaction with Sae2. Curr. Genet. 64, 697–712 (2018).

65. Li, Y. et al. The SNW Domain of SKIP Is Required for Its Integration into the Spliceosome and Its Interaction with the Paf1 Complex in Arabidopsis. Mol. Plant 9, 1040–1050 (2016).

66. Lopez, A. J. ALTERNATIVE SPLICING OF PRE-mRNA: Developmental Consequences and Mechanisms of Regulation. Annu. Rev. Genet. 32, 279–305 (1998).

67. Luco, R. F., Allo, M., Schor, I. E., Kornblihtt, A. R. & Misteli, T. Epigenetics in alternative pre-mRNA splicing. Cell 144, 16–26 (2011).

68. Lüders, J., Pyrowolakis, G. & Jentsch, S. The ubiquitin-like protein HUB1 forms SDS-resistant complexes with cellular proteins in the absence of ATP. EMBO Rep. 4, 1169–1174 (2003).

69. M.A., G.-B., A.P., B. & E.L., L. Alternative splicing in disease and therapy. Nat. Biotechnol. 22, 535–546 (2004).

70. Makarova, O. V. et al. A subset of human 35S U5 proteins, including Prp19, function prior to catalytic step 1 of splicing. EMBO J. 23, 2381–2391 (2004).

71. Maniatis, T. & Reed, R. An extensive network of coupling among gene expression machines. Nature 416, 499–506 (2002).

72. Maroney, P. A., Hannon, G. J., Denker, J. A. & Nilsen, T. W. The nematode spliced leader RNA participates in trans-splicing as an Sm snRNP. EMBO J. 9, 3667–3673 (1990).

73. Maslon, M. M. et al. A slow transcription rate causes embryonic lethality and perturbs kinetic coupling of neuronal genes. EMBO J. 38, 1–18 (2019).

74. Maudlin, I. E. & Beggs, J. D. Spt5 modulates co-transcriptional spliceosome assembly in Saccharomyces cerevisiae. Rna rna.070425.119 (2019).

75. McNally, T. et al. Structural analysis of UBL5, a novel ubiquitin-like modifier. Protein Sci. 12, 1562–1566 (2003).

76. Millhouse, S. & Manley, J. L. The C-Terminal Domain of RNA Polymerase II Functions as a Phosphorylation-Dependent Splicing Activator in a Heterologous Protein. Mol. Cell. Biol. 25, 533–544 (2005).

77. Mishra, S. K. & Thakran, P. Intron specificity in pre-mRNA splicing. Curr. Genet. 64, 777–784 (2018).

78. Mishra, S. K. et al. Role of the ubiquitin-like protein Hub1 in splice-site usage and alternative splicing. Nature 474, 173–178 (2011).

79. Moore, M. J. & Sharp, P.A. Evidence for two active sites in the spliceosome provided by stereochemistry of pre-mRNA splicing. Nature 365, (1993).

80. Moore, M. J., Schwartzfarb, E. M., Silver, P. A. A. & Yu, M. C. Differential Recruitment of the Splicing Machinery during Transcription Predicts Genome-Wide Patterns of mRNA Splicing. Mol. Cell 24, 903–915 (2006).

81. Mukhopadhyay, D. & Dasso, M. Modification in reverse: the SUMO proteases. Trends Biochem. Sci. 32, 286–295 (2007).

82. Müller, S., Hoege, C., Pyrowolakis, G. & Jentsch, S. Sumo, ubiquitin's mysterious cousin. Nat. Rev. Mol. Cell Biol. 2, 202–210 (2001).

83. Nakai, Y., Nakai, M. & Hayashi, H. Thio-modification of yeast cytosolic tRNA requires a ubiquitin-related system that resembles bacterial sulfur transfer systems. J. Biol. Chem. 283, 27469–27476 (2008).

84. Ngo, J. C. K. et al. Interplay between SRPK and Clk/Sty kinases in phosphorylation of the splicing factor ASF/SF2 is regulated by a docking motif in ASF/SF2. Mol. Cell 20, 77–89 (2005).

85. Nilsen, T. W. & Graveley, B. R. Expansion of the eukaryotic proteome by alternative splicing. Nature 463, 457–463 (2010).

86. Nilsen, T. W. RNA-RNA interactions in the spliceosome: Unraveling the ties that bind. Cell 78, 1–4 (1994).

87. Nilsen, T. W. The spliceosome: The most complex macromolecular machine in the cell? BioEssays 25, 1147–1149 (2003).

88. Oka, Y. et al. UBL5 is essential for pre- mRNA splicing and sister chromatid cohesion in human cells. EMBO Rep. 15, 1330–1330 (2014).

89. Oka, Y., Bekker-Jensen, S. & Mailand, N. Ubiquitin-like protein UBL 5 promotes the functional integrity of the Fanconi anemia pathway . EMBO J. 34, 1385–1398 (2015).

90. Palusa, S. G., Ali, G. S. & Reddy, A. S. N. Alternative splicing of pre-mRNAs of Arabidopsis serine/arginine-rich proteins: Regulation by hormones and stresses. Plant J. 49, 1091–1107 (2007).

91. Pan, Q., Shai, O., Lee, L. J., Frey, B. J. & Blencowe, B. J. Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. Nat. Genet. 40, 1413–1415 (2008).

92. Patel, M. et al. Overexpression of ubiquitin-like LpHUB1 gene confers drought tolerance in perennial ryegrass. Plant Biotechnol. J. 13, 689–699 (2015).

93. Pick, M., Flores-Flores, C. & Soreq, H. From brain to blood: Alternative splicing evidence for the cholinergic basis of mammalian stress responses. Ann. N. Y. Acad. Sci. 1018, 85–98 (2004).

94. Pollard, P. J. et al. Accumulation of Krebs cycle intermediates and over-expression of HIF1α in tumours which result from germline FH and SDH mutations. Hum. Mol. Genet. 14, 2231–2239 (2005).

95. Proudfoot, J. N. Ending the message: poly(A) signals then and now. Genes Dev. 25, 1770–1782 (2011).

96. Rabut, G. & Peter, M. Function and regulation of protein neddylation. 'Protein modifications: beyond the usual suspects' review series. EMBO Rep. 9, 969–976 (2008).

97. Rajkovic, A., Davis, R. E., Simonsen, J. N. & Rottman, F. M. A spliced leader is present on a subset of mRNAs from the human parasite Schistosoma mansoni. Proc. Natl. Acad. Sci. U. S. A. 87, 8879–8883 (1990).

98. Ramani, A. K. et al. Genome-wide analysis of alternative splicing in Caenorhabditis elegans. Genome Res. 21, 342–348 (2011).

99. Ramelot, T. A. et al. Solution structure of the yeast ubiquitin-like modifier protein Hub1. J. Struct. Funct. Genomics 4, 25–30 (2003).

100. Reyes-Turcu, F. E., Ventii, K. H. & Wilkinson, K. D. Regulation and Cellular Roles of Ubiquitin-Specific Deubiquitinating Enzymes. Annu. Rev. Biochem. 78, 363–397 (2009).

101. Romfo, C. M., Maroney, P. A., Wu, S. & Nilsen, T. W. 3′ splice site recognition in nematode trans-splicing involves enhancer-dependent recruitment of U2 snRNP. Rna 7, 785–792 (2001).

102. Roscigno, R. F. & Garcia-Blanco, M. A. SR proteins escort the tri-snRNP.pdf. Rna 1, 692–706 (1995).

103. Sakharkar, M. K., Kangueane, P. & Dmitri, A. Genes from Eukaryotes. Bioinformatics 18, 1266–1267 (2002).

104. Sanford, J. R. & Bruzik, J. P. SR proteins are required for nematode trans-splicing in vitro. Rna 5, 918–928 (1999).

105. Schneider, C., Will, C. L., Makarova, O. V., Makarov, E. M. & Luhrmann, R. Human U4/U6. U5 and U4atac/U6atac.U5 Tri-snRNPs Exhibit Similar Protein Compositions. Mol. Cell. Biol. 22, 3219–3229 (2002).

106. Selak, M. A. et al. Succinate links TCA cycle dysfunction to oncogenesis by inhibiting HIF-α prolyl hydroxylase. Cancer Cell 7, 77–85 (2005).

107. Sgorbissa, A. & Brancolini, C. IFNs, ISGylation and cancer: Cui prodest? Cytokine Growth Factor Rev. 23, 307–314 (2012).

108. Shen, H. & Green, M. R. RS domain-splicing signal interactions in splicing of U12-type and U2-type introns. Nat. Struct. Mol. Biol. 14, 597–603 (2007).

109. Shen, H. & Green, M. R. RS domain-splicing signal interactions in splicing of U12-type and U2-type introns. Nat. Struct. Mol. Biol. 14, 597–603 (2007).

110. Shepard, P. J. & Hertel, K. J. The SR protein family. Genome Biol. 10, 242 (2009).

111. Simonis, N. et al. Empirically controlled mapping of the Caenorhabditis elegans protein-protein interactome network. Nat. Methods 6, 47–54 (2009).

112. Singer, E., Silas, Y. B. H., Ben-Yehuda, S. & Pines, O. Bacterial fumarase and L-malic acid are evolutionary ancient components of the DNA damage response. Elife 6, 1–18 (2017).

113. Soller, M. Pre-messenger RNA processing and its regulation: A genomic perspective. Cell. Mol. Life Sci. 63, 796–819 (2006).

114. Song, E. J. et al. Deubiquitinating Enzyme Control Reversible Ubiquitination at the Spliceosome. Genes Dev. 24, 1434–1447 (2010).

115. Staley, J.P & Woolford Jr, J. L. Assembly of ribosomes and spliceosomes: complex ribonucleoprotein machines. Curr. Opin. Cell Biol. 46, 220–231 (2009).

116. Švéda, M., Častorálová, M., Lipov, J., Ruml, T. & Knejzlík, Z. Human UBL5 protein interacts with coilin and meets the Cajal bodies. Biochem. Biophys. Res. Commun. 436, 240–245 (2013).

117. Taherbhoy, A. M., Schulman, B. A. & Kaiser, S. E. Ubiquitin-like modifiers. Essays Biochem. 52, 51–63 (2012).

118. Tessier, L. et al. RNAs to pre-mature mRNAs by trans-splicing in Euglena. EMBO J. 10, 2621–2625 (1991).

119. Thakran, P. et al. Sde2 is an intron-specific pre-mRNA splicing regulator activated by ubiquitin-like processing. EMBO J. 37, 89–101 (2018).

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

121. Valcárcel, J. & Green, M. R. The SR protein family: Pleiotropic functions in pre-mRNA splicing. Trends Biochem. Sci. 21, 296–301 (1996).

122. Wahl, M. C., Will, C. L. & Lührmann, R. The Spliceosome: Design Principles of a Dynamic RNP Machine. Cell 136, 701–718 (2009).

123. WANG, Y. et al. Mechanism of alternative splicing and its regulation. Biomed. Reports 3, 152–158 (2015).

124. Waterhouse, A. M., Procter, J. B., Martin, D. M. A., Clamp, M. & Barton, G. J. Jalview Version 2-A multiple sequence alignment editor and analysis workbench. Bioinformatics 25, 1189–1191 (2009).

125. Weiner, A. M. mRNA splicing and autocatalytic introns: Distant cousins or the products of chemical determinism? Cell 72, 161–164 (1993).

126. Wilkinson, C. R. M. et al. Ubiquitin-like Protein Hub1 Is Required for Pre-mRNA Splicing and Localization of an Essential Splicing Factor in Fission Yeast. 14, 2283–2288 (2004).

127. Will, C. & Lührmann, R. Spliceosome structure and function. Cold Spring Harb Perspect Biol doi: 10.1101/cshperspect. a003707. Cold Spring Harb Perspect Biol 3, (2011).

128. Wise, J. A. Guides to the heart of the spliceosome. Science 262, 1978–1979 (1993).

129. Yashiroda, H. & Tanaka, K. Hub1 is an essential ubiquitin-like protein without functioning as a typical modifier in fission yeast. Genes to cells 9, 1189–1197 (2004).

130. Yogev, O. et al. Fumarase: A mitochondrial metabolic enzyme and a cytosolic/nuclear component of the dna damage response. PLoS Biol. 8, (2010).

131. Zhou, Z. & Fu, X.-D. Regulation of Splicing by SR proteins. Chromosoma 122, 191–207 (2013).

132. Zhu, J., Mayeda, A. & Krainer, A. Exon identity established through differential antagonism between exonic splicing silencer-bound hnRNP A1 and enhancer-bound SR proteins. Mol. Cell 8, 1351–1361 (2001).

133. Zlotorynski, E. Splicing: Phasing alternative exons. Nat. Rev. Mol. Cell Biol. 18, 529 (2017).

134. Zuo, P. & Manley, J. L. Functional domains of the human splicing factor ASF/SF2 the human splicing factor ASF/SF2 displays two predominant activities in in vitro splicing. EMBO J. 12, 4727–4737 (1993).

Publication

1. Thakran, P*., Pandit, P.A*., Datta, S., **Kolathur, K.K.,** Pleiss, J.A., and Mishra, S.K. (2018). *Sde2 is an intron-specific pre-mRNA splicing regulator activated by ubiquitin-like processing*. **The EMBO Journal** 37(1): 89-101.