Influence of spliceosome proteins on RNA interactions in pre-mRNA splicing

By

Karli A. Lipinski

A dissertation submitted in partial fulfillment of

the requirements for the degree of

Doctor of Philosophy

(Chemistry)

at the

UNIVERSITY OF WISCONSIN-MADISON

2024

Date of final oral examination: 5/8/2024

The dissertation is approved by the following members of the Final Oral Committee:

Aaron A. Hoskins, Professor, Biochemistry

David A. Brow, Professor, Biomolecular Chemistry

Andrew R. Buller, Assistant Professor, Chemistry

Tina Wang, Assistant Professor, Chemistry

ABSTRACT

Pre-mRNA splicing is a fundamental process governing gene expression in eukaryotes and is orchestrated by the spliceosome, a dynamic ribonucleoprotein complex. Spliceosomes catalyze splicing in two steps and are composed of both small nuclear RNAs (snRNAs; U1, U2, U4, U5, U6) and numerous proteins. Dynamic interactions between snRNAs, proteins, and the pre-mRNA substrate occur during splicing which are central to regulating splice site recognition, catalysis, and fidelity. This thesis examines multiple facets of RNA-protein interactions within spliceosome complexes beginning with snRNA biogenesis, through formation of the spliceosome active site, and during the second step of chemistry, exon ligation. Functional mechanisms into splicing, such as selection of intron recognition sites, are examined.

Function of the main catalytic component of the spliceosome, the U6 snRNA, is first explored. Transcription and post-transcriptional processing of U6 is unique among snRNAs and therefore may be required to generate a functional U6. Transcription by RNAP II instead of RNAP III produces a functional U6 molecule. Defects in stability, likely a result of incorrect post-transcriptional processing and binding of stabilizing proteins, result in changes to in vivo distributions of spliceosome sub particles called snRNPs. Transcription of U6 by RNAP II is useful for the incorporation of genetic tags for endogenous fluorescent labeling or purification.

Additionally, new applications of endogenous fluorescent labeling techniques within the U4 snRNA are presented, paving the way for single molecule studies of snRNA dynamics and Brr2 helicase function during activation. Mango and MS2 tags are well tolerated in the yeast U4 snRNA and with several of the tagged U4 constructs also minimally impacting splicing activities. Future single molecule experiments will examine the timing of the U4 snRNA release from the spliceosome compared to the release of Prp3, a protein associated with U6 and U4 snRNAs.

Finally, validation of a proposed novel splicing factor, Fyv6, and study of its influence on 3′ SS represent major contributions to the field of splicing. Utilizing a new high resolution P complex structure of the spliceosome containing Fyv6 solved by Max Wilkinson, I examined multiple contacts of Fyv6 with other splicing factors, notably Prp22. Genetic studies in yeast show that interactions with the protein Syf1 and Prp22 are important for Fyv6 function. The absence of Fyv6 from spliceosomes results in transcriptome-wide splicing defects, largely due to changes in 3′ SS usage. From these studies, Fyv6 can be added to the list of splicing factors that impact the second step of splicing and affect fidelity of 3′ SS selection.

ACKNOWLEDGEMENTS

I would like to extend my deepest appreciation to several individuals who have been instrumental in my laboratory experiences. Firstly, I am profoundly grateful to Aaron Hoskins for his invaluable mentorship during my graduate studies. The lessons he imparted to me will undoubtedly shape the rigor and quality of my work throughout my career. I am particularly thankful for his unwavering support as I explored new techniques, as well as for the depth and thoughtfulness of the feedback he provided on my work. I also owe a debt of gratitude to Katherine Senn, who was my close scientific partner for a significant portion of my graduate work. Her extensive knowledge of literature, precise laboratory skills, and thoughtful approach continue to inspire me. I am deeply appreciative of the undergraduate students I had the privilege to work with and mentor in the laboratory, including Cade Harkner, Natalie Zeps, and Amory Griffin. Their enthusiasm and dedication made my time in the lab both enjoyable and productive. Some of my fondest memories in the laboratory involve collaborating with all of the aforementioned individuals. Special thanks are also due to Justin Mabin for his invaluable assistance in navigating RNA-seq analysis, and to all of my lab mates for their insightful discussions and unwavering support. Additionally, I thank Cheri Barta and Léa Gustin for their guidance in the classroom, mentorship practices, and life advice.

Beyond the laboratory, I am grateful to my parents, Barbara and Jeff Lipinski, who supported my dream of becoming a scientist since I was six years old. Their steadfast belief in me has been a constant source of motivation. I am also thankful to my in-laws, Gabriell and Richard Mills, for embracing me as part of their family and providing support throughout my graduate studies. Lastly, I want to express my deepest gratitude to my spouse, Perry Mills. Your support, encouragement, and unwavering belief in me inspire me every day. Thank you for everything you do and for being the incredible person you are.

Dedicated to my family and friends, whose unwavering belief in me
is uplifting and a continued source of inspiration.

TABLE OF CONTENTS

TABLE OF FIGURES

CONTRIBUTIONS

CHAPTER 1

All figures were made by Karli Lipinski.

CHAPTER 2

Figure 2.7 – Data collected by Xin Chen, analyzed by Karli. Karli made the figure.

Figure 2.8 – Data collected by Karli and analyzed by Karli. Karli made the figure.

Figure 2.9 – Data collected and analyzed by Karli. Karli made the figure.

Figure 2.10 – Data collected and analyzed by Karli. Karli made the figure.

Figure 2.11 – Data collected and analyzed by Karli. Karli made the figure.

Figure 2.12 – Data collected and analyzed by Karli. Karli made the figure.

CHAPTER 3

Karli collected and analyzed all data. All figures were made by Karli.

CHAPTER 4

Figure 4.1D – Karli conducted the growth assay and made the figure.

Figure 4.5 – Karli conducted A1C1 assays and made the figure.

Figure 4.6 – Karli conducted splicing assays, analyzed data, and made the figure.

Figure 4.7 – Karli conducted growth and splicing assays, analyzed data, and made the figure.

Figure 4.8 – Karli conducted primer extension assays, analyzed data, and made the figure.

CHAPTER 5

Figure 5.1 – Kathy made schematic in A. Karli conducted growth assays in B. Karli and Natalie* prepared samples for RNA-seq. Karli analyzed RNA-seq data resulting in C and D. Karli and Amory* analyzed RNA-seq data resulting in E and F. Karli made the figure except for A.

Figure 5.2 – Karli analyzed data and made figures A and B. Karli and Natalie analyzed data and made figure C. Karli conducted RT-PCR in E and made the figure. Karli and Amory analyzed data for F and Karli made the figure.

*Natalie Zeps and Amory Griffin are undergraduate students mentored by Karli Lipinski.

LIST OF ABBREVIATIONS

3′ SS ..... 3′ splice site
5′ SS ..... 5′ splice site
5FOA ..... 5-Fluoroorotic acid
A5SS ..... alternate 5' SS
A3SS ..... alternate 3' SS
BP ..... branch point
CoSMoS ..... colocalization single molecule spectroscopy
cs ..... cold sensitivity
cryo-EM ..... cryo-electron microscopy
FAnS ..... fraction of annotated splicing
FRET ..... Förster resonance energy transfer
fyv6Δ ..... FYV6 deletion
gal ..... galactose
GFP ..... green fluoresent protein
gluc ..... glucose
ILS ..... intron lariat spliceosome
IP ..... immunoprecipitation
IR ..... intron retention
ISL ..... internal stem loop
mAb ..... monoclonal antibody
MCP ..... MS2 coat protein
MMG ..... monomethylguanosine
mRNA ..... messenger RNA
NLS ..... nuclear localization signal
NMD ..... nonsense mediated decay
NNS Nrd1–Nab3–Sen1
nt ..... nucleotide
NTC ..... ninteen complex
NTR ..... NTC-related factors
OD600 ..... optical density at 600 nm
PCR ..... polymerase chain reaction
PPT ..... polypyrimidine tract
PSI ..... percent spliced in
PTC ..... premature stop codon
raff ..... raffinose
RCP ..... Rnt1 cleavage product
RNAP ..... RNA polymerase
RT ..... reverse transcriptase

SD ..... standard deviation
SL ..... stem loop
SNR14 ..... U4 snRNA gene
SNR6 ..... U6 snRNA gene
snRNP ..... small nuclear ribonucleoprotein
syf1Δ ..... *SYF1 deletion*
TMG 2,2,7-trimethylguanosine
TO1 ..... thiazole orange
ts ..... temperature sensitivity
U6-II ..... U6 transcribed by RNAP II
UAS ..... upstream activating sequence
upf1Δ ..... *UPF1 deletion*
(y)WCE ..... (yeast) whole cell extract
Ψ-WT ..... pseudo-wild type

CHAPTER 1

Introduction

CHAPTER 1: Introduction

1.1 Splicing is essential for proper gene expression in eukaryotes

Most transcripts produced by RNA polymerase II (RNAP II) in eukaryotes contain intervening sequences called introns that interrupt protein coding segments. Removal of introns and subsequent joining of the protein-coding segments called exons is required at nucleotide precision for in-frame protein translation. Intron removal and exon ligation is known as pre-mRNA splicing and is essential for eukaryotic gene expression. In the yeast *Saccharomyces cerevisiae*, there are ~280 genes that contain a single intron out of ~5,000 protein coding genes (Kuang et al. 2017). Humans have approximately ~25,000 protein coding genes which contain an average of 8 introns per gene (Hnilicova and Stanek, 2011). Exons in human pre-mRNAs have an average length of ~145 bp (Hnilicova and Stanek, 2011), while introns are either short at under 100 nt or broadly distributed with most around several thousand nucleotides extending up to over a million nucleotides (Yu et al. 2022; Lim et al. 2001). In contrast, yeast have shorter intron sizes with intron lengths cluster around either 100 bp or 400 bp (Parker and Patterson, 1987; Rymond and Rosbash, 1992). In organisms with many and long introns, identifying exons sequences from intron sequences presents an additional challenge.

Exons within a gene can be joined in a combination of ways, most simply by either inclusion or exclusion, in the final spliced mRNA in a process called alternative splicing. How the introns are removed and the timing of removal allows for regulation of gene expression (Gehring and Roignant, 2021). Whereas some introns are removed constitutively, removal of others is highly regulated (Matlin et al. 2005). Different mRNA isoforms that result from alternative splicing increase the quantity of in-frame proteins that can be produced from a single gene, thus diversifying in the proteome. In humans, approximately 95% of genes are alternatively spliced, further adding to the complexity of splicing in humans (Ule and Blencowe, 2019). However, the intron-poor yeast *S. cerevisiae* presents an opportunity for characterization of pre-mRNA splicing mechanisms within a comparatively less complicated model system.

1.2 Splicing occurs in two transesterification steps

Several short, conserved sequences allow for recognition of intron boundaries: the 5′ splice site (5′ SS), the branch point (BP), and the 3′ splice site (3′ SS). Both the 5′ SS and BP are highly conserved in yeast with the consensus sequences GUAUGU and UACUA**A**C (BP adenosine in bold; **Fig. 1.1A**) (Spingola et al. 1999; Qin et al. 2016). In contrast, the 3′ SS has a much shorter consensus sequence of YAG (Y = C or U). Human introns contain degenerate 5′ SS and BP sequences compared to yeast. Sequences preceding the 3′ SS that are rich in pyrimidines (polypyrimidine tract; PPT) also help with

intron recognition although PPT sequences are much weaker in yeast (Sheth et al. 2006; Abiovich et al. 1994; Patterson and Guthrie, 1991). These sequences are recognized multiple times during splicing to maintain splicing fidelity.

Introns are removed from pre-mRNA in two sequential transesterification reactions (branching and exon ligation) that are catalyzed in a single active site within the spliceosome, a large protein and RNA complex (**Fig, 1.1B**) (Padgett et al. 1984; Ruskin et al. 1984; Rodriguez et al. 1984; Domdey et al. 1984; Brody et al. 1985). During the first step, the BP and 5′ SS are positioned within the active site and the 2′ hydroxyl group of the adenosine at the BP attacks the phosphodiester group at the 5′ SS. Two intermediate products are formed after this step, a cleaved 5′ exon and a lariat-3′-exon intermediate. In the lariat intermediate, the 5′ phosphate of the first intron nucleotide is linked to the 2′ oxygen of the adenosine at the BP. In the second step, the 5′ exon remains in the active site, but the BP adenosine is removed to allow for docking of the 3′ SS. The 3′ hydroxyl group of the 5′ exon attacks the phosphodiester group at the 3′ SS, resulting in ligated exons that form mRNA and an intron lariat.

Within the spliceosome active site, two metal ions stabilize the pentacovalent transition states of the splicing transesterification reactions (Steitz et al. 1993). During the first step, one metal ion stabilizes the 3′ hydroxyl of the last nucleotide in the 5′ exon (leaving group) while the second metal ion activates the BP adenosine 2′ hydroxyl group (attacking nucleophile). The BP adenosine must leave the active site before the 3′ SS is bound. During the second step, one metal ion activates the 3′ hydroxyl of the 5′ exon while the second stabilizes the 3′ hydroxyl group of the last nucleotide within the intron. RNAs assembled within the spliceosome coordinate both metal ions for catalysis, making the spliceosome a ribozyme (Fica et al. 2013). Both steps of chemistry catalyzed by the spliceosome are reversible (Tseng and Cheng, 2008).

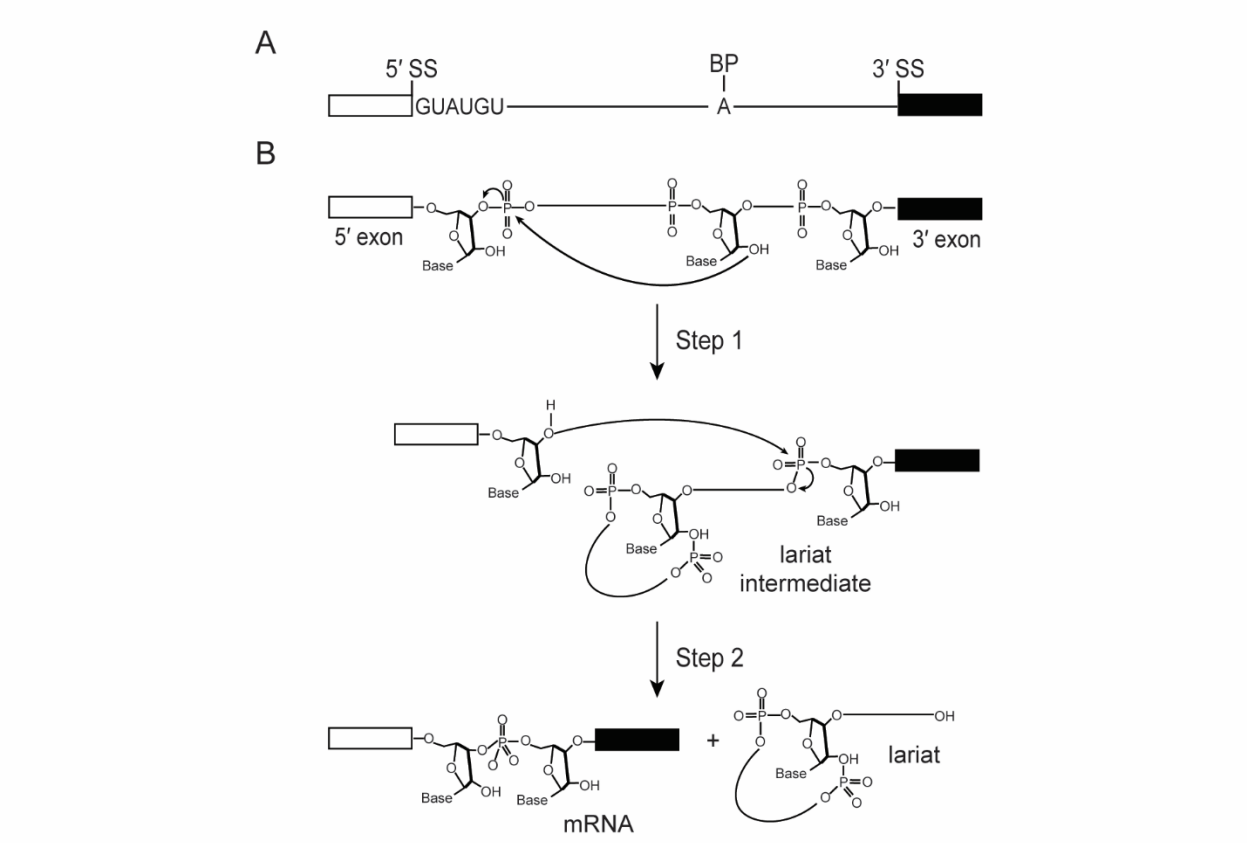

Figure 1.1 Consensus sequences within *S. cerevisiae* pre-mRNAs and splicing mechanism (adapted from Woll et al. 2017).

1.3 Spliceosome are complex macromolecular machines that catalyze splicing

The spliceosome is composed of dozens of proteins (over 170 protein factors in humans and ~90 in *Saccharomyces cerevisiae*) and five small nuclear RNAs (U1, U2, U4, U5, U6; snRNAs) (Will and Lührmann, 2011; Kaur et al, 2022). Recognition of substrate consensus sequences and catalysis are performed by snRNAs during splicing. Each snRNA binds a specific set of proteins to form a spliceosome subunit called a small nuclear ribonucleoprotein (snRNP). Spliceosomes also contain a large, protein-only subunit called the Nineteen complex (NTC) and several non-snRNP associated proteins, including splicing factors and at least eight DExD/H-box ATPases.

For each intron within a pre-mRNA, spliceosomes must assemble do novo from snRNPs. Spliceosome assembly occurs in an ordered and defined manner called the splicing cycle (**Fig. 1.2**). Each step within the splicing cycle represents a stalled spliceosome intermediate that has been characterized through biochemical, genetic, and/or structural experiments. Transitions between complexes are facilitated by ATP hydrolysis of one of eight DExD/H-box ATPases that are recruited transiently to the

spliceosome independently of snRNPs. They mediate remodeling of RNA-protein interactions including the removal or repositioning of protein and RNA components during splicing.

Spliceosome assembly begins with binding of the U1 and U2 snRNPs to the 5′ SS and BP, respectively, within an intron as it is transcribed. Recognition of these two consensus sequences occurs via base pairing of the snRNAs and is further aided by associated proteins within each snRNP. Next, the remainder of the snRNPs join the spliceosome together as a preassembled particle called the tri-snRNP. During spliceosome activation, rearrangements occur to position the 5′ SS within the active site adjacent to the BP for the first transesterification reaction. The U1 and U4 snRNPs are released during activation, and the NTC arrives, marking the largest compositional change during the splicing cycle where dozens of proteins are exchanged. With the release of U4, U6 undergoes a conformational change to form the catalytic center of the spliceosome and bring the 5′ SS into the active site.

After completion of the first step, the spliceosome undergoes structural rearrangements to remove the lariat intermediate from the active site including the release of first step splicing factors. The free 5′ exon and 3′ SS are repositioned within the active site aided by second step splicing factors. Unlike recognition of the 5′ SS and BP, recognition of the 3′ SS occurs only via protein interactions with the pre-mRNA. After completion of the second step, the mRNA is released by Prp22 ATP hydrolysis, and the intron-lariat spliceosome (ILS) is disassembled by helicase activity of Prp43. Splicing components are reassembled into snRNPs and recycled for another round of splicing.

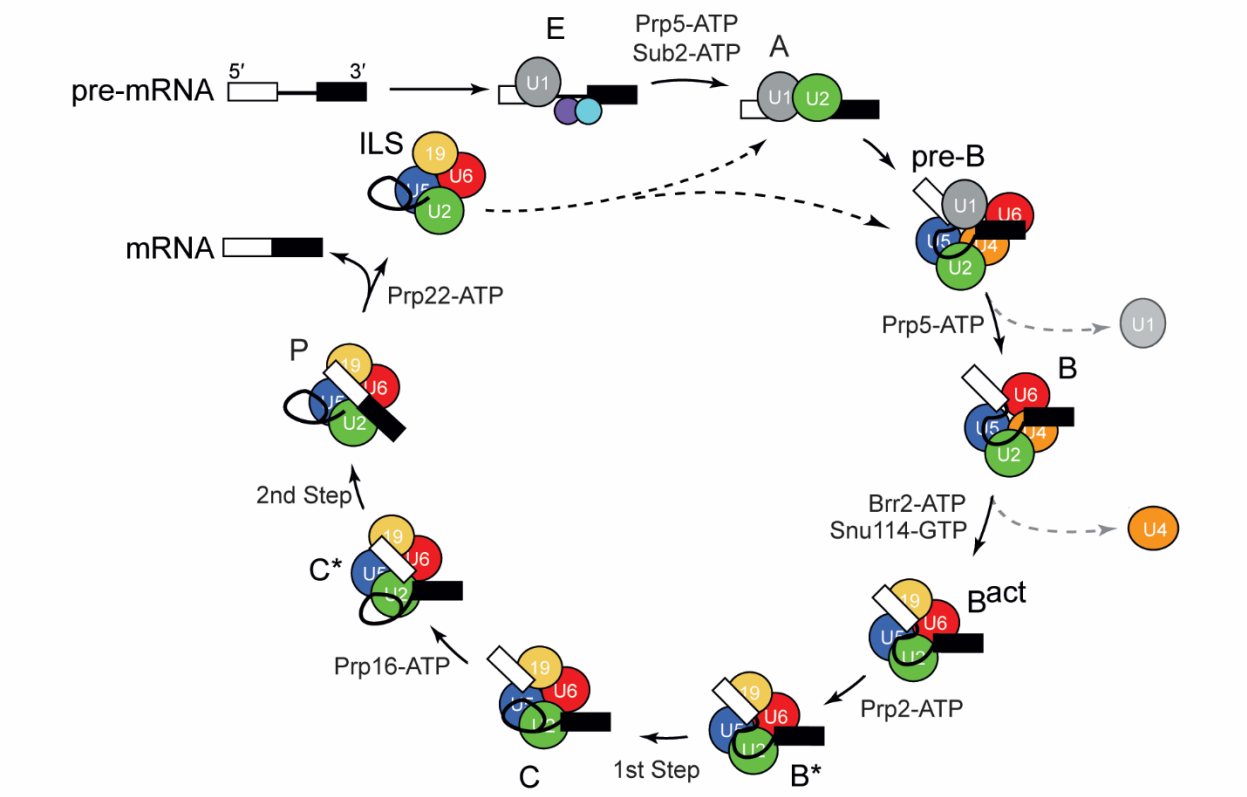

Fig. 1.2 Spliceosomes assemble in sub particles called snRNPs in a sequential manner on pre-mRNA substrates. snRNPs are modeled as colored balls assembled in stalled spliceosome intermediates. ATPases that drive the splicing reaction forward with ATP hydrolysis are listed between the intermediates where they act. Splicing cycle as adapted from Hoskins et. al, Trends in Biochemical Sciences, 2012.

1.4 Alternative splicing patterns cause disease in humans

Alternative splicing of pre-mRNAs that produces at least two unique transcript isoforms occurs in ~90% of human genes (Wang et al. 2008). In many cases, these unique isoforms produce different protein variants as confirmed by mass spectroscopy thereby diversifying the proteome (Sinitcyn et al. 2023; Nilsen et al. 2010). Production of alternative transcript isoforms arises from changes in splice site recognition by the spliceosome machinery itself or its regulators (Love et al. 2023). Alternative inclusion or exclusion of exons, recognition of different 5 and 3′ SS, and use of different transcription star sites or polyadenylation sites can result in the production of diverse mRNA products (**Fig. 1.3**). However, some alternatively spliced isoforms are not translated due to changes in the reading frame that result in a premature stop codon (PTC). Spliced mRNAs that contain PTCs are recognized by Upf1 ATPase as part of the nonsense mediated decay pathway (NMD) (Kurosaki et al. 2019). At least 15% of human hereditary diseases and cancers are associated with alterative splicing (Marquez et al. 2012; Cui et al. 2017). The most commonly seen alternative splicing events are failure to remove an intron (intron

retention) or failure to recognize an exon (exon skipping) (Matlin et al. 2005). Alternative RNA splicing patterns are generally identified bioinformatically through application of statistical algorithms, like rMATS, to RNA-sequencing datasets (Shen et al. 2014).

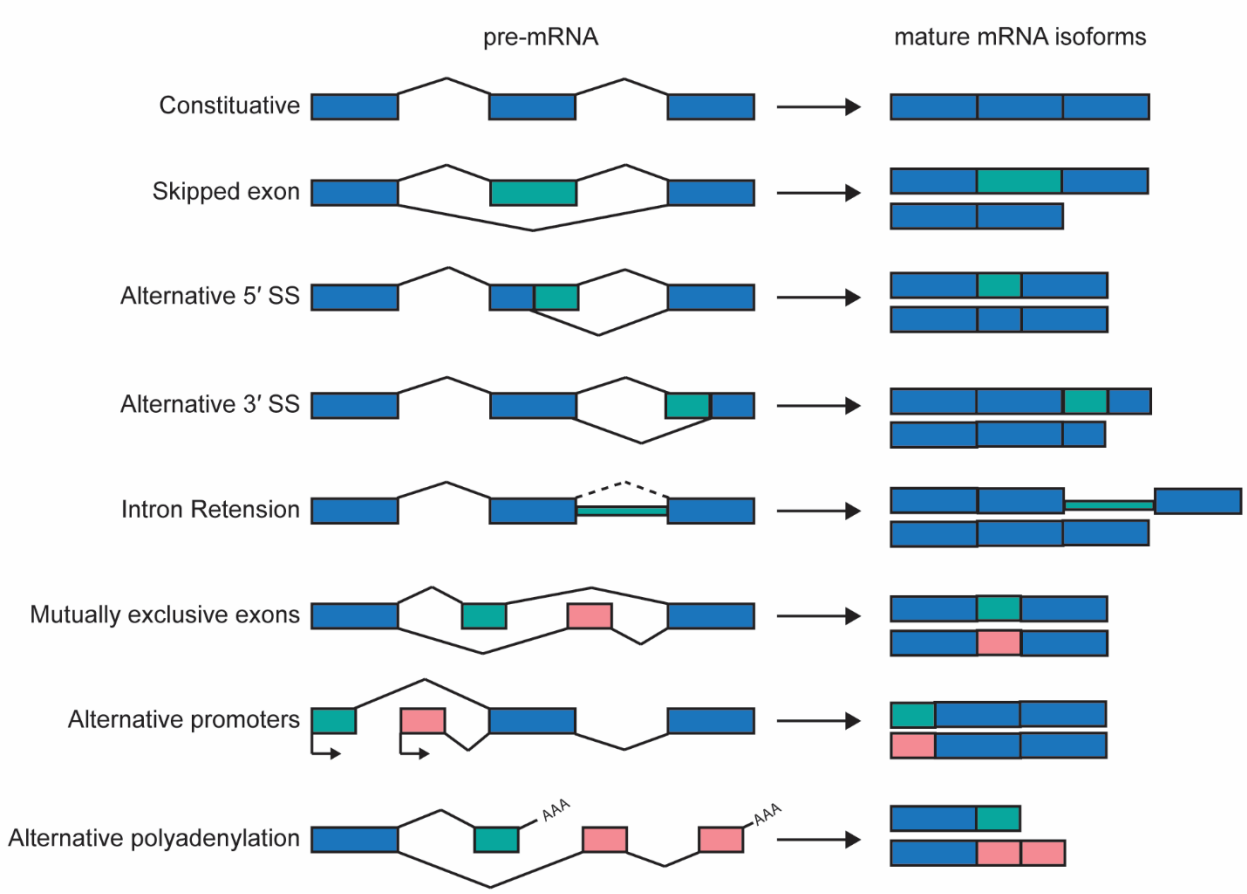

Figure 1.3 Isoforms produced through multiple alternative splicing pathways. Alternative segments are shown in green and pink. Constitutive segments shown in blue. Adapted from Bhadra et al. 2020.

1.5 snRNAs undergo post-transcriptional processing before snRNP assembly

All snRNAs are extensively processed post-transcriptionally wherein they are transcribed, cleaved, trimmed, exported, reimported, bound by proteins, and modified. The snRNAs U1, U2, U4, and U5 share a similar maturation pathway whereas U6 has a distinct maturation pathway (**Figure 1.4**). The U6 snRNA is the most conserved snRNA as it is central to splicing catalysis, therefore it is maybe unsurprising that U6 is processed distinctly.

Beginning with transcription, U6 is transcribed by RNAP III while the other four snRNAs are transcribed by RNAP II (Reddy et al. 1987; Moenne et al. 1990). The expression level of U6 varies across tissues (Spaniel et al. 2013), yet regulation of transcriptional activity or its effect on splicing is poorly understood. Whether transcription

of U6 by RNAP III allows for tighter regulation of U6 levels in the cell compared to RNAP II or helps to facilitate post-transcriptional processing is unknown. Termination of RNAP III on the U6 gene occurs at a stretch of 10 dAs, creating a heterogeneous tail length of 4-7 uridines (Brow and Guthrie, 1990). Newly transcribed U6 molecules are bound by Lhp1 (La in humans) following RNAP III release (Rinke and Steitz 1985; Pannone et al. 1998).

The heterologous end of U6 in yeast is shortened during post-transcriptional processing in yeast to produce a mature U6 with an end of uniform length. Usb1 is a 3′ to 5′ exoribonuclease that trims the oligo(U) tail, leaving a phosphate group on the terminal uridine (Lund and Dahlberg 1992; Didychuk et al. 2017). The timing of Usb1 modification is unclear but may happen coincident with splicing (Tazi et al. 1993). Trimming of the oligo(U) tail reduces affinity of La for binding U6 and enhances binding of the Lsm2-8 ring (Terns et al. 1992; Licht et al. 2008).

RNAP II transcribed snRNAs are 3′ end cleaved by Rnt1 coincident with extrusion of the cleavage site from RNAP II, thus releasing the snRNA from the polymerase (Chanfreau et al. 1998; Braglia et al. 2011). Also thought to occur cotranscriptionally as with other RNAP II transcribed RNAs, the cap is modified to a m^7G cap and is bound by the cap binding complex (Köhler and Hurt, 2007, Matera and Wang, 2014). RNA export complexes (Mex67 and Xpo1/Crm1) assemble onto the immature snRNA and facilitate nuclear export (Becker et al. 2019). Once in the cytoplasm, snRNAs undergo several maturation steps before being reimported to the nucleus in a process known as shuttling.

It is not known why snRNAs are exported only to be reimported after assembly into stable snRNP particles. The sequestration of immature RNPs in destinations remote from their substrates is not unique to snRNAs: ribosomal subunits are primarily assembled in the nucleolus, but function in the cytoplasm (Panse and Johnson 2010; Strunk et al. 2011). Cytoplasmic shuttling may provide a mechanism for quality control as immature snRNPs cannot encounter their substrates (Matera and Wang, 2014). Historically, all snRNAs have been shown to shuttle in yeast except for the U6 snRNA, which remains in the nucleus during its entire maturation process (Bertrand and Bordonné, 2004, Matera et al., 2007, Sloan et al., 2016). However, recently it has been proposed that U6 also shuttles to the cytoplasm during maturation via Mex67 (Becker et al. 2019).

snRNP complexes are likely assembled in the cytoplasm, including the piecewise addition of the heteroheptameric Sm ring (Becker et al. 2019; Achsel et al. 1999). The addition of an Sm ring protects the snRNA from degradation by 3′ - 5′ exoribonucleases, providing stability and thus extending the half-life of snRNAs within the cell (Coy et al., 2013, Shukla and Parker, 2014). It was proposed that if U6 shuttles to the cytoplasm it is reimported only after it is assembled with its inactivator, the U4 snRNA, as the U4/U6 di-snRNP (Becker et al. 2019). The U6 snRNA binds an analogous ring to the RNAP II transcribed snRNAs, the Like-Sm (Lsm) ring, as a preassembled complex (Achsel et al. 1999). Lsm rings perform a similar function to Sm rings in stabilization of the U6 snRNA

(Matera et al., 2007). Within the context of U6 shuttling, it is then unclear at what point in maturation the Lsm ring binds, as a bound Lsm ring is also important for nuclear localization and retention of U6 (Spiller et al. 2007a,b). Immature snRNAs are reimported into the nucleus via Mtr10 and Cse1 (Becker et al. 2019). Further processing of RNAP II transcribed snRNAs occurs with cap modification from m^7G to 2,2,7-trimethylguanosine (TMG) by TGS1 in the nucleolus and 3′ end trimming in the nucleus (Becker et al. 2019). Once snRNAs complete the final steps of maturation, they are now ready to join the splicing cycle.

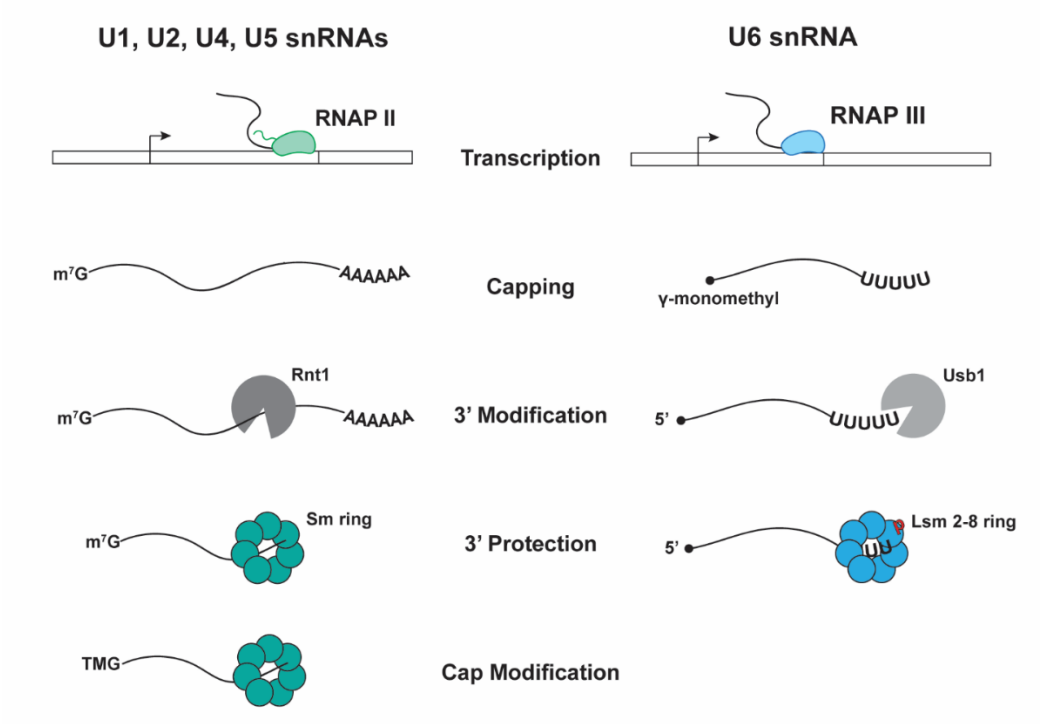

Figure 1.4 Comparison between snRNA biogenesis steps for RNAP II transcribed (U1, U2, U4, and U5) and RNAP III transcribed snRNAs (U6). RNAP II and Sm rings are shown in green while RNAP II and Lsm2-8 rings are shown in blue.

1.6 Brr2 mediated disruption of U4/U6 di-snRNAs during spliceosome activation

Splicing requires ATP hydrolysis by eight conserved RNA dependent ATPases at key points to rearrange spliceosomal RNP interaction networks (Staley and Guthrie, 1998; Cordin et al. 2012; Cordin and Beggs, 2013). Spliceosomal helicases can be grouped into families based on conserved motifs: DEAD box proteins (Prp5, Sub2, Prp28), DEAH box proteins (Prp2, Prp16, Prp22, Prp43), and one Ski2-like helicase (Brr2) (Staley and Gurthrie, 1998). Brr2 is unique among other spliceosomal ATPases as it is a constituent part of the spliceosome instead of transiently interacting. Brr2 hydrolyzes ATP

during activation to facilitate unwinding of the U4/U6 di-snRNA duplex resulting in the release of the U4 snRNA and its associated proteins (Agafonov et al 2011; Theuser et al. 2016). Release of U4 results in a conformational change in the U6 snRNA, which forms the active site and positions substrate in proximity for catalysis (Hang et al. 2015). Whereas most steps along the splicing cycle are reversible, activation represented an irreversible step in splicing where the spliceosome exchanges ~50 factors (Hoskins et al. 2016). Brr2 activity must therefore be tightly regulated to ensure correct timing of ATP hydrolysis during activation.

The substrate for the Brr2 helicase is the U4/U6 di-snRNA. Brr2 first encounters its substrate during tri-snRNP formation when the U4-U6 di-snRNP joins the free U5 snRNP to form the U4/U6.U5 tri-snRNP (Absmeier et al. 2016). U6 is extensively paired to its regulator, the U4 snRNA, prior to activation in both the di-snRNP and tri-snRNP (Stevens and Abelson, 1999). The stem II U4/U6 duplex is formed from nucleotides (nt) 1-17 of the U4 snRNA and nt 64-80 of the U6 snRNA. Between Stems I and II, the U4 snRNA forms a k-turn motif to which the U4 snRNP associated proteins Snu14 and Prp31 bind (Nguyen et al. 2016). Stem I consists of the U4 nt 56-64 and the U6 nt 54-63. The ATPase Brr2 is loaded onto the single-stranded region of U4 between the 5 SL and the 3 SL and translocates 3’ to 5’ on the U4 snRNA to disrupt base pairing interactions with the U6 snRNA (Raghunathan and Guthrie, 1998).

Upon release of U4, U6 forms the internal stem loop (ISL) within the active site coordinating magnesium ions for catalysis (Yean et al. 2000) and forming new RNA contacts (5’ SS and U2 snRNA) to position the BP in proximity to the 5′ SS for first transesterification reaction (Sashital et al. 2004). Competing RNA structures are proposed to form transiently prior to activation that may aid Brr2 in displacing U4 (Rodgers et al. 2016). Single molecule studies have elucidated intermediates within the activation pathway between B and B^{ACT} spliceosome structures (Fu et al. 2023; Hoskins et al 2016), but there is still much to understand about the mechanisms that occur during activation resulting in the formation of the spliceosome active site.

1.7 Second-step splicing factors mediate 3′ SS selection and fidelity

The BP and 5′ SS sequence are recognized by base pairing of an snRNA; however, the 3′ SS is identified by several protein factors, rather than an snRNA, that join after completion of the first step of splicing. The limited information content of the 3′ SS presents a challenge to identification of the correct nucleotides for exon ligation, and understanding of precise mechanisms required for 3′ SS selection is limited. Several factors are known to be involved in selection of the 3′ SS in yeast: Prp22, Slu7, Prp18, Prp8, and Fyv6. Prp17 also aids in exon-ligation, likely by stabilizing the exon-ligation conformation over the branching conformation as it changes position after Prp16 activity results in remodeling of the spliceosome (Jones et al. 1995; Wilkinsons et al. 2017).

Docking of the 3′ SS is aided by the second-step factors Prp18 and Slu7, scaffolded by the largest protein in the spliceosome Prp8 (Semlow et al. 2016, Warkocki et al. 2009). Additionally, the alpha finger of Prp8 is thought to facilitate 3′ SS selection as shown by specific interactions of Gln1948 with the **Y**AG nucleotide (-3 position bolded; Wilkinson et al. 2017). Prp22 remodels pre-mRNA to promote alternative 3′ SS sampling (Semlow et al. 2016) and proofreads the ligated exons after the second transesterification reaction (Mayas et al. 2006). Both functions of Prp22 in ensuring 3′ SS fidelity occur before or simultaneously with functions of Prp22 in releasing the mRNA product during the transition between P and ILS complexes.

The mechanism for selection of 3′ SS by the aforementioned factors is not well understood. Two of the five second-step factors, Prp18 and Fyv6, are nonessential in yeast, but their deletion produces severe growth defects suggesting these factors enhance in vivo splicing efficiency (Lipinski et al. 2023; Kawashima et al. 2009; Roy et al. 2023). Prior works established functions of Prp18 in stabilizing interactions between nucleotides in the exon and U5 snRNA using splicing reporters (Wilkinson et al. 2017; Liu et al. 2017) and stabilizing spliceosome conformations that favor the second step of splicing (Wilkinson et al. 2017; Wilkinson et al. 2021). Positioning of a conserved loop region in Prp18 near the 3′ SS -3 and -4 positions combined with upregulation of upstream 3′ SS when Prp18 is deleted suggest a role in splicing fidelity (Roy et al. 2023; Kawashima et al. 2014). Deletion of Prp18 results in a global loss of splicing efficiency, particular in nonribosomal protein genes, coupled with a loss in splicing fidelity seen by widespread activation of nonconsensus 3′ SS (Roy et al. 2023). Recruitment of Prp18 by Slu7 and interactions with the Prp8 RNaseH domains are required for promoting 3′ SS fidelity (Roy et al. 2023).

The mechanism of Slu7-mediated 3′ SS selection is less clear. Slu7 is inserted along the cavity through which the 3′ SS docks after the first step, suggesting interactions with the 3′ SS to either guide it into the active site or stabilize it once docked (Semlow et al. 2016; Brys et al. 1996; Chau et al. 1999). Prp18 and Slu7 have been shown to fill nonredundant roles in 3′ SS as rescue of Prp18 splicing defects cannot be rescued by an excess of Slu7 (Roy et al. 2023). Recently a new second-step factor, Fyv6, was identified in yeast that also influences 3′ SS selection (Zhan et al. 2022; Lipinski et al. 2023). Mechanisms as to how Fyv6 promotes 3′ SS fidelity are an active area of research. Prp18, Slu7, Prp22, and Fyv6 are dispensable for splicing of some substrates, but only when the distance between the BP and 3′ SS is short (Brys and Schwer, 1996; Schwer and Gross, 1998; Zhang and Schwer, 1997; Chapter 5).

Prp22 acts together with Slu7, Prp18, and also likely Fyv6 to promote exon ligation in an ATP-independent manner (Schwer et al 1998; Chapter 5). Prp18 dependent 3′ SS fidelity is synergized by Prp22 proofreading activity, which is thought to occur downstream of Prp18 function (Roy et al. 2023). After exon ligation, the spliceosomal P complex retains both the excised intron lariat and the spliced mRNA, which is held by proteins and the U5

snRNA (Wilkinson et al. 2019). ATP-dependent activity of Prp22 releases the mature mRNA, which also releases exon ligation factors Prp18 and Slu7 (Schwer et al. 1998; Company et al. 1991; Wagner et al. 1998). It is unclear whether Fyv6 is also released during this step (Chapter 5). Prp22 crosslinks to the second exon downstream of the splice junction and is thought to translocate 3′ to 5′ to disrupt the U5 snRNA-mRNA duplex (Schwer et al. 2008; McPheeters et al. 2000; McPheeters et al. 2003). However, psiCLIP revealed that Prp22 exhibits dynamic RNA binding before and after exon ligation, suggesting that proofreading activity of Prp22 may be influenced by spliceosome stability during exon ligation (Strittmatter et al. 2021).

1.8 Tools for fluorescent tagging of RNA molecules in vivo

The toolbox for fluorescently labeling RNAs in vivo for visualization and tracking is ever-expanding to provide a diversity of genetic tags in a wide spectrum of colors. One of the first systems was GFP (green fluorescent protein) fused to the RNA phage capsid protein MS2, also called MS2 coat protein (MCP) (Bertrand et al. 1998). The stem loop binding site (MS2) was inserted as a tandem array of six binding sites into the ASH1 mRNA in *S. cerevisiae* for visualization of mRNA localization within the cell (Bertrand et al. 1998). Longer tandem arrays of MS2 (24 copies) have been used for single molecule resolution of mRNAs (Fusco et al. 2003). These longer tandem arrays extend the mRNA by over 500 nts and disrupt RNA metabolism due to tight binding of the MCP (Garcia and Parker 2015, 2016; Henrich et al. 2017). Improvements to the MS2 system have been made to improve degradation and turnover of reporter mRNAs and reduce background from unbound fluorescent protein among others (Tutucci et al. 2018; Park et al. 2020). Other phage-derived stem loop systems have also been developed, notably the PP7 system (Chao et al. 2008). Despite several disadvantages of stem loop systems, they are easily genetically encodable and applicable to a wide set of RNAs.

Another genetically encodable fluorescent system are fluorogenic RNA aptamers, which are RNA sequences selected in vitro for binding and activation of fluorescence in otherwise non fluorescent small molecules (Bouhedda et al., 2017; Swetha et al., 2020). Fluorescence enhancement by RNA aptamers occurs via rotational inhibition of their small molecule ligands (Huang et al 2023). The RNA aptamers and their fluorogenic ligands are now available in a wide spectrum of fluorescence emission and have been broadly applied for in vivo localization and tracking of RNA molecules. One such aptamer is Mango, selected for increased brightness and smaller size (~30 nt) than previously developed RNA aptamers (Dolgosheina et al., 2014). Mango binds the fluorogenic small molecule TO1-Biotin with nanomolar affinity and enhances fluorescence of TO1 (thiazole orange) up to 1100-fold. Crystallization of Mango bound to TO1 showed a three-tiered G-quadruplex structure with TO1-Biotin bound against the top quadruplex (Trachman III et al., 2017). Studies of Mango in vivo showed that the aptamer exhibits comparatively high

photostability and folding efficiency (Cawte et al. 2020). Reselected variants of Mango resulted in Mango-II, Mango-III, iMango-III, and Mango-IV. Variants of Mango have slight variations in stability but, more notably, different binding of TO1-Biotin on the G-quadruplex face and dimerization of aptamers in Mango-IV (Trachman III et al. 2018; Trachman III et al. 2020). More recently, the reselection of Mango to bind a different fluorescent ligand resulted in Peach (Kong et al. 2021). Peach is useful for both orthogonal fluorescent labeling with Mango and observation of Förster resonance energy transfer (FRET) between RNAs with encoded Mango or Peach aptamers (Kong et al. 2021). The expanded toolset of fluorogenic RNA aptamers makes tagging of RNA for fluorescence visualization easy and highly customizable.

1.9 Thesis Overview

My thesis work aims to better understand the regulation of RNA-protein interactions that occur between components of the splicing machinery. Chapter 2 focuses on understanding how the U6 snRNA is produced as it is uniquely processed from the other spliceosomal snRNAs. U6 transcribed by RNAP II (U6-II) instead of RNAP III is sufficient for viability when expressed in *S. cerevisiae*. However, low cellular levels and instability of U6-II were observed as well as disruptions to endogenous snRNP distributions. Addition of a binding site for the Sm ring (U6-II-Sm), normally found on the RNAP II-transcribed snRNAs, provided stabilization and increased cellular levels of U6. To determine if U6-II-Sm is stabilized by binding of the Sm ring, we immunoprecipitated epitope-tagged Sm or Lsm rings and examined the associated RNAs by primer extension. U6-II-Sm is indeed found to be associated with Sm rings, likely explaining the increased stability. RNAP II transcription does produce a functional U6 but affects downstream processing step and assembly into snRNPs.

Continuing to explore how RNA-protein interactions shape RNA dynamics in the spliceosome, Chapter 3 focuses on developing implementing methods for endogenous fluorescent labeling of the U4 snRNA for use in single molecule studies of spliceosome activation. Monitoring U4 snRNA release separate from U4 associated proteins will yield insight into both regulation of Brr2-mediated snRNA release and conformational dynamics of the U6 snRNA during splicing. Two strategies were examined: 1) insertion into U4 of RNA aptamers for tethering fluorophores suitable for imaging of single molecules and 2) incorporation of a MS2 stem loop for binding recombinant MS2 coat protein (MCP) fused to a SNAP tag for fluorescent labeling. RNA tags inserted into U4 were generally well tolerated as yeast were viable, grew robustly at optimal temperatures, and had few growth defects when placed at nonoptimal temperatures. Tagged U4 variants that displayed some growth defects also showed correlated decreases in splicing efficiency. Initial single molecule experiments did not yield promising results, so further optimization is required of these fluorescent systems.

Whereas substrates for the 1st step of splicing catalysis are selected and mainly coordinated by RNA-RNA interactions between the intron and snRNAs, substrates for the 2nd step of catalysis are selected by both protein components and intramolecular interactions within the intron. Interactions of proteins with the 5' SS and 3' SS aid in positioning and stable docking within the spliceosome active site prior to catalysis. In Chapters 4 and 5, we have characterized a new splicing factor, the nonessential and uncharacterized protein Fyv6. Deletion of *FYV6* produces a dramatic growth phenotype, results in accumulation of splicing intermediates, and causes selection of an alternative 3' SS within the *SUS1* gene in vivo. Changes in 3' SS selection occur widely within the transcriptome where BP proximal 3' SS are used more frequently in spliceosomes lacking Fyv6. Genetic experiments suggest that Fyv6 acts during the second step of splicing additionally supported by presence of Fyv6 in a high-resolution P complex spliceosome structure solved by our collaborator Max Wilkinson. Fyv6 is an elongated protein with extensive contacts spanning the spliceosome, but notably with the Prp22 helicase that releases the mRNA product after the second step. Structure-function analysis of Fyv6 using both truncations of Fyv6, aided by the P complex structure, and a screen for suppressors of the deletion phenotype was conducted. Truncation analysis revealed a functionally important interaction with the NTC component Syf1. Suppression of the *fyv6Δ* growth phenotypes can be accomplished via single point mutations within a number of splicing factors. Quantitation of splicing with nonconsensus 3' SS reporters correlated *fyv6Δ* splicing phenotypes with Prp22 release-defect mutations.

Chapter 6 provides a summary of major contributions detailed in this thesis and discusses future directions. Methods described in this thesis are broadly applicable for isolating U4 snRNP species for further downstream characterization of structure or function, have laid the groundwork for future single molecule studies of the transition from the B to B^{ACT} spliceosome during activation, and provide an opportunity to characterize the effects of Brr2 mutations on activation. Further, there is still much unknown about Fyv6, particularly its interaction with Prp22.

1.10 References

Absmeier E, Santos KF, Wahl MC. 2016. Functions and regulation of the Brr2 RNA helicase during splicing. *Cell Cycle* **15**(24):3362-3377. doi:10.1080/15384101.2016.1249549

Abovich N, Liao XC, Rosbash M. 1994. The yeast MUD2 protein: An interaction with PRP11 defines a bridge between commitment complexes and U2 snRNP addition. *Genes Dev.* **8**:843–854. doi: 10.1101/gad.8.7.843.

Achsel T, Brahms H, Kastner B, Bachi A, Wilm M, Lührmann R. 1999. A doughnut-shaped heteromer of human Sm-like proteins binds to the 3′-end of U6 snRNA, thereby facilitating U4/U6 duplex formation in vitro. *EMBO J* **18**: 5789–5802. doi: 10.1093/emboj/18.20.5789

Agafonov DE, Deckert J, Wolf E, Odenwalder P, Bessonov S, Will CL, Urlaub H, Lührmann R. 2011. Semiquantitative proteomic analysis of the human spliceosome via a novel two-dimensional gel electrophoresis method. *Mol Cell Biol* **31**:2667-82. doi:10.1128/MCB.05266-11

Bertrand E, Bordonné R. 2004. Assembly and traffic of small nuclear RNPs. *Prog. Mol. Subcell. Biol.* **35**:79-97. doi: 10.1007/978-3-540-74266-1_4

Bhadra M, Howell P, Dutta S, Heintz C, Mair WB. 2020. Alternative splicing in aging and longevity. *Hum. Genet.* **139**(3):357-369. doi: 10.1007/s00439-019-02094-6.

Bertrand E, Chartrand P, Schaefer M, Shenoy SM, Singer RH, Long RM. 1998. Localization of ASH1 mRNA Particles in Living Yeast. *Mol. Cell* **2**(4):437-445. doi: 10.1016/s1097-2765(00)80143-4

Bouhedda F, Autour A, Ryckelynck M. 2017. Light-Up RNA Aptamers and Their Cognate Fluorogens: From Their Development to Their Applications. *Int. J. Mol. Sci.* **19**(1):44. doi: 10.3390/ijms19010044

Braglia P, Kawauchi J, Proudfoot NJ. 2011. Co-transcriptional RNA cleavage provides a failsafe termination mechanism for yeast RNA polymerase I. *Nucleic Acids Res* **39**(4):1439-1448. doi:10.1093/nar/gkq894

Brody E, Abelson J. 1985. The “spliceosome”: yeast pre-messenger RNA associates with a 40S complex in a splicing-dependent reaction. *Science* **228**:4702963–67. doi: 10.1126/science.3890181

Brow DA, Guthrie C. 1990. Transcription of a yeast U6 snRNA gene requires a polymerase III promoter element in a novel position. *Genes Dev* **4**: 1345–1356. doi: 10.1101/gad.4.8.1345

Brys A, Schwer B. 1996. Requirement for SLU7 in yeast pre-mRNA splicing is dictated by the distance between the branchpoint and the 3′ splice site. *RNA* **2**:7707–17.

Cawte AD, Unrau PJ, Rueda DS. 2020. Live cell imaging of single RNA molecules with fluorogenic Mango II arrays. *Nat Commun* **11**:1283. doi:10.1038/s41467-020-14932-7

Chao JA, Patskovsky Y, Almo SC, Singer RH (2008) Structural basis for the coevolution of a viral RNA-protein complex. *Nat Struct Mol Biol* **15**(1):103–105. doi: 10.1038/nsmb1327.

Chanfreau G, Rotondo G, Legrain P, Jacquier A. 1998. Processing of a dicistronic small nucleolar RNA precursor by the RNA endonuclease Rnt1. *EMBO J* **17**(13):3726-3737. doi:10.1093/emboj/17.13.3726

Chua K, Reed R. 1999. The RNA splicing factor hSlu7 is required for correct 3′ splice-site choice. *Nature* **402**:6758207–10. doi: 10.1038/46086

Company M, Arenas J, Abelson J. 1991. Requirement of the RNA helicase-like protein PRP22 for release of messenger RNA from spliceosomes. *Nature* **349**:487–493. doi: 10.1038/349487a0.

Cordin O, Hahn D, Beggs JD. 2012. Structure, function and regulation of spliceosomal RNA helicases. *Curr Opin Cell Biol* **24**:431-8. doi: 10.1016/j.ceb.2012.03.004

Cordin O, Beggs JD. 2013. RNA helicases in splicing. *RNA Biol* **10**:83-95. doi:10.4161/rna.22547

Coy S, Volanakis A, Shah S, Vasiljeva L. 2013. The Sm complex is required for the processing of non-coding RNAs by the exosome. *PLoS ONE* **8**:e65606. doi: 10.1371/journal.pone.0065606

Cui Y, Cai M, Stanley HE. 2017. Comparative analysis and classification of cassette exons and constitutive exons. *Biomed Res Int.* **2017**:7323508. doi: 10.1155/2017/7323508

Didychuk AL, Montemayor EJ, Carrocci TJ, DeLaitsch AT, Lucarelli SE, Westler WM, Brow DA, Hoskins AA, Butcher SE. 2017. Usb1 controls U6 snRNP assembly through evolutionarily divergent cyclic phosphodiesterase activities. *Nat Commun* **8**: 497. doi: 10.1038/s41467-017-00484-w

Domdey H, Apostol B, Lin RJ, Newman AJ, Brody E, Abelson J 1984. Lariat structures are in vivo intermediates in yeast pre-mRNA splicing. *Cell* **39**:3, Part 2611–21. doi: 10.1016/0092-8674(84)90468-9

Fica SM, Tuttle N, Novak T, Li N-S, Lu J et al. 2013. RNA catalyses nuclear pre-mRNA splicing. *Nature* **503**:7475229–34. doi: 10.1038/nature12734

Fu X, Kaur H, Rodgers ML, Montemayor EJ, Butcher SE, Hoskins AA. 2022. Identification of transient intermediates during spliceosome activation by single molecule

fluorescence microscopy. *PNAS* **119**(48): e2206815119. doi: 10.1073/pnas.2206815119

Fusco D, Accornero N, Lavoie B, Shenoy SM, Blanchard J-M, Singer RH, Bertrand E. 2003. Single mRNA Molecules Demonstrate Probabilistic Movement in Living Mammalian Cells. *Curr. Biol.* **13**(2):161-167. doi: 10.1016/s0960-9822(02)01436-7

Gehring NH, Roignant JY. 2021. Anything but ordinary – emerging splicing mechanisms in eukaryotic gene regulation. *Trends Genet.* **37**(4):355–372. doi: 10.1016/j.tig.2020.10.008

Garcia JF, Parker R. 2015. MS2 coat proteins bound to yeast mRNAs block 5′ to 3′ degradation and trap mRNA decay products: implications for the localization of mRNAs by MS2-MCP system. *RNA* **21**: 1393–1395. doi: 10.1261/rna.051797.115

Garcia JF, Parker R. 2016. Ubiquitous accumulation of 3′ mRNA decay fragments in Saccharomyces cerevisiae mRNAs with chromosomally integrated MS2 arrays. *RNA* **22**: 657–659. doi: 10.1261/rna.056325.116

Hang J, Wan R, Yan C, Shi Y. 2015. Structural basis of pre-mRNA splicing. *Science* **349**:1191-8. doi: 10.1126/science.aac8159

Henrich S, Sidler CL, Azzalin CM, Weis K. 2017. Stem–loop RNA labeling can affect nuclear and cytoplasmic mRNA processing. *RNA* **23**(2): 134–141. doi: 10.1261/rna.057786.116.

Hnilicova J, Stanek D. 2011. Where splicing joins chromatin. *Nucleus* **2**(3):182–188. doi: 10.4161/nucl.2.3.15876

Hoskins AA, Moore MJ. 2012. The spliceosome: a flexible, reversible macromolecular machine. *Trends Biochem Sci.* **37**(5):179-88. doi: 10.1016/j.tibs.2012.02.009.

Hoskins AA, Rodgers ML, Friedman LJ, Gelles J, Moore MJ. 2016. Single molecule analysis reveals reversible and irreversible steps during spliceosome activation. *eLife* **5**: e14166. doi: 10.7554/eLife.14166

Huang Z, Guo X, Ma X, Wang F, Jiang J-H. 2023. Genetically encodable tagging and sensing systems for fluorescent RNA imaging. *Biosensors and Bioelectronics* **219**:114769. doi: 10.1016/j.bios.2022.114769.

Jones MH, Frank DN, Guthrie C 1995. Characterization and functional ordering of Slu7p and Prp17p during the second step of pre-mRNA splicing in yeast. *PNAS* **92**:219687–91. doi: 10.1073/pnas.92.21.9687

Kaur H, van der Feltz C, Sun Y, Hoskins AA. 2022. Network theory reveals principles of spliceosome structure and dynamics. *Structure* **30**: 190–200. e192. doi: 10.1016/j.str.2021.09.003

Kawashima T, Pellegrini M, Chanfreau GF. 2009. Nonsense-mediated mRNA decay mutes the splicing defects of spliceosome component mutations. *RNA* **15**:2236–2247. doi: 10.1261/rna.1736809

Kawashima T, Douglass S, Gabunilas J, Pellegrini M, Chanfreau GF. 2014. Widespread use of non-productive alternative splice sites in Saccharomyces cerevisiae. *PLoS Genet.* **10**:e1004249. doi: 10.1371/journal.pgen.1004249

Köhler A, E. Hurt E. 2007. Exporting RNA from the nucleus to the cytoplasm. *Nat. Rev. Mol. Cell Biol.* **8**:761-773. doi: 10.1038/nrm2255

Kong KYS, Jeng SCY, Rayyan B, Unrau PJ. 2021. RNA Peach and Mango: Orthogonal two-color fluorogenic aptamers distinguish nearly identical ligands. *RNA* **11**:1283. doi: 10.1261/rna.078493.120

Kuang Z, Boeke JD, Canzar S. 2017. The dynamic landscape of fission yeast meiosis alternative-splice isoforms. *Genome Res.* **27**(1):145-156. doi:10.1101/gr.208041.116

Kurosaki T, Popp MW, Maquat LE. 2019. Quality and quantity control of gene expression by nonsense-mediated mRNA decay. *Nat Rev Mol Cell Biol.* **20**(7):406–420. doi: 10.1038/s41580-019-0126-2

Licht K, Medenbach J, Lührmann R, Kambach C, Bindereif A. 2008. 3′-cyclic phosphorylation of U6 snRNA leads to recruitment of recycling factor p110 through LSm proteins. *RNA* **14**: 1532–1538. doi: 10.1261/rna.1129608

Lipinski KA, Senn KA, Zeps NJ, Hoskins AA. 2023. Biochemical and genetic evidence supports Fyv6 as a second-step splicing factor in Saccharomyces cerevisiae. *RNA* **29**(11):1792-1802. doi:10.1261/rna.079607.123.

Lim LP, Burge CB. 2001. A computational analysis of sequence features involved in recognition of short introns. *Proc. Natl. Acad. Sci. USA.* **98**:11193–11198. doi: 10.1073/pnas.201407298.

Liu S, Li X, Zhang L, Jiang J, Hill RC et al. 2017. Structure of the yeast spliceosomal postcatalytic P complex. *Science* **358**:63681278–83. doi: 10.1126/science.aar3462

Love, SL, Emerson JD, Koide K, Hoskins AA. 2023. Pre-mRNA splicing-associated diseases and therapies. *RNA Biology* **20**(1):525–538. doi: 10.1080/15476286.2023.2239601

Lund E, Dahlberg JE. 1992. Cyclic 2′,3′-phosphates and nontemplated nucleotides at the 3′ end of spliceosomal U6 small nuclear RNA's. *Science* **255**: 327–330. doi: 10.1126/science.1549778

Marquez Y, Brown JW, Simpson C, Barta A, Kalyna M. 2012. Transcriptome survey reveals increased complexity of the alternative splicing landscape in Arabidopsis. *Genome Res.* **22**(6):1184–1195. doi: 10.1101/gr.134106.111

Matera AG, Wang Z. 2014. A day in the life of the spliceosome. *Nat. Rev. Mol. Cell Biol.*, **15**:108-121. doi: 10.1038/nrm3742

Matlin AJ, Clark F, Smith CWJ. 2005. Understanding alternative splicing: towards a cellular code. *Nat Rev Mol Cell Biol.* **6**(5):386–398. doi: 10.1038/nrm1645

McPheeters DS, Schwer B, Muhlenkamp P. 2000. Interaction of the yeast DExH-box RNA helicase prp22p with the 3′ splice site during the second step of nuclear pre-mRNA splicing. *Nucleic Acids Res.* **28**:1313–1321. doi: 10.1093/nar/28.6.1313.

McPheeters DS, Muhlenkamp P. 2003. Spatial organization of protein-RNA interactions in the branch site-3′ splice site region during pre-mRNA splicing in yeast. *Mol. Cell Biol.* **23**:4174–4186. doi: 10.1128/MCB.23.12.4174-4186.2003.

Mayas RM, Maita H, Staley JP. 2006. Exon ligation is proofread by the DExD/H-box ATPase Prp22p. *Nat. Struct. Mol. Biol.* **13**:482–490. doi: 10.1038/nsmb1093

Moenne A, Camier S, Anderson G, Margottin F, Beggs J, Sentenac A. 1990. The U6 gene of Saccharomyces cerevisiae is transcribed by RNA polymerase C (III) in vivo and in vitro. *EMBO J* **9**: 271–277. doi: 10.1002/j.1460-2075.1990.tb08105.x

Mouaikel J, Verheggen C, Bertrand E, Tazi J, Bordonné R. 2022. Hypermethylation of the cap structure of both yeast snRNAs and snoRNAs requires a conserved methyltransferase that is localized to the nucleolus. *Mol. Cell.* **9**:891–901. doi: 10.1016/s1097-2765(02)00484-7

Nilsen TW, Graveley BR. 2010. Expansion of the eukaryotic proteome by alternative splicing. *Nature* **463**(7280):457–463. doi: 10.1038/nature08909

Nguyen TH, Galej WP, Bai XC, Oubridge C, Newman AJ, Scheres SH, Nagai K. 2016. Cryo-EM structure of the yeast U4/U6.U5 tri-snRNP at 3.7 A resolution. *Nature* **530**:298-302. doi: 10.1038/nature16940

Padgett RA, Konarska MM, Grabowski PJ, Hardy SF, Sharp PA. 1984. Lariat RNA's as intermediates and products in the splicing of messenger RNA precursors. *Science* **225**:4665898–903. doi: 10.1126/science.6206566

Pannone BK, Xue D, Wolin SL. 1998. A role for the yeast La protein in U6 snRNP assembly: evidence that the La protein is a molecular chaperone for RNA polymerase III transcripts. *EMBO J* **17**: 7442–7495. doi: 10.1093/emboj/17.24.7442

Panse VG, Johnson AW. 2010. Maturation of eukaryotic ribosomes: Acquisition of functionality. *Trends Biochem Sci* **35**: 260–266. doi: 10.1016/j.tibs.2010.01.001

Park SY, Moon HC, Park HY. 2020. Live-cell imaging of single mRNA dynamics using split superfolder green fluorescent proteins with minimal background. *RNA* **26**(1):101-109. doi: 10.1261/rna.067835.118.

Parker R, Patterson B. 1987. Architecture of fungal introns: Implications for spliceosome assembly. In Molecular biology of RNA: New perspectives (ed. M. Inouye and B. Dudock), pp. 133–149. Academic Press, San Diego, CA.

Patterson B, Guthrie C. 1991. A U-rich tract enhances usage of an alternative 3′ splice site in yeast. *Cell.* **64**:181–187. doi: 10.1016/0092-8674(91)90219-O.

Qin D, Huang L, Wlodaver A, Andrade J, Staley JP. 2016. Sequencing of lariat termini in S. cerevisiae reveals 5′ splice sites, branch points, and novel splicing events. *RNA* **22**:2237–53. doi: 10.1261/rna.052829.115

Raghunathan PL, Guthrie C. 1998. RNA unwinding in U4/U6 snRNPs requires ATP hydrolysis and the DEIH-box splicing factor Brr2. *Curr Biol.* **8**:847–855. doi: 10.1016/S0960-9822(07)00345-4

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

Rinke J, Steitz JA. 1985. Association of the lupus antigen La with a subset of U6 snRNA molecules. *Nucleic Acids Res* **13**: 2617–2629.

Rodgers ML, Didychuk AL, Butcher SE, Brow DA, Hoskins AA. 2016. A multi-step model for facilitated unwinding of the yeast U4/U6 RNA duplex. *Nucleic Acids Res.* **44**(22):10912-10928. doi: 10.1093/nar/gkw686.

Rodriguez JR, Pikielny CW, Rosbash M. 1984. In vivo characterization of yeast mRNA processing intermediates. *Cell* **39**:3, Part 2603–10. doi: 10.1016/0092-8674(84)90467-7

Roy KR, Gabunilas J, Neutel D, Ai M, Yeh Z, Samson J, Lyu G, Chanfreau GF. 2023. Splicing factor Prp18p promotes genome-wide fidelity of consensus 3′-splice sites. *Nucleic Acids Research* **5**(22):12428–12442. doi: 10.1093/nar/gkad968.

Ruskin B, Krainer AR, Maniatis T, Green MR. 1984. Excision of an intact intron as a novel lariat structure during pre-mRNA splicing in vitro. *Cell* **38**:1317–31. doi: 10.1016/0092-8674(84)90553-1

Rymond BC, Rosbash M. 1992. Yeast pre-mRNA splicing. J.R. Broach, J. Pringle, E.W. Jones (Eds.), The Molecular and Cellular Biology of the Yeast Saccharomyces, Volume 2, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Santos KF, Mozaffari-Jovin S, Weber G, Pena V, Lührmann R, Wahl MC. 2012. Structural basis for functional cooperation between tandem helicase cassettes in Brr2-mediated remodeling of the spliceosome. *Proc Natl Acad Sci USA* **109**:17418-23 doi: 10.1073/pnas.1208098109.

Sashital DG, Cornilescu G, McManus CJ, Brow DA, Butcher SE. 2004. U2-U6 RNA folding reveals a group II intron-like domain and a four-helix junction. *Nat Struct Mol Biol* **11**:1237-42. doi: 10.1038/nsmb863

Semlow DR, Blanco MR, Walter NG, Staley JP. 2016. Spliceosomal DEAH-Box ATPases remodel pre-mRNA to activate alternative splice sites. *Cell* **164**:5985–98. doi: 10.1016/j.cell.2016.01.025

Schwer B, Gross CH. 1998. Prp22, a DExH-box RNA helicase, plays two distinct roles in yeast pre-mRNA splicing. *EMBO J.* **17**:2086–2094. doi: 10.1093/emboj/17.7.2086.

Schwer B. 2008. A conformational rearrangement in the spliceosome sets the stage for Prp22-dependent mRNA release. *Mol. Cell.* **30**:743–754. doi: 10.1016/j.molcel.2008.05.003.

Shen S, Park JW, Lu Z-X. 2014. rMATS: Robust and flexible detection of differential alternative splicing from replicate RNA-Seq data. *Biological Sciences* **111**(51): E5593-E5601. doi: 10.1073/pnas.1419161111

Sheth N, Roca X, Hastings ML, Roeder T, Krainer AR, Sachidanandam R. 2006. Comprehensive splice-site analysis using comparative genomics. *Nucleic Acids Res* **34**:143955–67. doi: 10.1093/nar/gkl556

Sinitcyn P, Richards AL, Weatheritt RJ, et al. 2023. Global detection of human variants and isoforms by deep proteome sequencing. *Nat Biotechnol.* **41**:1776–1786. doi:10.1038/s41587-023-01714-x.

Sloan KE, Gleizes PE, Bohnsack MT. 2016. Nucleocytoplasmic transport of RNAs and RNA-protein complexes. *J. Mol. Biol.* **428**:2040-2059. doi: 10.1016/j.jmb.2015.09.023

Spaniel C, Honda M, Selitsky SR, Yamane D, Shimakami T, Kaneko S, Lanford RE, Lemon SM. 2013. microRNA-122 abundance in hepatocellular carcinoma and

non-tumor liver tissue from Japanese patients with persistent HCV versus HBV infection. *PLoS One* **8**: e76867. doi: 10.1371/journal.pone.0076867

Spiller MP, Boon KL, Reijns MA, Beggs JD. 2007a. The Lsm2-8 complex determines nuclear localization of the spliceosomal U6 snRNA. *Nucleic Acids Res* **35**: 923–929. doi: 10.1093/nar/gkl1130

Spiller MP, Reijns MA, Beggs JD. 2007b. Requirements for nuclear localization of the Lsm2-8p complex and competition between nuclear and cytoplasmic Lsm complexes. *J Cell Sci* **120**: 4310–4320. doi: 10.1242/jcs.019943

Spingola M, Grate L, Haussler D, Ares M. 1999. Genome-wide bioinformatic and molecular analysis of introns in Saccharomyces cerevisiae. *RNA* **5**:2221–34. doi: 10.1017/s1355838299981682

Staley JP, Guthrie C. 1998. Mechanical devices of the spliceosome: motors, clocks, springs, and things. *Cell* **92**:315-26. doi: 10.1016/S0092-8674(00)80925-3

Steitz TA, Steitz JA. 1993. A general two-metal-ion mechanism for catalytic RNA. *PNAS* **90**:146498–502. doi: 10.1073/pnas.90.14.6498

Stevens SW, Abelson J. 1999. Purification of the yeast U4/U6·U5 small nuclear ribonucleoprotein particle and identification of its proteins. *Biochemistry* **96**:(13) 7226-7231. doi: 10.1073/pnas.96.13.7226

Strittmatter, LM, Capitanchik C, Newman AJ, et al. 2021. psiCLIP reveals dynamic RNA binding by DEAH-box helicases before and after exon ligation. *Nat Commun* **12:**1488. doi:10.1038/s41467-021-21745-9.

Strunk BS, Loucks CR, Su M, Vashisth H, Cheng S, Schilling J, Brooks CL III, Karbstein K, Skiniotis G. 2011. Ribosome assembly factors prevent premature translation initiation by 40S assembly intermediates. *Science* **333**: 1449–1453. doi: 10.1126/science.1208245

Swetha P, Fan PZ, Wang F, Jiang J-H. 2020. Genetically encoded light-up RNA aptamers and their applications for imaging and biosensing. *J. Mater. Chem. B* **8**:3382-3392. doi: 10.1039/C9TB02668A

Tazi J, Forne T, Jeanteur P, Cathala G, Brunel C. 1993. Mammalian U6 small nuclear RNA undergoes 3′ end modifications within the spliceosome. *Mol Cell Biol* **13**: 1641–1650. doi: 10.1128/mcb.13.3.1641-1650.1993

Terns MP, Lund E, Dahlberg JE. 1992. 3′-end-dependent formation of U6 small nuclear ribonucleoprotein particles in Xenopus laevis oocyte nuclei. *Mol Cell Biol* **12**: 3032–3040. doi: 10.1128/mcb.12.7.3032

Theuser M, Hobartner C, Wahl MC, Santos KF. 2016. Substrate-assisted mechanism of RNP disruption by the spliceosomal Brr2 RNA helicase. *Proc Natl Acad Sci USA* **113**:7798-803. doi:10.1073/pnas.1524616113.

Tseng C-K, Cheng S-C. 2008. Both Catalytic Steps of Nuclear Pre-mRNA Splicing Are Reversible. *Science* **320**:1782-1784. doi:10.1126/science.1158993.

Trachman III RJ, Demeshkina N, Lau M, et al. 2017. Structural basis for high-affinity fluorophore binding and activation by RNA Mango. *Nat Chem Biol* **13**: 807–813 doi: 10.1038/nchembio.2392

Trachman III RJ, Abdolahzadeh A, Andreoni A, Cojocaru R, Knutson JR, Rycklynck M, Unrau PJ, Ferré-D'Amaré AR. 2018. Crystal Structures of the Mango-II RNA Aptamer Reveal Heterogeneous Fluorophore Binding and Guide Engineering of Variants with Improved Selectivity and Brightness. *Biochemistry* **57**(26):3544–3548. doi: 10.1021/acs.biochem.8b00399

Trachman III RJ, Cojocaru R, Wu D, Piszczek G, Ryckelynck M, Unrau PJ, Ferré-D'Amaré AR. 2020. Structure-Guided Engineering of the Homodimeric Mango-IV Fluorescence Turn-on Aptamer Yields an RNA FRET Pair. *Structure* **28**(7):776-785.e3. doi: 10.1016/j.str.2020.04.007.

Tutucci E, Vera M, Biswas J, Garcia J, Parker R, Singer RH. 2018. An improved MS2 system for accurate reporting of the mRNA life cycle. *Nat Methods.* **15**(1):81-89. doi: 10.1038/nmeth.4502.

Ule J, Blencowe BJ. 2019. Alternative splicing regulatory networks: functions, mechanisms and evolution. *Mol. Cell* **76**:329–45. doi: 10.1016/j.molcel.2019.09.017

Wagner JDO. 1998. The DEAH-box protein PRP22 is an ATPase that mediates ATP-dependent mRNA release from the spliceosome and unwinds RNA duplexes. *EMBO J.* **17**:2926–2937. doi:10.1093/emboj/17.10.2926.

Wang ET, Sandberg R, Luo S, et al. 2008. Alternative isoform regulation in human tissue transcriptomes. *Nature* **456**(7221):470–476. doi: 10.1038/nature07509

Warkocki Z, Odenwälder P, Schmitzová J, Platzmann F, Stark H et al. 2009. Reconstitution of both steps of Saccharomyces cerevisiae splicing with purified spliceosomal components. *Nat. Struct. Mol. Biol.* **16**:121237–43. doi: 10.1038/nsmb.1729

Wilkinson ME, Fica SM, Galej WP, Norman CM, Newman AJ, Nagai K. 2017. Postcatalytic spliceosome structure reveals mechanism of 3′-splice site selection. *Science* **358**:63681283–88. doi: 10.1126/science.aar3729

Wilkinson ME, Charenton C, Nagai K. 2019. RNA splicing by the spliceosome. Annu. Rev. *Biochem.* **89**:1–30. doi: 10.1146/annurev-biochem-091719-064225

Wilkinson ME, Fica SM, Galej WP, Nagai K. 2021. Structural basis for conformational equilibrium of the catalytic spliceosome. *Mol. Cell.* **81**:1439–1452. doi: 10.1016/j.molcel.2021.02.021

Will CL, Lührmann R. 2011. Spliceosome structure and function. *Cold Spring Harb. Perspect. Biol.* **3**: a003707. doi: 10.1101/cshperspect.a003707

Woll MG, Naryshkin NA, Karp GA. 2017. Drugging pre-mRNA splicing. RNA Therapeutics Topics in Medicinal Chemistry, 27. Springer, Cham. doi: 10.1007/7355_2017_12.

Yean SL, Wuenschell G, Termini J, Lin RJ. 2000. Metal-ion coordination by U6 small nuclear RNA contributes to catalysis in the spliceosome. *Nature* **408**(6814):881-884. doi:10.1038/35048617

Yu J, Yang Z, Kibukawa M, Paddock M, Passey DA, Wong GK. 2002. Minimal introns are not “junk” *Genome Res.* **12**:1185–1189. doi: 10.1101/gr.224602.

Zhan X, Lu Y, Zhang X, Yan C, Shi Y. 2022. Mechanism of exon ligation by human spliceosome. *Molecular Cell* **82**:2769-2778. doi: 10.1016/j.molcel.2022.05.021

Zhang X, Schwer B. 1997. Functional and physical interaction between the yeast splicing factors Slu7 and Prp18. *Nucleic Acids Res.* **25**:2146–2152. doi: 10.1093/nar/25.11.2146

CHAPTER 2

Yeast U6 snRNA made by RNA polymerase II is less stable but functional

A version of this chapter is published as: Karli A. Lipinski*, Jing Chi*, Xin Chen, Aaron A. Hoskins, and David A. Brow. 2022. Yeast U6 snRNA made by RNA polymerase II is less stable but functional. *RNA* **28**(12): 1606–1620. doi: 10.1261/rna.079328.122

*Indicates co-first authors

DAB and JC originated the project. KAL, JC, and XC performed experiments and analyzed data. KAL, JC, AAH, and DB wrote the manuscript. See contributions (pg. xii).

CHAPTER 2: Yeast U6 snRNA made by RNA polymerase II is less stable but functional

2.1 Abstract

U6 small nuclear (sn)RNA is the shortest and most conserved snRNA in the spliceosome and forms a substantial portion of its active site. Unlike the other four spliceosomal snRNAs, which are synthesized by RNA polymerase (RNAP) II, U6 is made by RNAP III. To determine if some aspect of U6 function is incompatible with synthesis by RNAP II, we created a U6 snRNA gene with RNAP II promoter and terminator sequences. This “U6-II” gene is functional as the sole source of U6 snRNA in yeast, but its transcript is much less stable than U6 snRNA made by RNAP III. Addition of the U4 snRNA Sm protein binding site to U6-II increased its stability and led to formation of U6-II•Sm complexes. We conclude that synthesis of U6 snRNA by RNAP III is not required for its function and that U6 snRNPs containing the Sm complex can form in vivo. The ability to synthesize U6 snRNA with RNAP II relaxes sequence restraints imposed by intragenic RNAP III promoter and terminator elements and allows facile control of U6 levels via regulators of RNAP II transcription.

2.2 Introduction

The five small nuclear RNAs U1, U2, U4, U5, and U6 are essential components of the eukaryotic spliceosome, necessary both for precise identification of the intron–exon boundaries and splicing catalysis (Guthrie and Patterson 1988). Of the five spliceosomal snRNAs, four are synthesized by RNA polymerase (RNAP) II. Only U6, the smallest of the spliceosomal RNAs, is synthesized by RNAP III (Didychuk et al. 2018; Dergai and Hernandez 2019). Indeed, it appears that RNAP II is actively excluded from the vicinity of the *Saccharomyces cerevisiae* yeast U6 snRNA gene, *SNR6* (Steinmetz et al. 2006). In contrast, while human major and minor (U6atac) U6 snRNAs are also made by RNAP III, the presence of RNAP II upstream of the U6 gene promoters appears to enhance transcription by RNAP III (Listerman et al. 2007; Younis et al. 2013).

It is not known why U6 is made by RNAP III, but it could be because some characteristic of RNAP II synthesis is inconsistent with U6 snRNA function. For example, RNAP III typically starts transcription at a single position, defined primarily by spacing from an intragenic “A block” promoter element (Gerlach et al. 1995), whereas RNAP II often starts at multiple positions defined by a less stringent initiator element (Kuehner and Brow 2006). Also, RNAP II transcripts receive a reverse 7-methylguanosine “cap” on their 5′ triphosphate, whereas RNAP III transcripts generally do not. Human U6 snRNA instead receives a γ-monomethylphosphate cap on its 5′ triphosphate (Singh and Reddy 1989),

which may be present in yeast and other organisms as well (Didychuk et al. 2018). However, yeast U6 molecules made by RNAP III but bearing altered 5′ ends and methylguanosine caps are apparently functional in vivo (Kwan et al. 2000).

Another possible reason that U6 snRNA is made by RNAP III is to promote assembly of the Lsm2–8 protein ring on its 3′ tail (Mayes et al. 1999; Karaduman et al. 2006). All transcripts made by RNAP III terminate with an oligo(U) tail of four or more residues (Arimbasseri and Maraia 2015). On yeast U6 snRNA the oligo(U) tail, which initially has a terminal 2′,3′ cis diol, is bound by the La-homologous protein Lhp1 (Pannone et al. 1998; Didychuk et al. 2017). Subsequently, the terminal nucleoside of the oligo(U) tail is removed by the 3′–5′ exonuclease Usb1 leaving a 3′ phosphate group on the new terminal nucleotide (Mroczek et al. 2012; Shchepachev et al. 2012). The trimmed tail is then bound by the Lsm2–8 heteroheptameric protein ring (Mayes et al. 1999; Salgado-Garrido et al. 1999; Didychuk et al. 2017; Montemayor et al. 2018). Retention of mature U6 in the nucleus is dependent on binding of Lsm2–8 (Spiller et al. 2007a,b). In contrast, the 3′ tails of the spliceosomal snRNAs made by RNAP II instead bind to a paralogous, seven-protein Sm ring, which assembles on an AUUUUUG sequence internal to the 3′ terminus (Jones and Guthrie 1990; Salgado-Garrido et al. 1999).

To determine if synthesis of U6 snRNA by RNAP III is required for its function, we replaced the upstream and downstream sequences of *SNR6* with RNAP II control elements. Here we report that U6 snRNA synthesized by RNAP II (U6-II) is functional in vivo but is highly unstable. U6 snRNA made from a glucose-repressible variant of the U4 snRNA (*SNR14*) promoter has a half-life of about 15 min, as compared to an estimated half-life of more than 6 h for U6 made by RNAP III. We hypothesized that the instability of U6-II is due to inefficient recruitment of the Lsm protein complex. Consistent with this hypothesis, a tagged Lsm8 protein strongly exacerbates the growth defect associated with U6-II and addition of the Sm binding site of U4 snRNA to the 3′ tail of U6-II (U6-II-Sm) strongly stabilizes the RNA, increasing the half-life to about 2 h. Binding of the Sm ring to U6-II-Sm was confirmed by coimmunoprecipitation of free U6 RNA with the tagged Sm protein Smd1. We conclude that biosynthesis by RNAP II is compatible with yeast U6 snRNA function. The ability to control U6 snRNA synthesis with regulated RNAP II promoters expands the tools available to probe U6 biogenesis and function.

2.3 Results

Functional U6 snRNA can be made by RNAP II

SNR6 has an unusual RNAP III promoter structure, with upstream, intragenic, and downstream elements (**Fig. 2.1A**; Eschenlauer et al. 1993). To convert U6 snRNA into an RNAP II transcript, which we call U6-II, we replaced the upstream and downstream DNA flanking the U6 snRNA-coding region with the promoter and terminator from the RNAP II-transcribed U4 snRNA gene, *SNR14* (Steinmetz et al. 2001; Kuehner and Brow 2006). Our initial construct, named *SNR14-6-14*, contained *SNR14* sequence from −224 to +701 relative to the transcription start site (at position +1), but with the 160-nt U4 snRNA coding region replaced with the 112-nt U6 snRNA coding region (**Fig. 2.1A**). Note that the intragenic A block promoter element, while present in this construct, is not able to direct RNAP III transcription of the U6 gene in the absence of the downstream B block element (Brow and Guthrie 1990). This construct did not support viability in the absence of *SNR6* unless present on a high-copy plasmid.

Northern blot analysis of RNA produced in vivo from the high-copy *SNR14-6-14* construct indicated that most of the U6 transcript was 3′-extended, with the major species 250–300 nt in length (**Fig. 2.1B**, lane 3, “RCP”). This length corresponds to the known position of a Rnt1 cleavage site important for the 3′-processing of the pre-U4 snRNA (Allmang et al. 1999). Normally, the Rnt1 cleavage product of pre-U4 snRNA does not accumulate significantly, due to rapid 3′-trimming by the exosome. However, it is possible that 3′-trimming is inhibited in the context of the U6 snRNA coding sequence. To increase the yield of U6-II with the proper 3′ end, we deleted the first 135 nt of *SNR14* sequence downstream from the U6 snRNA coding region, which places the upstream Rnt1 cleavage site at position +112 of the U6 snRNA coding region. We named this construct *SNR6-II* (**Fig. 2.1A**; **Fig. 2.2**). When expressed from a high copy-number plasmid in yeast cells, the *SNR6-II* allele produced U6 snRNA of the correct size (U6-II), as well as a 3′-extended product corresponding in length to uncleaved primary transcript (**Fig. 2.1B**, lane 2).

To generate a regulated version of the *SNR6-II* construct with a potentially higher expression level, we deleted DNA upstream of the *SNR14* TATA box and replaced it with the 461 base pairs of DNA immediately upstream of the *GAL1* TATA box to create *GAL-SNR6-II* (**Fig. 2.1A**). This substitution put the *SNR6-II* construct under control of the galactose-inducible activator Gal4 (Johnston and Davis 1984). We found that the *GAL-SNR6-II* construct on a low copy-number plasmid rescues deletion of the chromosomal U6 gene, but only on growth medium containing galactose and not on medium containing glucose or raffinose, which are expected to repress or not induce Gal4-regulated genes, respectively (**Fig. 2.1C**).

To confirm expression or repression of the Gal4-regulated U6-II gene, we generated a merodiploid strain containing two different *SNR6* alleles: *GAL-SNR6-II* and a “pseudo-wild type” (Ψ-WT) U6 gene containing a shortened 5′ stem–loop and under

control of the native RNAP III promoter (Madhani et al. 1990). Ψ-WT U6 maintains yeast viability in the presence of raffinose or glucose and allows quantification of the levels of U6-II as the two RNAs are of different lengths. Primer extension analysis confirmed that no detectable U6-II was expressed from the *GAL-SNR6-II* construct in the presence of glucose or raffinose (**Fig. 2.1D**, lanes 2,3) and explains why yeast failed to grow under those conditions without Ψ-WT U6.

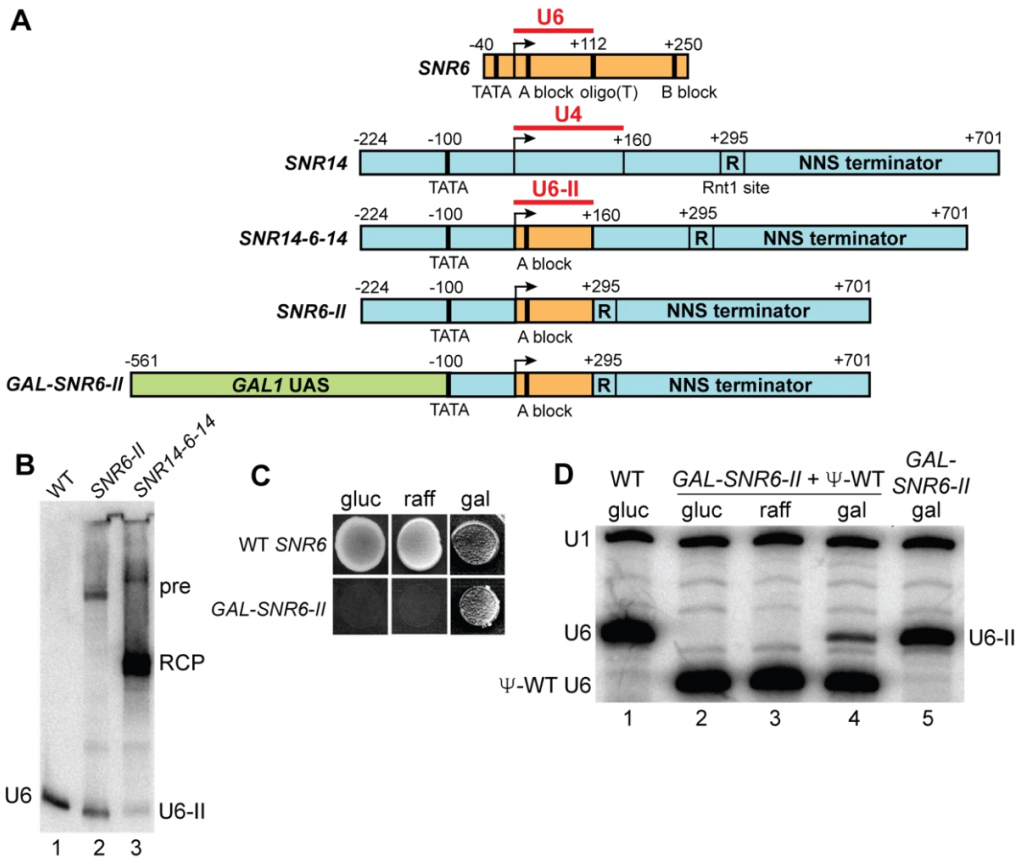

Figure 2.1 Expression of functional yeast U6 snRNA by RNAP II. (A) Diagram of construction of the *GAL-SNR6-II* allele. The wild type U6 (*SNR6*, orange) and U4 (*SNR14*, blue) snRNA genes are shown at *top*. Red lines span the mature RNA-coding regions. The *SNR14* Nrd1–Nab3–Sen1 (NNS) terminator is downstream from the Rnt1 cleavage site (R). The *GAL1* upstream activating sequence (UAS) is shown in green. Selected promoter and terminator elements are indicated by thick vertical lines. Numbers are base pairs upstream (−) or downstream (+) of the transcription start sites (bent arrow). **(B)** Northern blot of total cellular RNA from an *SNR6* disruption strain containing a low-copy plasmid with WT *SNR6* or a high-copy plasmid with the *SNR6-II* or *SNR14-6-14* allele, using a U6 probe. (pre) presumptive primary transcript, (RCP) presumptive Rnt1 cleavage product, (U6-II) U6 snRNA synthesized by RNAP II. **(C)** Growth of an *SNR6* disruption strain containing the indicated *SNR6* allele on a low-copy plasmid on YEP medium containing glucose (gluc), raffinose (raff), or galactose (gal). **(D)** Primer extension analysis of U1 and U6 snRNAs in a U6 disruption strain containing the indicated *SNR6* allele(s) on low-copy plasmid(s) and grown in YEP medium containing the indicated sugars. A shortened "pseudo-WT" (Ψ-WT) U6 allele is present in lanes *2–4* to provide U6 function in conditions where *GAL-SNR6-II* is repressed. Note that the level of U6-II is increased in the absence of Ψ-WT U6. U1 RNA serves as a normalization control.

Primer extension analysis also showed that the low copy-number *GAL-SNR6-II* construct produces about one-third the WT amount of U6 snRNA in the presence of medium containing galactose (**Fig. 2.1D**, cf. lanes 1,5). We showed previously that a similar amount of RNAP III-made U6 snRNA is sufficient for normal growth (Kaiser and Brow 1995). It is interesting that the steady-state level of U6-II (relative to U1 snRNA) is diminished in the presence of Ψ-WT U6 (**Fig. 2.1D**, cf. lanes 4,5). This observation suggests that either the two U6 genes compete for transcription factors, which should not be the case since they are presumably transcribed by different RNAPs, or that the RNAs compete for binding proteins that stabilize them. The latter is more likely since both should bind Prp24 and the Lsm2–8 ring (Montemayor et al. 2018). If U6-II is out-competed by Ψ-WT U6 for these proteins, then it may be destabilized and degraded leading to the lower steady-state level. Weaker binding of the Lsm2–8 ring to U6-II could be due to lack of 3′-end processing by Usb1 to produce a terminal 3′-phosphate or shortening of the U6 Lsm binding site from four to three or fewer uridines by Usb1. In either case the binding affinity of U6-II for Lsm2–8 would be expected to be reduced (Didychuk et al. 2017) and result in a competitive disadvantage for protein binding by U6-II. Importantly, the combined observations show that a functional U6 snRNA can be made by RNAP II using a galactose-inducible promoter. Thus, transcription by RNAP III is not required for U6 function.

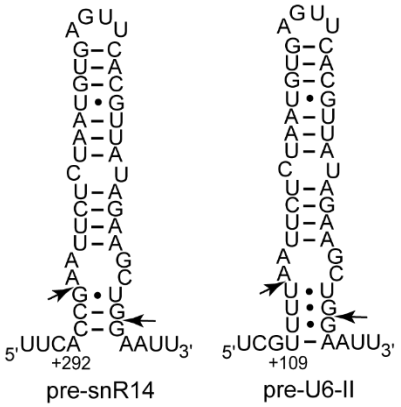

Figure 2.2 Predicted positions of Rnt1 cleavage sites (arrows) in the primary transcript 21 of *SNR6-II* (right) compared to the previously mapped cleavage sites in pre-U4 snRNA 22 (left; Allmang et al. 1999). The first 135 base pairs downstream of U6 position +112 were 23 deleted from the *SNR14-6-14* allele to align the mature 3' end of the shortest version of 24 yeast U6 snRNA with the upstream Rnt1 cleavage site (see **Fig. 2.1A**).

U6 snRNA made by RNAP II is methylguanosine-capped

Human U6 snRNA is O-methylated on the γ-phosphate group of its initiating GTP residue (Singh and Reddy 1989). It is not known if yeast U6 snRNA receives the same modification, but it is known that disruption of the recognition element for the human U6 capping enzyme in yeast U6 results in the acquisition of an RNAP II-like monomethylguanosine (MMG) 5′-cap in vivo, even when the U6 is produced by RNAP III (Kwan et al. 2000). This MMG cap is subsequently hypermethylated to a trimethylguanosine (TMG) cap. Since methylguanosine capping of RNAP II transcripts normally occurs cotranscriptionally, we would expect some fraction of U6-II to receive a

MMG cap, even though it has an intact determinant for a potential γ-O-methylphosphate cap. We assessed methylguanosine-capping of U6-II by antibody gel shift with the monoclonal antibody (mAb) H20 that recognizes both the TMG and, to a lesser extent, MMG cap (Bochnig et al. 1987). U4 snRNA acts as an internal control, and its mobility is efficiently shifted by binding of H20 mAb (**Fig. 2.3**, cf. lanes 1,2,5,6). RNAP III-made U6 exhibits no binding to H20 mAb, as expected (cf. lanes 3,4), but RNAP II-made U6 is partially shifted by H20 mAb (cf. lanes 7–10). The incomplete shifting of U6-II could be due to competing γ-O-methylphosphate capping, incomplete hypermethylation of MMG to TMG cap, or occlusion of mAb binding by the U6 5′ stem–loop (Kwan et al. 2000).

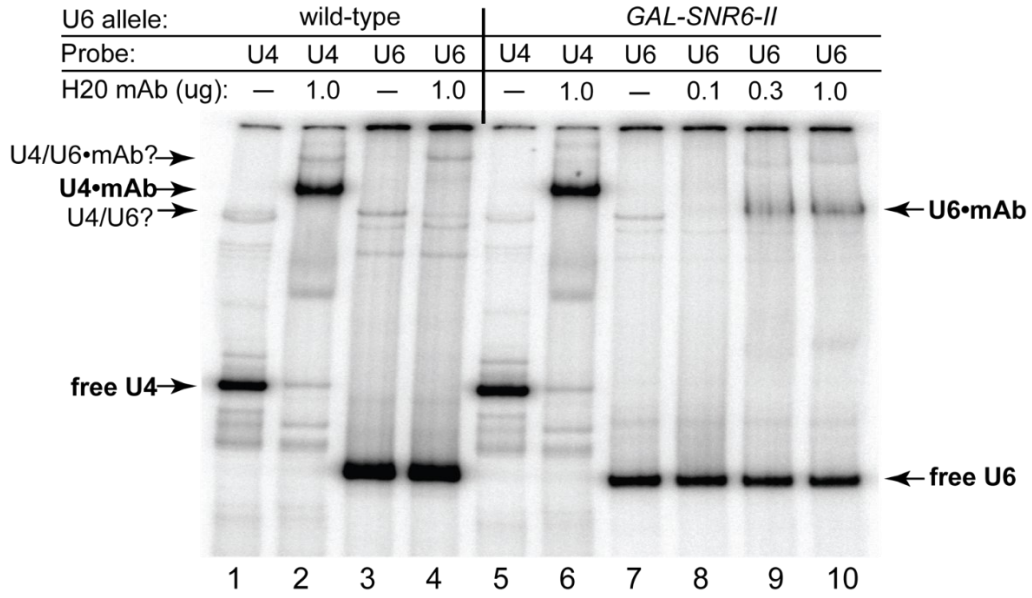

Figure 2.3 Evidence for methylguanosine-capping of U6 snRNA made by RNAP II. Total RNA from an *SNR6* disruption strain with pRS314-SNR6 (WT) or pRS314-GAL-SNR6-II (*GAL-SNR6-II*) was heated for 1 min at 90°C, then hybridized to 32P-labeled oligonucleotide probes specific for U4 or U6 snRNA. The RNA was then incubated with the indicated amount of mAb H20, which binds TMG or MMG caps, prior to electrophoresis on a nondenaturing gel. The positions of free U4 and U6 snRNAs, and mAb-shifted snRNAs are indicated. A minor band consistent with base-paired U4/U6 is also seen in both samples.

U6 snRNA made by RNAP II is highly unstable

Creation of a U6 gene under control of the *GAL1* UAS made it possible to shut off U6 snRNA synthesis by switching the carbon source from galactose to glucose, which allowed us to measure the in vivo half-life of U6-II. When transferred to glucose medium, yeast cells bearing the *GAL-SNR6-II* construct stopped growing in about 10 h (**Fig. 2.4A**), more rapidly than yeast cells bearing a Gal4-regulated U1, U2, or U5 RNA gene (Patterson and Guthrie 1987; Seraphin and Rosbash 1989). By analyzing U6 snRNA

levels at different times after shift of the *GAL-SNR6-II* strain to glucose, we infer a half-life of ~15 min for U6-II (**Fig. 2.4B,C**). It has not been possible to measure a precise half-life of RNAP III-made U6 RNA, but studies using the lac repressor to block RNAP III transcription (Luukkonen and Seraphin 1998) or a temperature-sensitive mutation in an RNAP III subunit (Kwan et al. 2000) suggest that RNAP III-made U6 has a half-life of greater than 6 h. Thus, U6-II is at least 20-fold less stable than U6 snRNA made by RNAP III.

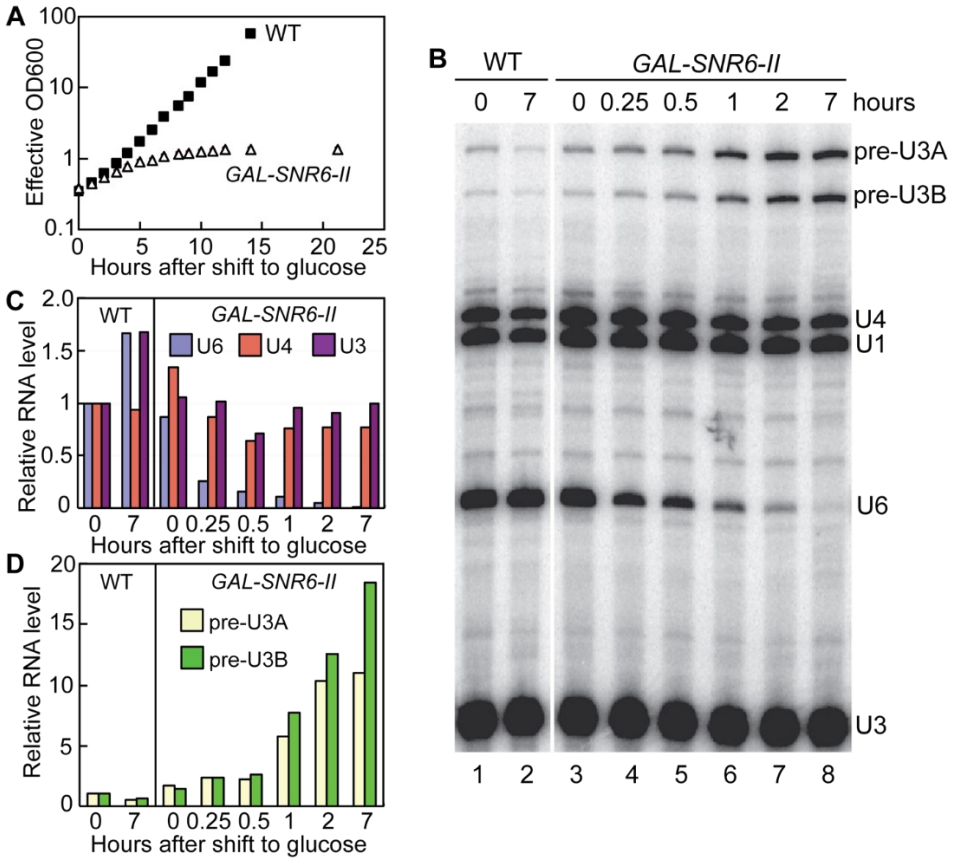

Figure 2.4 U6 snRNA made by RNAP II is unstable in vivo. A) Growth curve of an *SNR6* disruption strain containing pRS314-*SNR6* (WT) or pRS314-*GAL-SNR6-II* after replacement of YEP-galactose medium with YEP-glucose at time 0 to repress transcription of the *GAL-SNR6-II* allele. Both strains were diluted to keep the optical density at 600 nm (OD600) less than 1. The "effective OD600" value is corrected for this dilution. **B)** Total RNA extracted from cells collected at the indicated times was subjected to primer extension to detect U1, U4, and U6 snRNAs and U3 snoRNA, which is transcribed from two loci (A,B) as an intron-containing precursor (pre-U3). **C)** Quantitation of U6, U4, and mature U3 RNA present in the gel from panel *B* relative to the WT strain immediately before shift from galactose to glucose (0 h). Values are normalized to U1 snRNA in each sample to control for variable RNA recovery. **D)** Quantitation of pre-U3 snoRNA present in the gel from panel *B* relative to the WT strain immediately before shift from galactose to glucose (0 h). Values are normalized as in panel *C*.

We monitored splicing as a function of U6 snRNA abundance by measuring the level of pre-U3 snoRNA, which is made from two genes (SNR17A and SNR17B) that have spliceosomal introns of different lengths. In the presence of galactose, the steady-state level of pre-U3 RNA in the *GAL-SNR6-II* strain is roughly the same as in a WT strain (**Fig. 2.4B,D**). However, as soon as an hour after shift to glucose, an increase in pre-U3 is observed in the *GAL-SNR6-II* strain. (Because mature U3 snoRNA is very stable, little change in its level is observed over the time course.) Thus, the rapid kinetics of glucose repression and U6-II turnover result in quick disablement of the spliceosome in response to the change in medium. Interestingly, human U6atac snRNA is also very unstable and inhibition of its synthesis has been shown to decrease levels of mRNAs containing minor introns (Younis et al. 2013).

U6-II is stabilized by inclusion of an Sm protein-binding sequence

The spliceosomal snRNAs that are normally synthesized by RNAP II bind the heteroheptameric Sm protein ring, which contributes to their stability (Jones and Guthrie 1990). In contrast, U6 snRNA is bound and stabilized by the paralogous Like-Sm (Lsm) heteroheptamer (Mayes et al. 1999; Salgado-Garrido et al. 1999). Binding of the Lsm proteins to U6 snRNA may be coupled to RNAP III transcription via the La-homologous protein, Lhp1, which binds the oligo(U) tail of RNAP III transcripts (Pannone et al. 1998, 2001). In addition, the affinity of Lsm2–8 for U6 is determined by post-transcriptional processing of the 3′ end of U6 by Usb1, as mentioned previously. Therefore, it is not clear if U6-II would efficiently bind and be stabilized by the Lsm proteins (but see below).

To test if binding of Sm proteins could stabilize U6-II and is compatible with U6 snRNA function, we replaced 11 nt at the 3′ end of U6 with 14 nt from the 3′ end of U4, creating the *GAL-SNR6-II-Sm* allele (**Fig. 2.5A**). This substitution creates a consensus Sm binding site (A2U5G2) beginning at position 98 of the U6 snRNA but retains a four-uridine Lsm binding site at the 3′ end. A strain bearing the *GAL-SNR6-II-Sm* allele on a low-copy plasmid was viable, indicating that U6-II-Sm is functional. RNA analysis showed that U6-II-Sm is much more stable than U6-II, with a half-life of ∼2 h (**Fig. 2.5B,C**). Two hours after transfer from galactose to glucose medium about five times as much U6-II-Sm (relative to U1 snRNA) is present than U6-II, yet the U6-II-Sm strain accumulated at least as much unspliced pre-U3 snoRNA (**Fig. 2.5D**), suggesting that U6-II-Sm may not be as functional as U6-II. Thus, creation of an Sm-binding site in the 3′ tail of U6-II increases its stability but may decrease its function.

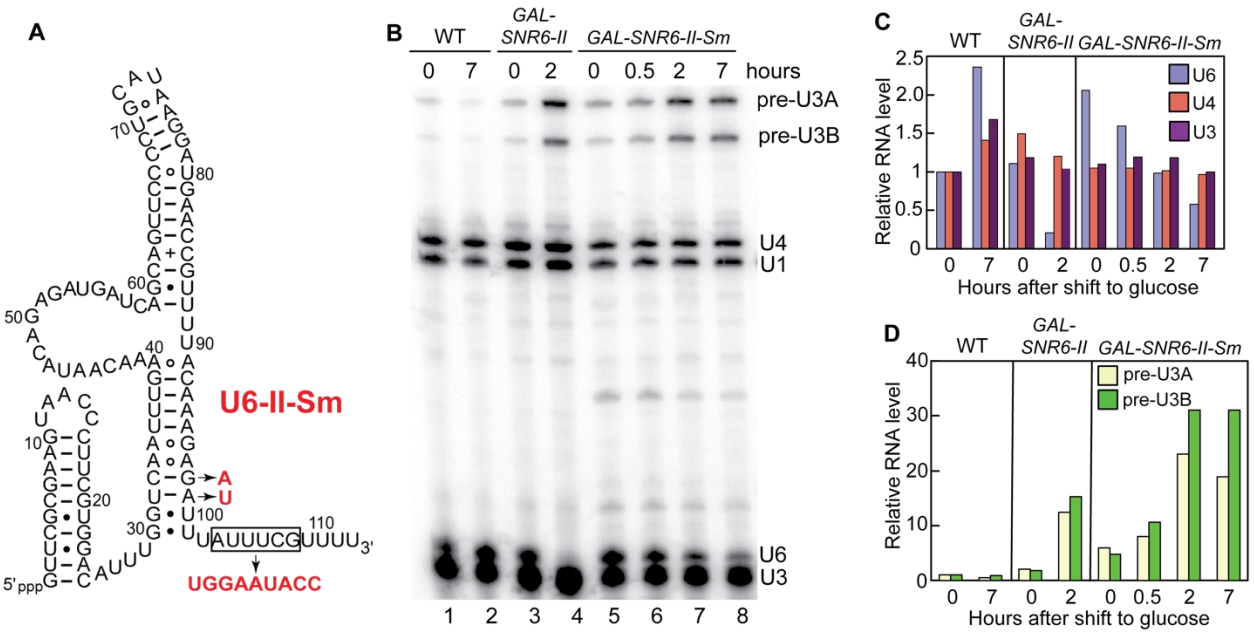

Figure 2.5 Creation of a consensus Sm-binding site in U6-II increases its stability. (A) Substitutions (in red) introduced in U6-II to replace its 3′ end with that of U4 snRNA, including the Sm-binding site ($A_2U_5G_2$). Replacement of 6 nt of U6 with nine of U4 elongate U6-II-Sm by 3 nt. **(B)** Total RNA was extracted from cells with WT U6, U6-II, or U6-II-Sm collected at the indicated times after transfer from YEP-galactose to YEP-glucose. The RNA was subjected to primer extension to detect U1, U4, and U6 snRNAs as well as unspliced and spliced U3 snoRNA. (Note that a different U6 primer, U6B, was used than in Fig. 2.4.) **(C)** Quantitation of U6, U4, and mature U3 RNA present in the gel from part B relative to the WT strain immediately before shift from galactose to glucose (0 h). **(D)** Quantitation of pre-U3 snoRNA present in the gel from part B relative to the WT strain immediately before shift from galactose to glucose (0 h). Values are normalized to U1 snRNA in each sample to control for variable RNA recovery.

Free U4 snRNP accumulates in the presence of U6-II

To examine how U6-II impacts the distribution of snRNAs among splicing complexes, extracts made from yeast expressing either WT U6, U6-II, or U6-II-Sm were sedimented through a glycerol gradient to resolve different snRNPs and spliceosomes. Primer extension of the RNA extracted from each gradient fraction was then used to identify the snRNAs present (**Fig. 2.6**). It should be noted that the haploid yeast strain used in these experiments also contained a deletion of the *SNR14* gene with U4 snRNA being provided from a low-copy plasmid, so U4 gene dosage could potentially increase under selective pressure.

U6 assembly into spliceosomes normally requires base-pairing with U4 to form a di-snRNP particle, which then binds the U5 snRNP to form the U4/U6.U5 tri-snRNP. Therefore, the U4 to U6 ratio is expected to be equal to one in fractions containing only tri-snRNP. Fractions at the bottom of the gradient (24 and 25) have U4/U6 ratios of less than one relative to tri-snRNP fractions, consistent with activated spliceosomes that have lost U4 snRNP (**Fig. 2.7**).

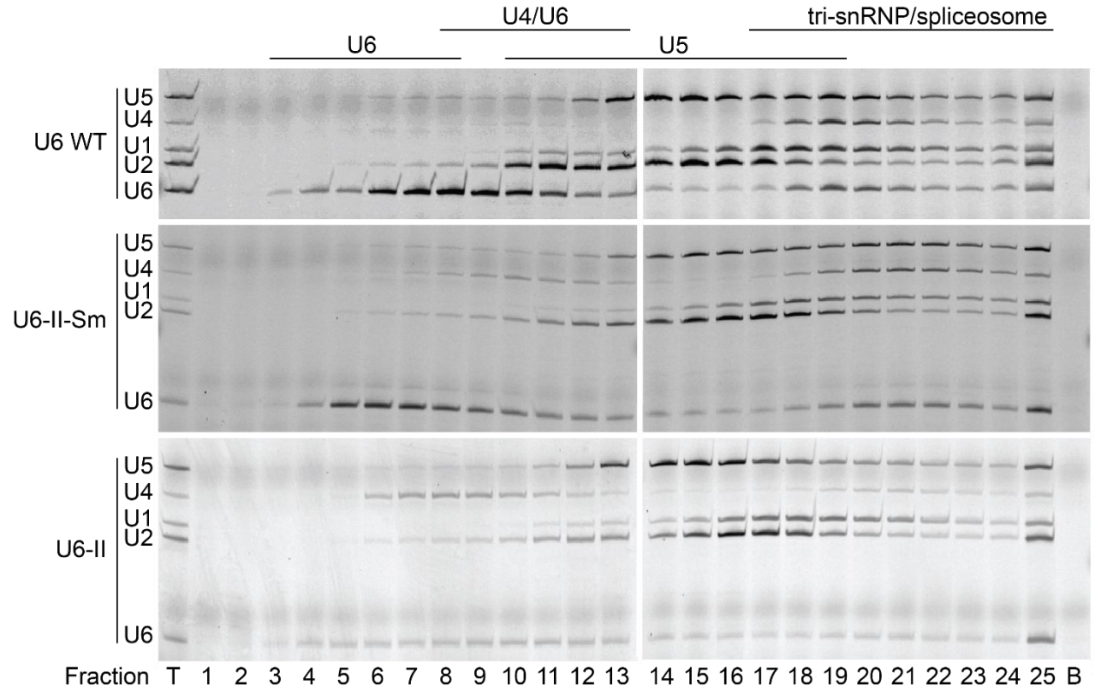

Figure 2.6 Primer extension of snRNAs from glycerol gradient fractions. RNAs were isolated from fractions taken from glycerol gradient sedimentation. Levels of primer extension products were quantitated as seen in Figure 2.7.

In cell extract with WT U6, there is a large excess of U6 over U4 and the U4 to U6 ratio is very low in fractions not containing tri-snRNP (**Fig. 2.7A,D**; **Fig. 2.6**). These results are consistent with previous studies of snRNP distribution in WT yeast extracts (Bordonné et al. 1990). In contrast, extract with U6-II exhibited increased levels of free U4 snRNP (**Fig. 2.7B,D**, fractions 6–10) and free U5 snRNP (**Fig. 2.7B**, fractions 13–18), indicating inefficient formation of U4/U6 di-snRNP and U4/U6.U5 tri-snRNP. U6-II extract also showed a large amount of U6 sedimenting at the bottom of the gradient (Fig. 5B, fraction 25). When the U6-II-Sm extract was analyzed, the snRNA distribution in gradient fractions much more closely resembled the distribution seen with WT U6 (cf. **Fig. 2.7A,C**).

Accumulation of U4 snRNA that is not base-paired to U6-II was also observed by solution hybridization of radiolabeled oligonucleotides to cold-extracted cellular RNAs followed by nondenaturing gel electrophoresis (**Fig. 2.8**). The total amount of U6 in the U6-II strain is less than the amounts present in either WT U6 or U6-II-Sm (**Fig. 2.8C**) and the U6-II extract contained a higher percentage of U4 not paired to U6 than either the U6-II-Sm or WT U6 extracts (**Fig. 2.8D**). Excess free U4 snRNA was also observed when RNAP III-made U6 levels were reduced by extragenic promoter mutations (Kaiser and Brow 1995), so reduced U6-II stoichiometry may be the main reason for accumulation of free U4. However, it is also possible that U6-II assembles less efficiently into U4/U6 di-snRNP, for example, due to decreased Prp24 or Lsm2–8 binding (Didychuk et al. 2015).

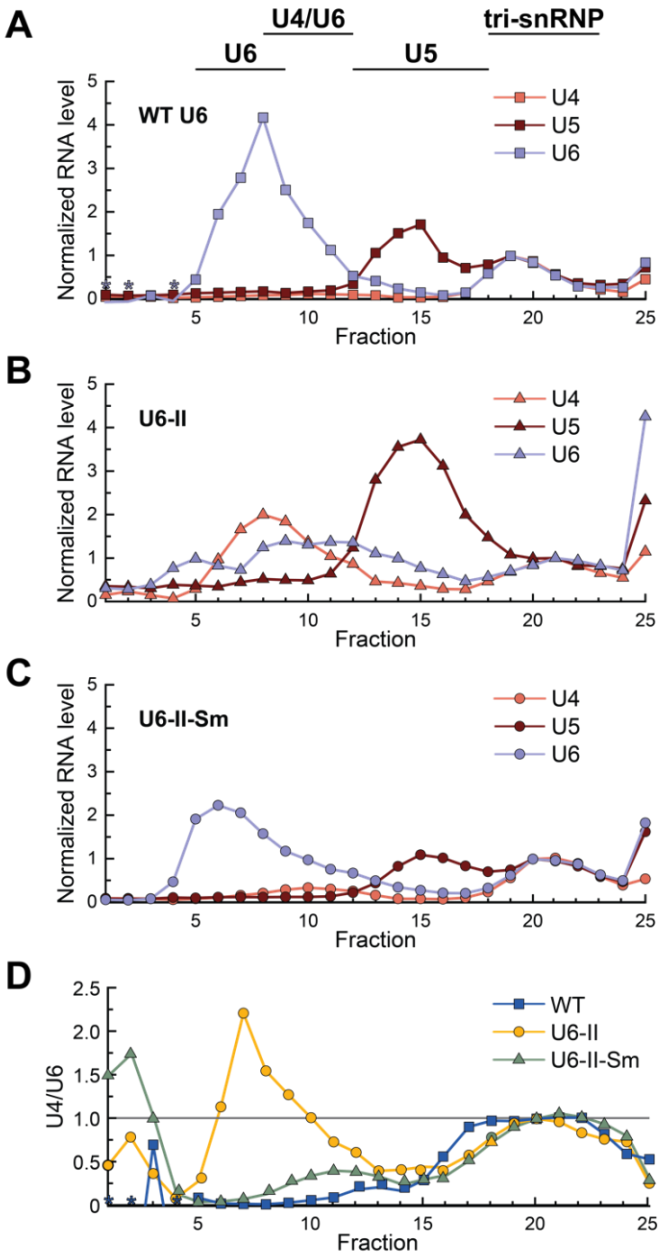

Figure 2.7 Glycerol gradient fractionation reveals altered snRNP distributions in U6-II extracts. Yeast cell extracts were sedimented through a 10%–30% glycerol gradient, fractionated, and snRNAs analyzed by primer extension. To normalize the data, the U4, U5, and U6 primer extension products in peak tri-snRNP-containing fractions for each gradient (20 or 21) were set to values of "1" since these RNAs are expected to be equimolar in tri-snRNPs. Asterisks denote fractions for which no detectable U4 RNA above background was detected. **(A–C)** Relative abundances of the U4 (red), U5 (brown), and U6 (blue) RNAs in each glycerol gradient fraction obtained using WT **(A)**, U6-II **(B)**, or U6-II-Sm **(C)** yeast extract. **(D)** The normalized U4/U6 ratio is plotted for each fraction from gradients of extracts prepared from WT (blue squares), U6-II (yellow circles), or U6-II-Sm (green triangles) strains.

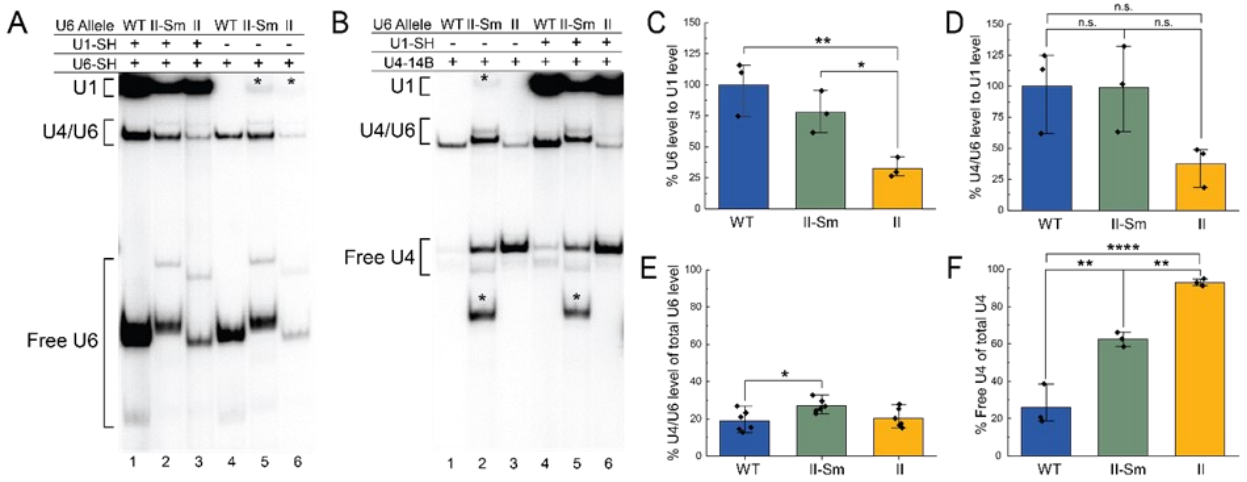

Figure 2.8 Free U4 snRNA accumulates and U4/U6 is diminished in yeast cells producing U6-II. A, B) Non-denaturing gel electrophoresis of cold-extracted whole-cell RNA from strains bearing the indicated U6 allele hybridized to probes against the indicated snRNAs. Positions of U4/U6 di-snRNA and free U1, U4, and U6 snRNAs are indicated. The bands indicated with asterisks are believed to be U6-II-Sm snRNA hybridized with the U4-14B probe due to inclusion of the U4 Sm binding site. **C, D, E, F)** Quantification of di-snRNA ratios and snRNA abundance from gel bands present in solution hybridization assays. di-snRNA ratios and relative abundance were calculated from three biological replicates. Band volume for U1 was adjusted to exclude contributions from a faint band in U6-II-Sm lanes running at the same height as U1. Sample means were compared with one-way ANOVA followed by a post hoc Tukey multi-pairwise analysis. (p = 0.05*, 0.01**, 0.001***, 0.0001****)

U6-II and U6-II-Sm are associated with Lsm proteins

It is possible that synthesis of U6 by RNAP II, with or without inclusion of an Sm protein binding site, could reduce or even eliminate Lsm2–8 protein binding. To test if this occurs, we first looked for genetic interactions between U6-II-Sm and Lsm proteins. Single-gene deletions of lsm5 and lsm8 are not lethal if U6 snRNA is over-expressed, consistent with the primary function of the Lsm2–8 proteins in yeast being stabilization of U6 snRNA (Pannone et al. 2001; Roth et al. 2018). We predicted that if Lsm5 or Lsm8 association with U6-II-Sm is not essential, then U6-II-Sm yeast cells should be viable in *lsm5Δ* or *lsm8Δ* strains after selection for loss of complementing URA3-marked plasmids. Instead, the lsm5 and lsm8 deletions were lethal in cells expressing U6-II-Sm (**Fig. 2.9**). This result suggests that the Lsm2–8 ring associates with U6-II-Sm snRNA. It should be noted that lsm5 and lsm8 deletions were viable in the presence of WT U6 (**Fig. 2.9**), likely due to amplification of the low-copy plasmid containing the U6 gene (Pannone et al. 2001; Roth et al. 2018). The stronger requirement for Lsm5 and Lsm8 observed with U6-II-Sm may have to do with its decreased function, noted above.

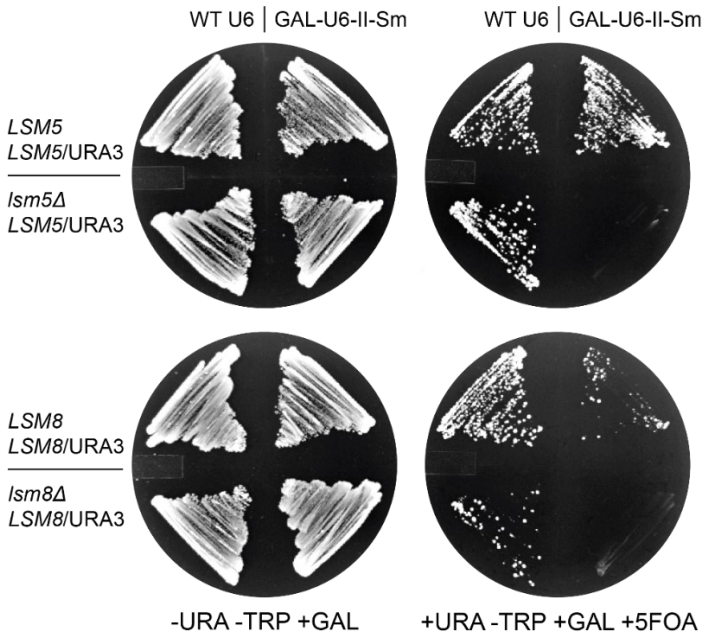

Figure 2.9 Genetic interaction between U6-II-Sm and Lsm Proteins. lsm5 or lsm8 genes were deleted from strains expressing either WT U6 or U6-II-Sm RNAs and complemented with Lsm5 or Lsm8 expression from a *URA3/CEN*-marked plasmid. These deletion strains or their corresponding controls were then grown on medium selecting for (-URA -TRP +GAL) or against (+URA -TRP +GAL +5FOA) the URA3-marked plasmid bearing the WT *LSM* gene. Deletion of the lsm5 or lsm8 genes and loss of the corresponding URA3-marked plasmid is lethal in the presence of U6-II-Sm but not WT U6.

To provide evidence for a physical interaction between the Lsm2–8 ring and U6-II or U6-II-Sm, we created yeast strains with an integrated, carboxyl-terminal 2xV5 epitope tag on the Lsm8 protein. The carboxyl terminus of Lsm8 is on the surface of the Lsm ring far from U6 and Prp24, so the tag is not expected to directly interfere with U6 binding or be occluded in the U6 snRNP (Montemayor et al. 2018). Indeed, carboxyl-terminally tagged Lsm8 was used previously to pull down yeast U6 snRNA-containing complexes (Salgado-Garrido et al. 1999; Fernandez et al. 2004). Neither the V5 tag on Lsm8 nor a carboxyl-terminal 3xFLAG epitope tag on the Smd1 protein altered the growth rate of the strain bearing WT U6 (**Fig. 2.10**). Strains expressing either U6-II or U6-II-Sm grow more slowly than those expressing WT U6 in the absence of tagged proteins, particularly at 37°C (**Fig. 2.10A,E**). Strikingly, however, the combination of U6-II with Lsm8-2xV5 causes a severe growth defect (**Fig. 2.10B,D**). The fact that the Sm binding site in U6-II-Sm substantially rescues this growth defect implies the Sm ring functionally replaces, at least in part, the Lsm2–8 ring for U6-II-Sm. In contrast, tagged Smd1 exhibits a less severe genetic interaction with U6-II-Sm (**Fig. 2.10C–E**), possibly due to a lesser effect of the tag on Sm ring function and/or Lsm2–8 ring binding to U6-II-Sm.

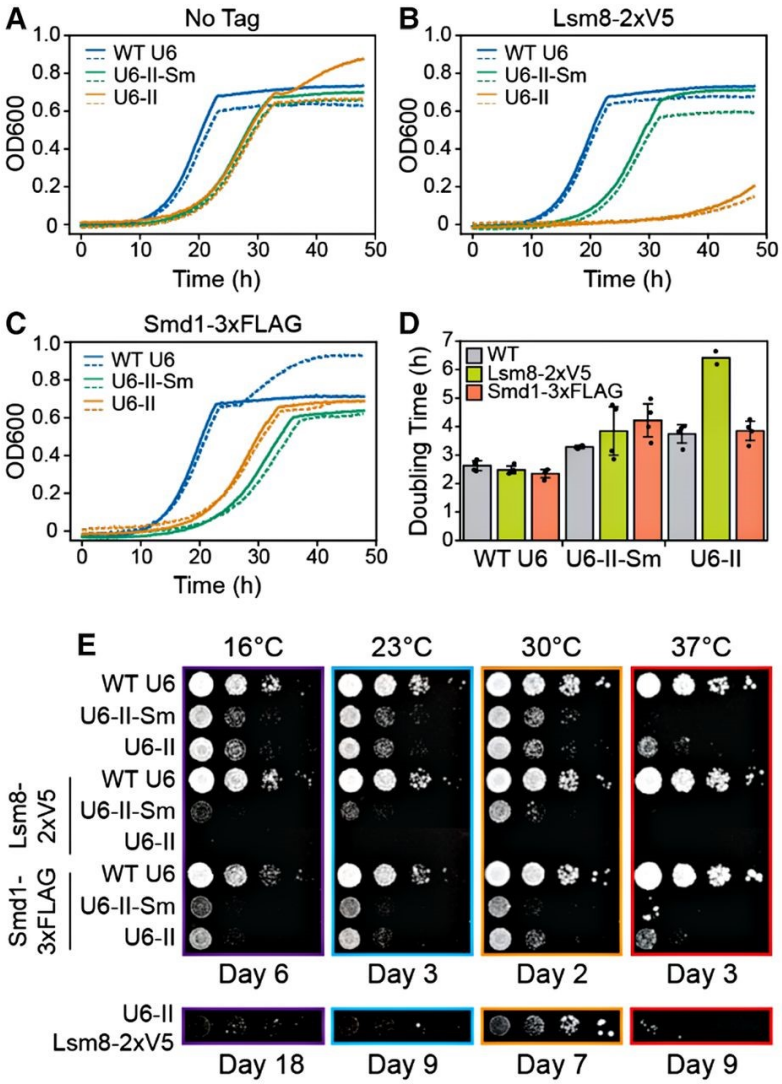

Figure 2.10 U6-II strains grow more slowly than WT and are more sensitive to tagging the Lsm2–8 and Sm rings. (A–D) Strains were grown in YPGal liquid media for 48 h at 30°C. Two technical replicates are shown as solid and dotted lines. Strains expressing WT U6, U6-II-Sm, and U6-II are separated by strain background based on the presence of (A) no epitope tagged proteins, **(B)** Lsm8-2xV5, and **(C)** Smd1-3xFLAG. **(D)** Calculated doubling times for strains shown in panels A–C. Plotted are the averages ± SD (bars) from four replicates (dots), except for U6-II/Lsm8-2xV5, which only has two technical replicates. **(E)** Yeast spotted onto YPGal solid media were incubated at the indicated temperatures and imaged on the indicated days.

Despite the growth defect of U6-II with Lsm8-2xV5, it was possible to obtain cell extract from this strain for immunoprecipitation (IP). In WT U6-expressing cells, the U4, U5, and U6 snRNAs but not the U1 and U2 snRNAs co-IP'ed with Lsm8-2xV5 (**Fig. 2.11**, lane 6). This result is consistent with Lsm2–8-bound U6 snRNA being a component of the U6 snRNP, U4/U6 di-snRNP, and the U4/U6.U5 tri-snRNP, while the Lsm ring is absent from activated or catalytic spliceosomes (Chan and Cheng 2005). U6-II and U6-II-Sm

snRNAs were also co-IP'ed with Lsm8-2xV5 (**Fig. 2.11**, lanes 7,8). In both cases, U4 and U5 but not U1 and U2 RNAs were also present, albeit in lower amounts. These data support binding of U6-II and U6-II-Sm by Lsm8, likely as part of the Lsm2–8 ring. Thus, it appears that U6 synthesis by RNAP III is not required for Lsm2–8 protein binding and that the inclusion of an Sm-binding site into U6-II-Sm does not prevent Lsm ring binding. However, the efficiency of U6-II-Sm IP by Lsm2–8 appears to be lower than for WT U6 or U6-II, raising the possibility that Sm ring binding competes for Lsm2–8 binding.

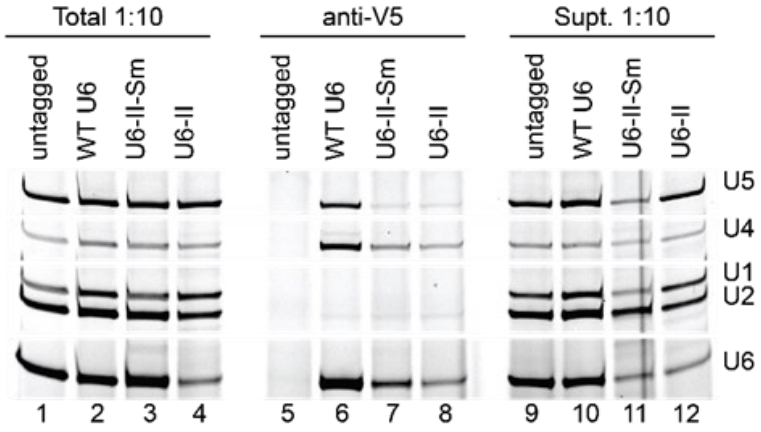

Figure 2.11 U6-II and U6-II-Sm co-IP with the Lsm8 protein. Extracts from Lsm8-2xV5 strains or a control lacking the 2xV5 epitope tag were immunoprecipitated using anti-V5 antibodies. RNAs from the immunoprecipitate, supernatant, or total extract were extracted using phenol, precipitated, and analyzed by primer extension.

U6-II-Sm binds Sm proteins

The stability conferred to U6-II transcripts by the addition of an Sm binding site suggests that U6-II-Sm is bound by Sm proteins. To investigate whether the Sm ring binds U6-II-Sm, we used anti-FLAG antibody to IP Smd1-associated RNAs from yeast strains with an integrated carboxyl-terminal 3xFLAG epitope tag on the Smd1 protein. Pulldown of WT U6 with Smd1 is possible due to its formation of complexes containing U1, U2, U4, and/or U5 with bound Sm rings. Not surprisingly, all five snRNAs co-IP'ed with Smd1 in all three strains (**Fig. 2.12A**). We quantified the ratio of U6 to U4 in the WT U6 sample and set the value equal to 1, since the RNAs should be stoichiometric in the U4/U6 and U4/U6.U5 snRNPs, with little free U4 snRNP present (**Fig. 2.7A**). In contrast, the U6 to U4 ratio in the U6-II-Sm IP (normalized to WT U6) was 1.5, indicating that about a third of the U6-II-Sm in the IP may originate from free U6 snRNP that contains Smd1 (**Fig. 2.12B).** Notably, the same calculation for U6-II IP gave a U6/U4 ratio of ~0.8, consistent with the presence of excess free U4 snRNP in this extract (**Fig. 2.7B**).

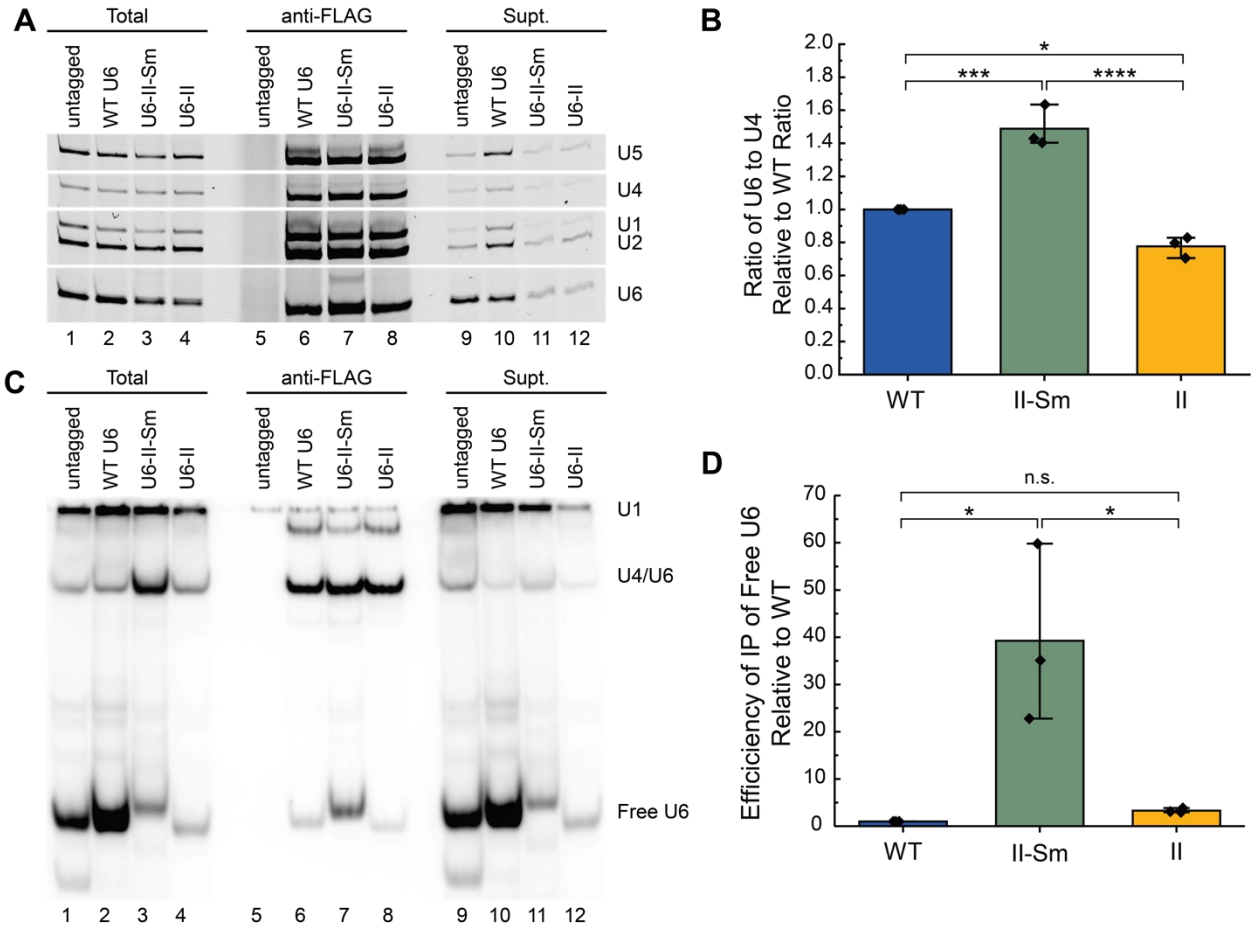

Figure 2.12 Free U6-II-Sm but not U6-II co-IPs with the SmD1 protein. (A) Extracts from SmD1-3xFLAG strains or a control lacking the 3xFLAG epitope tag were immunoprecipitated using anti-FLAG antibodies. RNAs from the immunoprecipitate, supernatant, or total extract were extracted using phenol, precipitated, and analyzed by primer extension. **(B)** Averages of the ratios of U6 to U4 in anti-FLAG IPs relative to WT samples from the experiment in panel A (n = 3). **(C)** Analysis of U6 base-pairing interactions with U4 snRNA. RNAs from the immunoprecipitate, supernatant, or total extract were isolated under conditions which maintain intact base-pairing. RNAs were then analyzed via hybridization with radiolabeled DNA probes complementary to U6 and U1 snRNAs. **(D)** Average efficiencies of free U6 IP relative to WT U6 and after normalization to U1 from the experiment in panel C (n = 3).

To confirm that increased co-IP of U6-II-Sm with Smd1 is due to Sm protein binding directly to the snRNA, we isolated RNAs from the co-IP under conditions that maintain U4/U6 base-pairing interactions to distinguish between U6 snRNAs co-IP'd as part of U4/U6 di-snRNA versus free U6. U1, U4/U6, and U6 were detected and quantified by solution hybridization and native gel electrophoresis. As expected, IPs from all three extracts contained base-paired U4/U6 (**Fig. 2.12C**, lanes 6–8). We quantified the efficiency of free U6 snRNP IP (anti-FLAG/Total) relative to WT after normalizing to U1 snRNA because the amount of free U6 snRNP varied greatly between the extracts (**Fig.**

2.12C, lanes 2–4). Anti-FLAG-Smd1 samples from strains expressing U6-II-Sm showed an ~40-fold higher efficiency of free U6 (U6 snRNP) IP compared to WT while those from strains expressing U6-II showed no significant increase over WT (**Fig. 2.12C,D**). Furthermore, the average ratio of U4/U6 to free U6 in the U6-II-Sm Smd1 IP sample is 1.8 (**Fig. 2.12,** lane 7), close to the 2:1 ratio predicted from Figure 2.12B. Thus, we conclude that the strong stabilization of U6-II by addition of an Sm-binding site is due to direct binding of the Sm ring.

2.4 Discussion

Eukaryotic RNAPs I, II, and III are specialist enzymes that transcribe specific sets of genes, each with defined characteristics and unique downstream post-transcriptional processing steps. Few examples exist in the literature of genes that are actively transcribed by different polymerases in different organisms or have been manipulated to switch transcription from one polymerase to another (Gunnery and Mathews 1995; Dergai and Hernandez 2019). Herein, we have shown that an snRNA central to pre-mRNA splicing, U6, is unexpectedly adaptable in both the route of its biosynthesis and its cellular interaction partners. We have shown that functional U6 can be synthesized by RNAP II under direction of the U4 snRNA promoter and terminator sequences in *S. cerevisiae*. However, high instability of the U6-II transcript was observed. Addition of an Sm binding site to U6-II increased stability, likely as the result of Sm ring binding. Uncoupling U6 synthesis from RNAP III allows for investigation of processing steps solely dependent on recognition of elements within the U6 snRNA rather than dependence on RNAP III transcription and provides the ability to control U6 snRNA synthesis via regulated RNAP II promoters.

Synthesis of U6 snRNA by RNAP II

The main challenge in converting U6 snRNA from an RNAP III transcript to an RNAP II transcript was proper 3′ end processing. There was no difficulty in generating the correct 5′ end of U6-II with the *SNR14* promoter, since the U4 gene is known to have a near-optimal initiator sequence that directs initiation predominantly at a single site (Kuehner and Brow 2006). Obtaining adequate levels of transcription was not a problem, even though the steady-state level of yeast U4 snRNA is more than twofold lower than U6, since we were able to increase expression with a strong upstream activating sequence from the *GAL1* gene. The surprising result was that, despite apparently normal recognition of the *SNR14* NNS terminator and Rnt1 cleavage site, the cleaved intermediate was not efficiently processed by the nuclear exosome. The reason for accumulation of the 3′-extended processing intermediate is not clear but could be due to interaction of the downstream region with U6 sequences in the same transcript. This

hypothesis could be examined by selecting for cis-acting mutations that improve mature U6 snRNA production from the *SNR14-6-14* allele.

Even when the 3′ processing defect was corrected and the promoter strength increased via construction of the *GAL-SNR6-II* allele, the steady-state level of U6-II is lower than for WT U6. An important contributor to this lower expression is a much shorter half-life of U6-II. The further decrease in steady-state U6-II level in the presence of RNAP III-made Ψ-WT U6 RNA (**Fig. 2.1D**) suggests that U6-II competes poorly for limiting U6 snRNP proteins. Since the U6-II sequence is the same as WT-U6, the most likely explanation is decreased Lsm2–8 ring recruitment due to differences in 3′ end structure. Indeed, U6-II appears to have slightly faster mobility than WT-U6 on a native gel (**Fig. 2.3**) consistent with missing a terminal nucleoside and/or 3′ phosphate, either of which is expected to weaken Lsm2–8 binding (Montemayor et al. 2018). The strongly negative genetic interaction between U6-II and carboxyl-terminally tagged Lsm8 (**Fig. 2.10**) is further evidence of weak Lsm2–8 binding. Thus, yeast U6 snRNA evolved for optimal synthesis by RNAP III is maladapted for synthesis by RNAP II.

Formation of a U6•Sm snRNP

The short half-life of U6-II is substantially improved by addition of an Sm ring binding site. The U6-II-Sm snRNA is bound by both Lsm and Sm proteins but we do not know if it can be bound by both the Lsm and Sm complexes simultaneously, if one complex is exchanged for the other during U6 biogenesis or splicing, or if one complex is more likely to be involved in catalytically competent spliceosome formation over the other. Nonetheless, this observation indicates flexibility in what proteins interact with the 3′ region of U6. A crystal structure of the WT U6 snRNP reveals a 660 Å2 binding interface between Lsm2 and the core U6-binding protein Prp24, as well as binding of the carboxyl-terminal peptide of Prp24 to Lsm5 and Lsm7 (Montemayor et al. 2018). Furthermore, the Lsm2–8 ring strongly enhances Prp24-dependent annealing of U4 and U6 in vitro (Didychuk et al. 2015). The Sm ring presumably lacks these interactions with Prp24, so it is possible that it serves primarily to stabilize free U6 until it can be exchanged with Lsm2–8 for di- and tri-snRNP assembly.

Sequential exchange of Sm for Lsm protein complexes on an RNA has previously been observed for the *S. pombe* telomerase RNA, TER1 (Tang et al. 2012). In this case, TER1 RNA initially is bound by Sm proteins, which facilitate its cleavage by the spliceosome and hypermethylation of its RNA cap. The Sm ring is then replaced by the Lsm ring for assembly of TER1 into telomerase RNPs. Whether or not the Sm and Lsm complexes need to be exchanged to make functional spliceosomes is not yet clear, nor have we excluded the possibility of simultaneous occupancy of U6 by both complexes. Inspection of the yeast U6 snRNP structure suggests it is possible that the Sm and Lsm rings can bind U6-II-Sm simultaneously as the Sm binding site should not be occluded by

the Lsm ring or U6 snRNA secondary structure. However, the Sm ring would potentially block the Lsm2–8 ring's interaction with Prp24.

It is possible that the Sm ring provides 3′ end protection during nuclear-cytoplasmic shuttling of U6-II-Sm. Although human U6 snRNA is roughly equally distributed in nuclear and cytoplasmic fractions of human cells (Pessa et al. 2008; Mabin et al. 2021), it was thought that this distribution is an artifact of fractionation and that U6 is exclusively nuclear in live cells. However, heterokaryon studies in yeast cells provided early evidence for nucleocytoplasmic shuttling of U6 snRNA (Olson and Siliciano 2003). More recently, additional evidence for a cytoplasmic phase of U6 snRNP assembly in yeast has been obtained (Becker et al. 2019). Although the specifics of a U6 shuttling pathway are not yet fully determined, it is possible that the Sm binding site on U6-II-Sm perturbs this pathway, for example, by sequestering U6-II-Sm in the cytoplasm. Such an effect could explain the fact that the substantial stabilization of U6-II by addition of an Sm site does not result in increased function as measured by growth rate or pre-U3 snRNA splicing.

The Lsm ring on WT U6 is released during formation of the spliceosome active site to allow binding of the NineTeen Complex (NTC) (Townsend et al. 2020). If Sm or Lsm complexes are retained on U6-II-Sm snRNAs during splicing, it is likely that this can only be accomplished by major remodeling of essential splicing factors and formation of spliceosome architectures distinct from those currently known. For splicing to proceed normally, we would therefore expect that the Sm ring on U6-II-Sm would also be released during spliceosome activation. The Sm rings on U1, U2, U4, and U5 snRNAs are not released at any point during splicing nor is it clear how the Sm•snRNA interaction would be disassembled. It is intriguing to speculate that if U6-II-Sm RNPs are incorporated into spliceosomes, the same mechanism responsible for Lsm release can also dissociate a U6•Sm complex (or both U6•Sm and U6•Lsm interactions if both rings are simultaneously present). Inefficient dissociation of the U6•Sm interaction during activation could also contribute to the reduced splicing efficiency we observed for the U3 pre-mRNAs.

Conclusions

It remains to be determined what evolutionary advantage, if any, is conferred by RNAP III specialization of U6 snRNA synthesis over RNAP II. It is possible that transcription by RNAP III is a relic of the mechanism by which U6 snRNA was originally acquired by the eukaryotic genome from Group II self-splicing introns (Novikova and Belfort 2017). A relevant parallel is that all yeast small nucleolar RNAs are synthesized by RNAP II except snR52, which is synthesized by RNAP III (Harismendy et al. 2003; Moqtaderi and Struhl 2004). Integration of transcription-factor-like subunits into the RNAP III core during evolution may have promoted more efficient and rapid transcription, facilitating production of the highly abundant ncRNAs synthesized by RNAP III (Barba-Aliaga et al. 2021).

Regardless of the unanswered evolutionary questions, our system for synthesis of U6 by RNAP II in yeast provides a useful tool for dissecting how cis-acting sequences within U6 shape its biogenesis without the transcriptional restraints imposed by RNAP III regulatory elements found within the *SNR6* gene. For example, U6 synthesis independent of the A block internal promoter element will enable sequence manipulation of the 5′ stem–loop in vivo, and rapid induction or repression of U6-II synthesis using the galactose-inducible promoter will allow conditional expression of mutant U6 alleles. As the 5′ stem–loop is the sole region of U6 with no known role in spliceosome function, it is a promising target for engineering affinity tags or binding sites for fluorophores or fluorescent proteins.

Finally, this study may help explain why RNAP II is depleted from the *SNR6* locus in a Sen1-dependent fashion (Steinmetz et al. 2006). *SNR6* has a consensus TATA box element that could potentially direct transcription by RNAP II. Our results indicate that RNAP II transcripts of *SNR6* are unlikely to be correctly processed and, even if they are, are unlikely to be stable. Given that U6 snRNA is essential, there is expected to be a strong selective pressure in favor of mechanisms that prevent misrecognition of *SNR6* by RNAP II.

2.5 Materials and Methods

Yeast strains and growth

Yeast strains used in these experiments are described in Table 2.1. Yeast were grown in standard media as indicated in the figure legends (Treco and Lundblad 1993). Plasmid transformations and shuffling assays were carried out using standard procedures as described (Sikorski and Boeke 1991).

Strains carrying either a carboxyl-terminal 2xV5 epitope tag on Lsm8 or a carboxyl-terminal 3xFLAG epitope tag on Smd1 were created using PCR-based tagging (Janke et al. 2004). Briefly, synthetic, linear DNA fragments coding for the 2xV5 or 3xFLAG epitope tags (IDT) were digested with HindIII and SalI and then ligated into those same restriction sites of a digested pAG32 plasmid, which also contained an expression cassette for the yeast hygromycin resistance (HygR) marker (Goldstein and McCusker 1999). The resultant plasmids were then used as templates for PCR to prepare linear DNAs for gene tagging by homologous recombination. After transformation of the DNA fragments, yeast containing the integration cassettes were selected by growth on YPD plates containing hygromycin (300 µg/mL; Sigma Aldrich). Colonies were screened by PCR to confirm proper integration of the tag at the correct chromosomal location.

Temperature growth assay

Overnight cultures of yeast grown in YPGal were adjusted to OD600 = 0.5, used to make 10-fold serial dilutions, and then stamped on YPGal plates. Plates were incubated at the indicated temperatures and imaged after the indicated number of days.

Calculation of yeast doubling times

Overnight liquid cultures of yeast in YPGal were diluted to OD600 = 0.02 in YPGal and grown at 30°C for 48 h in a Tecan plate reader (Infinite F200 PRO). Doubling times were calculated by first subtracting the background and then carrying out a linear fit of the log-transformed data between OD600 = 0.06 and 0.3 except for U6-II/Lsm8-2xV5 in which data were fit from between OD600 = 0.06 and the maximal OD reached after 48 h.

Oligonucleotides

The sequences for the DNA oligos used for cloning, primer extension, northern blotting, and solution hybridization assays are shown in Table 2.2.

Plasmids

Plasmids used in these experiments are listed in Table 2.3. To make pRS314-SNR14-6-14, the *SNR14* promoter from pRS313-SNR14 (Kuehner and Brow 2006) was amplified by PCR with primers PsnR14(223)-Xho-F and U4–U6-R, and the XhoI/NruI-digested product was ligated to XhoI/NruI-digested pRS314-554H6 (Martin et al. 2001) to create pRS314-SNR14-6. Then, recombinant PCR (Horton et al. 1993) was used to

replace the *SNR6* downstream DNA with *SNR14* downstream DNA. The first round of PCR used primers PsnR14(223)-Xho-F and U4(12)-U6(24)-R on pRS314-SNR14-6, and primers U6(12)-U4(23)F and snR14-BamHI-701R on pRS313-SNR14. These products were then joined and amplified by PCR with PsnR14(223)-Xho-F and snR14-BamHI-701R, and the digested product was ligated into XhoI/BamHI-cut pRS314.

pRS314-SNR6-II was created by recombinant PCR as described above but using primer pairs PsnR14(223)-Xho-F/Rnt1(12)-U6(24)R and U6(12)-Rnt1(23)F/snR14-BamHI-701R in the first round with pRS314-SNR14-6-14 as template.

pRS314-GAL-SNR6-II was created by recombinant PCR. The first round of PCR used primers GAL1-XhoI-608F and U4(14)-GAL1(24)R to amplify the *GAL1* upstream region from genomic DNA prepared from yeast strain PJ43-2b (Kaiser and Brow 1995), and primers Gal(11)-U4(25)F and snR14-KpnI-701R to amplify pRS424-SNR6-II. These products were then joined and amplified by PCR with primers GAL1-XhoI-608F and snR14-KpnI-701R, and the digested product was ligated into XhoI/KpnI-cut pRS314.

pRS314-GAL-SNR6-II-Sm was created by recombinant PCR using the following primer pairs and pRS314-GAL-SNR6-II as template: GAL1-XhoI-608F/Sew-U6::U4R and Sew-U6::U4-F/snR14-KpnI-701R. High-copy versions of these plasmids were made by subcloning the inserts into pRS424.

Northern blotting and primer extension

Total cellular RNA was isolated by phenol extraction in the presence of guanidinium thiocyanate (Wise et al. 1983). Northern blot analysis used 25 µg of RNA per lane of a 20 cm × 20 cm × 1.5 mm 6% polyacrylamide, 8.3 M urea gel run in 50 mM TBE buffer at 400 V for 1.5 h. The gel was washed 2 × 15 min in Transfer buffer (12 mM Tris acetate, pH 7.6, 0.3 mM EDTA) at 4°C and transferred overnight to a Zeta-Probe nylon membrane (Bio-Rad) at 20 V and 4°C in a Hoefer Transphor TE 42 unit. The blot was hybridized to a 32P-labeled oligonucleotide complementary to U6 RNA (U6D, **Table 2.3**) in Rapid-hyb buffer (Amersham) according to the manufacturer's instructions.

Primer extension analysis used 4 µg total cellular RNA per reaction combined with 200 fmol of each 5′-32P-labeled oligonucleotide (U1-SH, U4-14B, U3 21-mer, U6D, or U6B) in a total volume of 2.5 µL of Annealing Buffer (10 mM TrisCl, pH 8.0, 250 mM KCl). Mixtures were incubated at 90°C for 3 min, on ice for 3 min, and at 45°C for 5 min. 6.5 µL of RT mix (35 mM TrisCl, pH 8.0, 11.5 mM MgCl2, 11.5 mM DTT, 0.6 mM each dNTP, 4.5 U AMV reverse transcriptase) was added and the reaction incubated at 45°C for 15 min. An amount of 2.5 µL of formamide loading buffer was added and the mixture heated at 90°C for 3 min before loading on a 6% polyacrylamide (19:1) gel with 8.3 M urea and 50 mM TBE.

TMG cap antibody gel shift

Ten µg of total cellular RNA per lane were analyzed as described previously (Kwan et al. 2000) using 32P-labeled oligonucleotides U6D or U4-14B.

Nondenaturing RNA isolation

All steps were performed on ice with cold solutions. Ten ODs of log phase culture (at 0.6–0.8 OD600) was added to cold MilliQ water up to 50 mL. Cells were pelleted and washed with RNA Buffer (200 mM TrisCl pH 7.5, 100 mM EDTA, and 500 mM NaCl) before resuspension in 300 µL RNA buffer. An amount of 200 µL 425–600 micron acid-washed glass beads (Scientific Industries) and 300 µL 1:1 phenol/CHCl3 were added and cells were lysed by vortexing twice at high speed for 1 min with a 1 min rest on ice between. The aqueous phase was extracted again with an equal volume of 1:3 phenol/CHCl3 and precipitated with three volumes 100% ethanol at −20°C. RNA was pelleted and washed with 70% v/v ethanol before resuspension in RNase-free water.

Solution hybridization

DNA probes (U1-SH, U4-14B, and U6-SH) were end labeled with [γ-32P]-ATP (PerkinElmer) using T4 PNK (Thermo Scientific) for 1 h at 37°C followed by enzyme inactivation by heating to 60°C for 20 min according to the manufacturer's instructions. Probes were hybridized to 5 µg total RNA at 37°C for 15 min in Solution Hybridization Buffer (50 mM Tris pH 7.5, 1 mM EDTA, and 150 mM NaCl) before mixing 1:1 with 50% v/v cold glycerol (22 µL total volume). Twenty microliters of the reaction was loaded onto a 20 cm × 22 cm × 1 mm 9% nondenaturing polyacrylamide gel in 50 mM TBE. RNA:DNA hybrids were separated for 5 h at 300 V at 4°C. The gel was dried under vacuum for 30 min before exposure to a phosphor screen overnight followed by imaging (Li and Brow 1993).

Measurement of cell growth and U6 RNA stability under repressing conditions

Yeast cells were grown at 30°C in YEP with 2% galactose to OD600 = 0.5–1.0, pelleted, washed with sterile water, and resuspended to the same density in YEP with 2% glucose. Growth curves were generated by continually diluting cultures as required to keep the OD600 between 0.2 and 1.0 (Patterson and Guthrie 1987). Aliquots (50 mL) of the culture were collected at the indicated times to prepare RNA for primer extension.

Analysis of snRNPs by glycerol gradients

Yeast whole cell extracts (yWCEs) were prepared using the liquid N2 method and processing with a ball mill (Mayas et al. 2006). yWCEs in dialysis buffer (25 mM HEPES-KOH pH 7.9, 50 mM KCl, 20% v/v glycerol, 0.5 mM DTT) were adjusted to 8% v/v glycerol and 500 µL of the dilution was layered onto 12.5 mL 10%–30% (v/v) linear glycerol gradients (Gradient Master) in 20 mM HEPES pH 7.9, 150 mM KCl, 1 mM Mg(OAc)2, 0.1% v/v NP40, 0.5 mM PMSF, 1 µg/mL Pepstatin A, and 1 µg/mL Leupeptin. Sedimentation was performed in a Beckman SW40 rotor for 22 h at 35,000 rpm, 4°C. Fractions of 500 µL were collected from the top.

RNA was isolated from glycerol gradient fractions with a 25:24:1 phenol/CHCl3/isoamyl alcohol extraction and precipitated for 30 min at −20°C with 0.1 volume 3 M NaOAc, 10 µg tRNA, and 2.5 volumes 100% EtOH. Precipitated RNA pellets were washed with 96% EtOH and resuspended in 8 µL dH2O. Total RNA samples were prepared from 50 µL yWCE resuspended in an equal amount of Splicing Dilution Buffer (100 mM TrisCl pH 7.5, 10 mM EDTA pH 8.0, 1% SDS (w/v), 150 mM NaCl, 0.3 M NaOAc pH 5.2–5.4) and 100 µL dH2O. After incubation at 37°C for 15 min, 200 µL dH2O was added to samples. Total RNA was extracted with 400 µL 25:24:1 phenol/CHCl3/isoamyl alcohol four times and precipitated as described above. Pellets were resuspended in 32 µL dH2O. Primer extension of isolated RNA was accomplished with the qScript Flex cDNA Synthesis Kit using 8 µL volumes with 5′ labeled Dy682 primers (U1RT136, U2RTALL124, U4RTALL, U5, U6D). Amplified cDNA products were separated on a 7% denaturing PAGE gel for 1 h at 22 W and imaged on an Amersham Typhoon NIR laser scanner (Cytiva).

Immunoprecipitation

yWCEs for IP were prepared using the liquid N2 method and processing with a ball mill (Mayas et al. 2006) without dialysis. A 20 µL volume of protein G magnetic beads (BioRad SureBeads #1614023) per IP reaction was washed three times with 1.5 mL cold IPP150 (150 mM NaCl, 0.1% v/v NP-40 Surfact-Amps Detergent Solution, 10 mM Tris pH 8.0, and 0.1% w/v azide) and three times with 1.25 mL cold IPP500 (500 mM NaCl, 0.1% NP-40 Surfact-Amps Detergent Solution, 10 mM Tris pH 8.0, and 0.1% w/v azide). Beads were coupled with 4.8 µg of either anti-V5 (Thermo Fisher Scientific, R960-25) or anti-FLAG (Millipore Sigma, F1804) antibody overnight at 4°C in 100 µL total volume of magnetic bead suspension in IPP150.

After incubation, beads were washed three times with 1.5 mL of IPP150 and resuspended in 80 µL IPP150. IP reactions contained the magnetic bead suspension, 40 µL yWCE, 1.25 µL RNasin Plus Ribonuclease Inhibitor (Promega, #N2611), 2.5 mM DTT, and IPP150 to 200 µL total volume. Samples of total RNA were assembled without beads and incubated as described. Reactions were agitated overnight at 4°C. Supernatant from reactions was saved and beads were washed five times in IPP150 containing 2.5 mM DTT. Protein bound RNA was released by Proteinase K incubation (0.8 µg/µL Proteinase K [Thermo Scientific], 50 mM Tris pH 7.5, 10 mM EDTA, 1% w/v SDS, and RNase-free water to 200 µL) at 37°C for 30 min. The supernatant was transferred to a fresh tube and 200 µL of IPP150 with 0.1 mg/mL E. coli tRNA (Roche) was added. To the IP supernatant and total RNA samples, 200 µL IPP150 containing 10 mM EDTA, 0.5% w/v SDS, and 0.1 mg/mL E. coli tRNA was added.

RNA was extracted once with phenol:chloroform and precipitated with 300 µL 3M NaOAc pH 5.2 and 1 mL ice cold 100% EtOH overnight at −80°C. RNA was pelleted for 30 min at 20,800 × g, 4°C and washed with 70% EtOH. Dried pellets were resuspended in 10–12 µL volumes of RNase-free water. Immunoprecipitated RNAs were reverse

transcribed into cDNA using SuperScript III Reverse Transcriptase (Thermo Fisher Scientific, # 18080044) according to the manufacturer's instructions using 0.2 pmol of each primer (U1RT136, U2RTALL124, U4RTALL, U5, U6B labeled with 5′ IRDye 700, IDT; see **Table 2.2**). Reactions were diluted in one volume of clear formamide loading dye and heat denatured at 95°C before loading onto a 22 cm × 22 cm × 0.75 mm denaturing 7% 19:1 PAGE gel. Fluorescent cDNA products were separated for 80 min at 35 W and imaged on an Amersham Typhoon NIR laser scanner (Cytiva). U6 to U4 ratios of the IP based on primer extension analysis (Fig. 7A) and reported in Figure 7B were calculated using Equation 1. The IP subscript represents band intensities determined from the anti-FLAG samples.

$$\text{U6 to U4 ratio} = \frac{\text{experimental U6}_{IP}/\text{experimental U4}_{IP}}{\text{WT U6}_{IP}/\text{WT U4}_{IP}} \quad (1)$$

For RNAs isolated by IP and analyzed by solution hybridization, RNAs were isolated as described in the preceding paragraph, but with no NaOAc added. Efficiency of free U6 IP relative to wild type was calculated using Equation 2 after normalization of the samples using the U1 snRNA bands.

$$\text{Efficiency of free U6 IP} = \frac{\text{experimental U6}_{IP}/\text{experimental U6}_{Total}}{\text{WT U6}_{IP}/\text{WT U6}_{Total}} \quad (2)$$

Statistical analysis

All statistical tests were conducted using R (R Core Team 2021). Data collected from solution hybridization experiments was assumed to be normally distributed. One-way ANOVA was applied to test for differences in sample means between the WT, U6-II, and U6-II-Sm experimental strains in solution hybridization and co-IP experiments. Means between two experimental groups within the data set were compared with a post-hoc Tukey multiple pairwise-comparisons test to determine the statistical significance level (* P = 0.05, ** P = 0.01, *** P = 0.001, **** P = 0.0001).

Table 2.1 Yeast strains

Strain	Parent	Genotype	Reference
MWK027	PJ43-2b	MATα ura3-52 trp1-1 his3-11,15 leu2-3,112 ade2-1 lys2Δ2 met2Δ1 can1-101 snr6::LEU2 [YCp50-ΨWT]	Kaiser & Brow 1995
CJM000	ANK640	MATa, snr6::LEU2, snr14::trp1::ADE2, trp1, ura3, lys2, his3, ade2, [pRS316-U4wt-U6mini]	McManus et al. 2007
DAB101	CJM000	MATa, snr6::LEU2, snr14::trp1::ADE2, trp1, ura3, lys2, his3, ade2, [pRS313-U4wt + pRS314-U6wt]	This work
yAAH1361	CJM000	MATa, snr6::LEU2, snr14::trp1::ADE2, trp1, ura3, lys2, his3, ade2, [pRS313-U4wt + pRS314-U6-II-Sm]	This work
yAAH1442	CJM000	MATa, snr6::LEU2, snr14::trp1::ADE2, trp1, ura3, lys2, his3, ade2, [pRS313-U4wt + pRS314-U6-II]	This work
yAAH2687	BY4743	MATa/α his3Δ1/his3Δ1 leu2Δ0/leu2Δ0 LYS2/lys2Δ0 met15Δ0/MET15 ura3Δ0/ura3Δ0 LSM5::kanMX [pRS316-LSM5]	Roth et al. 2018
yAAH2690	CJM000	MATa, snr6::LEU2, snr14::trp1::ADE2, trp1, ura3, lys2, his3, ade2, SMD1::3xFLAG-HygR [pRS316-U4wt-U6mini]	This work
yAAH2719	CJM000	MATa, snr6::LEU2, snr14::trp1::ADE2, trp1, ura3, lys2, his3, ade2, LSM8::2xV5-HygR, [pRS316-U4wt-U6mini]	This work
yAAH2720	CJM000	MATa, snr6::LEU2, snr14::trp1::ADE2, trp1, ura3, lys2, his3, ade2, [pRS313-U4wt + pRS314-U6wt + pRS416-LSM5]	This work
yAAH2721	CJM000	MATa, snr6::LEU2, snr14::trp1::ADE2, trp1, ura3, lys2, his3, ade2, [pRS313-U4wt + pRS314-U6wt + pRS416-LSM8]	This work
yAAH2722	CJM000	MATa, snr6::LEU2, snr14::trp1::ADE2, trp1, ura3, lys2, his3, ade2, [pRS313-U4wt + pRS314-U6-II-Sm + pRS416-LSM5]	This work
yAAH2723	CJM000	MATa, snr6::LEU2, snr14::trp1::ADE2, trp1, ura3, lys2, his3, ade2, [pRS313-U4wt + pRS314-U6-II-Sm + pRS416-LSM8]	This work
yAAH2736	CJM000	MATa, snr6::LEU2, snr14::trp1::ADE2, trp1, ura3, lys2, his3, ade2, LSM5Δ::KanMX [pRS313-U4wt + pRS314-U6wt + pRS416-LSM5]	This work
yAAH2737	CJM000	MATa, snr6::LEU2, snr14::trp1::ADE2, trp1, ura3, lys2, his3, ade2, LSM8Δ:: HygR	This work

		[pRS313-U4wt + pRS314-U6wt + pRS416-LSM8]	
yAAH2750	CJM000	MATa, snr6::LEU2, snr14::trp1::ADE2, trp1, ura3, lys2, his3, ade2, LSM5Δ::KanMX [pRS313-U4wt + pRS314-U6-II-Sm + pRS416-LSM5]	This work
yAAH2751	CJM000	MATa, snr6::LEU2, snr14::trp1::ADE2, trp1, ura3, lys2, his3, ade2, LSM5Δ:: HygR [pRS313-U4wt + pRS314-U6-II-Sm + pRS416-LSM5]	This work
yAAH2773	CJM000	MATa, snr6::LEU2, snr14::trp1::ADE2, trp1, ura3, lys2, his3, ade2, LSM8::2xV5 HygR, [pRS313-U4wt + pRS314-U6WT]	This work
yAAH2774	CJM000	MATa, snr6::LEU2, snr14::trp1::ADE2, trp1, ura3, lys2, his3, ade2, LSM8::2xV5 HygR, [pRS313-U4wt + pRS314-U6-II-Sm]	This work
yAAH2775	CJM000	MATa, snr6::LEU2, snr14::trp1::ADE2, trp1, ura3, lys2, his3, ade2, LSM8::2xV5 HygR, [pRS313-U4wt + pRS314-U6-II]	This work
yAAH2853	CJM000	MATa, snr6::LEU2, snr14::trp1::ADE2, trp1, ura3, lys2, his3, ade2, SMD1::3xFLAG HygR, [pRS313-U4wt + pRS314-U6wt]	This work
yAAH2854	CJM000	MATa, snr6::LEU2, snr14::trp1::ADE2, trp1, ura3, lys2, his3, ade2, SMD1::3xFLAG HygR, [pRS313-U4wt + pRS314-U6-II-Sm]	This work
yAAH2855	CJM000	MATa, snr6::LEU2, snr14::trp1::ADE2, trp1, ura3, lys2, his3, ade2, SMD1::3xFLAG HygR, [pRS313-U4wt + pRS314-U6-II]	This work

Table 2.2 Oligonucleotides

Oligo	Sequence (5' to 3')
PsnR14(223)-Xho-F	CCGCTCGAGTAAGTAACCTCTGCATTGTC
U4-U6-R	CCCTCGCGAACGGAGTATTAAGGAAGGAAGTG
U4(12)-U6(24)-R	TCGGTAATGAAAAAAACGAAATAAATCTCTTTGTAA
U6(12)-U4(23)F	TTATTTCGTTTTTTTCATTACCGATATTCATTCTT
snR14-BamHI-701R	CGGGATCCTTCCTCTCTGCTGTTTTAG
Rnt1(12)-U6(24)R	CATTAGAGAATTAAAACGAAATAAATCTCTTTGTAA
U6(12)-Rnt(23)F	TTATTTCGTTTTAATTCTCTAATGTGAGTTCACGT
GAL1-XhoI-608F	CCGCTCGAGATCATATTACATGGCATTACCA
U4(14)-GAL1(24)R	CTGTTCCTTTTATATCTGTTAATAGATCAAAAATCATC
GAL1(11)-U4(25)F	CTATTAACAGATATAAAAGGAACAGAATATTAGTTA
snR14-Kpn1-701R	CGGGGTACCTTCCTCTCTGCTGTTTTAG
Sew-U6::U4-R	AAAAGGTATTCCAAAAATTCTTTGTAAAACGGTTCATCCTT
Sew-U6::U4-F	CAAAGAATTTTTGGAATACCTTTTAATTCTCTAATGTGAGTTCA
U1-SH	CCGTATGTGTGTGTGACC
U4-14B	AGGTATTCCAAAAATTCCC
U6-SH	ATTGTTTCAAATTGACCAAAT
U1RT136	GAC CAA GGA GTT TGC ATC AAT GAC
U2RTALL124	TTT GGG TGC CAA AAA ATG TGT ATT GTA
U4RTALL	GGT ATT CCA AAA ATT CCC TAC ATA GTC
U5	AAG TTC CAA AAA ATA TGG CAA GC
U6B	TCATCCTTATGCAGGG
U6D	AAA ACG AAA TAA ATC TCT TTG
U3 21-mer	CCAAGTTGGATTCAGTGGCTC

Table 2.3 Plasmids

Plasmid	Description	Reference
pRS314	*TRP1*-marked, low-copy yeast shuttle plasmid	Sikorski and Hieter 1989
pRS424	*TRP1*-marked, high-copy yeast shuttle plasmid	Christianson et al. 1992
pRS313-SNR14	*SNR14* -224 to +701 cloned into the BamHI site	Kuehner and Brow 2006
pRS314-14-6-14	See Materials and Methods	This work
pRS424-14-6-14	See Materials and Methods	This work
pRS314-SNR6-II	See Materials and Methods	This work
pRS424-SNR6-II	See Materials and Methods	This work
pRS314-GAL-SNR6-II	See Materials and Methods	This work
pRS314-GAL-SNR6-II-Sm	See Materials and Methods	This work
pAAH0229	pRS314-GAL-SNR6-SM (U6-II-Sm)/TRP/AMP	This work
pAAH0412	pRS314 WT U6/TRP/AMP	McManus et al. 2007
pAAH0413	pRS313 WT U4/HIS/AMP	McManus et al. 2007
pAAH0667	pRS314-GAL-SNR6-II (U6-II)/TRP/AMP	This work
pAAH1285	pRS416 LSM5/URA/CEN	Roth et al. 2018
pAAH1286	pRS416 LSM8/URA/CEN	Roth et al. 2018

2.5 References

Allmang C, Kufel J, Chanfreau G, Mitchell P, Petfalski E, Tollervey D. 1999. Functions of the exosome in rRNA, snoRNA and snRNA synthesis. *EMBO J* 18: 5399–5410. 10.1093/emboj/18.19.5399

Arimbasseri AG, Maraia RJ. 2015. Mechanism of transcription termination by RNA polymerase III utilizes a non-template strand sequence-specific signal element. *Mol Cell* 58: 1124–1132. 10.1016/j.molcel.2015.04.002

Barba-Aliaga M, Alepuz P, Pérez-Ortín JE. 2021. Eukaryotic RNA polymerases: the many ways to transcribe a gene. *Front Mol Biosci* 8: 663209. 10.3389/fmolb.2021.663209

Becker D, Hirsh AG, Bender L, Lingner T, Salinas G, Krebber H. 2019. Nuclear pre-snRNA export is an essential quality assurance mechanism for functional spliceosomes. *Cell Rep* 27: 3199–3214. 10.1016/j.celrep.2019.05.031

Bochnig P, Reuter R, Bringmann P, Lührmann R. 1987. A monoclonal antibody against 2,2,7-trimethylguanosine that reacts with intact, class U, small nuclear ribonucleoproteins as well as with 7-methylguanosine-capped RNAs. *Eur J Biochem* 168: 461–467. 10.1111/j.1432-1033.1987.tb13439.x

Bordonné R, Banroques J, Abelson J, Guthrie C. 1990. Domains of yeast U4 spliceosomal RNA required for PRP4 protein binding, snRNP-snRNP interactions, and pre-mRNA splicing in vivo. *Genes Dev* 4: 1185–1196. 10.1101/gad.4.7.1185

Brow DA, Guthrie C. 1990. Transcription of a yeast U6 snRNA gene requires a polymerase III promoter element in a novel position. *Genes Dev* 4: 1345–1356. 10.1101/gad.4.8.1345

Chan SP, Cheng SC. 2005. The Prp19-associated complex is required for specifying interactions of U5 and U6 with pre-mRNA during spliceosome activation. *J Biol Chem* 280: 31190–31199. 10.1074/jbc.M505060200

Christianson TW, Sikorski RS, Dante M, Shero JH, and Hieter P. 1992. Multifunctional yeast high-copy-number shuttle vectors. *Gene* 110: 119–122. doi: 10.1016/0378-1119(92)90454-w.

Dergai O, Hernandez N. 2019. How to recruit the correct RNA polymerase? Lessons from snRNA genes. *Trends Genet* 35: 457–469. 10.1016/j.tig.2019.04.001

Didychuk AL, Montemayor EJ, Brow DA, Butcher SE. 2015. Structural requirements for protein-catalyzed annealing of U4 and U6 RNAs during di-snRNP assembly. *Nucleic Acids Res* 44: 1398–1410. 10.1093/nar/gkv1374

Didychuk AL, Montemayor EJ, Carrocci TJ, DeLaitsch AT, Lucarelli SE, Westler WM, Brow DA, Hoskins AA, Butcher SE. 2017. Usb1 controls U6 snRNP assembly through evolutionarily divergent cyclic phosphodiesterase activities. *Nat Commun* 8: 497. 10.1038/s41467-017-00484-w

Didychuk AL, Butcher SE, Brow DA. 2018. The life of U6 small nuclear RNA, from cradle to grave. *RNA* 24: 437–460. 10.1261/rna.065136.117

Eschenlauer JB, Kaiser MW, Gerlach VL, Brow DA. 1993. Architecture of a yeast U6 RNA gene promoter. *Mol Cell Biol* 13: 3105–3026. 10.1128/mcb.13.5.3015

Fernandez CF, Pannone BK, Chen X, Fuchs G, Wolin SL. 2004. An Lsm2-Lsm7 complex in Saccharomyces cerevisiae associates with the small nucleolar RNA snR5. *Mol Biol Cell* 15: 2842–2852. 10.1091/mbc.e04-02-0116

Gerlach VL, Whitehall SK, Geiduschek EP, Brow DA. 1995. TFIIIB placement on a yeast U6 RNA gene in vivo is directed primarily by TFIIIC rather than by sequence-specific DNA contacts. *Mol Cell Biol* 15: 1455–1466. 10.1128/mcb.15.3.1455

Goldstein AL, McCusker JH. 1999. Three new dominant drug resistance cassettes for gene disruption in Saccharomyces cerevisiae. *Yeast* 15: 1541–1553. 10.1002/(SICI)1097-0061(199910)15:14<1541::AID-YEA476>3.0.CO;2-K

Gunnery S, Mathews MB. 1995. Functional mRNA can be generated by RNA polymerase III. *Mol Cell Biol* 15: 3597–3607. 10.1128/MCB.15.3587

Guthrie C, Patterson B. 1988. Spliceosomal snRNAs. *Annu Rev Genet* 22: 387–419. 10.1146/annurev.ge.22.120188.002131

Harismendy O, Gendrel CG, Soularue P, Gidrol X, Sentenac A, Werner M, Lefebvre O. 2003. Genome-wide location of yeast RNA polymerase III transcription machinery. *EMBO J* 22: 4738–4747. 10.1093/emboj/cdg466

Horton RM, Ho SN, Pullen JK, Hunt HD, Cai Z, Pease LR. 1993. Gene splicing by overlap extension. *Methods Enzymol* 217: 270–279. 10.1016/0076-6879(93)17067-f

Janke C, Magiera MM, Rathfelder N, Taxis C, Reber S, Maekawa H, Moreno-Borchart A, Doenges G, Schwob E, Schiebel E, et al. 2004. A versatile toolbox for PCR-based tagging of yeast genes: new fluorescent proteins, more markers and promoter substitution cassettes. *Yeast* 21: 947–962. 10.1002/yea.1142

Johnston M, Davis RW. 1984. Sequences that regulate the divergent GAL1-GAL10 promoter in Saccharomyces cerevisiae. *Mol Cell Biol* 4: 1440–1448. 10.1128/mcb.4.8.1440-1448.1984

Jones MH, Guthrie C. 1990. Unexpected flexibility in an evolutionarily conserved protein-RNA interaction: genetic analysis of the Sm binding site. *EMBO J* 9: 2555–2561. 10.1002/j.1460-2075.1990.tb07436.x

Kaiser MW, Brow DA. 1995. Lethal mutations in a yeast U6 RNA gene B block promoter element identify essential contacts with transcription factor TFIIIC. *J Biol Chem* 270: 11398–11405. 10.1074/jbc.270.19.11398

Karaduman R, Fabrizio P, Hartmuth K, Urlaub H, Lührmann R. 2006. RNA structure and RNA-protein interactions in purified yeast U6 snRNPs. *J Mol Biol* 365: 1248–1262. 10.1016/j.jmb.2005.12.013

Kuehner JN, Brow DA. 2006. Quantitative analysis of in vivo initiator selection by yeast RNA polymerase II supports a scanning model. *J Biol Chem* 281: 14119–14128. 10.1074/jbc.M601937200

Kwan S, Gerlach VL, Brow DA. 2000. Disruption of the 5′ stem-loop of yeast U6 RNA induces trimethylguanosine capping of this RNA polymerase III transcript in vivo. *RNA* 6: 1859–1869. 10.1017/s1355838200991325

Li Z, Brow DA. 1993. A rapid assay for quantitative detection of specific RNAs. *Nucleic Acids Res* 21: 4645–4646. 10.1093/nar/21.19.4645

Listerman I, Bledau AS, Grishina I, Neugebauer KM. 2007. Extragenic accumulation of RNA polymerase II enhances transcription by RNA polymerase III. *PLoS Genet* 3: e212. 10.1371/journal.pgen.0030212

Luukkonen BGM, Séraphin B. 1998. Construction of an in vivo-regulated U6 snRNA transcription unit as a tool to study U6 function. *RNA* 4: 231–238.

Mabin JW, Lewis PW, Brow DA, Dvinge H. 2021. Human spliceosomal snRNA sequence variants generate variant spliceosomes. *RNA* 27: 1186–1203. 10.1261/rna.078768.121

Madhani HD, Bordonné R, Guthrie C. 1990. Multiple roles for U6 snRNA in the splicing pathway. *Genes & Dev* 4: 2264–2277. 10.1101/gad.4.12b.2264

Martin MP, Gerlach VL, Brow DA. 2001. A novel upstream RNA polymerase III promoter element becomes essential when the chromatin structure of the yeast U6 RNA gene is altered. *Mol Cell Biol* 21: 6429–6439. 10.1128/MCB.21.19.6429-6439.2001

Mayas RM, Maita H, Staley JP. 2006. Exon ligation is proofread by the DExD/H-box ATPase Prp22p. *Nat Struct Mol Biol* 13: 482–490. 10.1038/nsmb1093

Mayes AE, Verdone L, Legrain P, Beggs JD. 1999. Characterization of Sm-like proteins in yeast and their association with U6 RNA. *EMBO J* 18: 4321–4331. 10.1093/emboj/18.15.4321

McManus CJ, Schwartz ML, Butcher SE, Brow DA. 2007. A dynamic bulge in the U6 RNA internal stem-loop functions in spliceosome assembly and activation. *RNA* 13: 2252-2265. doi:10.1261/rna.699907.

Montemayor EJ, Didychuk AL, Yake AD, Sidhu GK, Brow DA, Butcher SE. 2018. Architecture of the U6 snRNP reveals specific recognition of 3′-end processed U6 snRNA. *Nat Commun* 9: 1749. 10.1038/s41467-018-04145-4

Moqtaderi Z, Struhl K. 2004. Genome-wide occupancy profile of the RNA polymerase III machinery in Saccharomyces cerevisiae reveals loci with incomplete transcription complexes. *Mol Cell Biol* 24: 4118–4127. 10.1128/MCB.24.10.4118-4127.2004

Mroczek S, Krwawicz J, Kutner J, Lazniewski M, Kucnski I, Ginalski K, Dziembowski A. 2012. C16orf57, a gene mutated in poikiloderma with neutropenia, encodes a putative phosphodiesterase responsible for the U6 snRNA 3′ end modification. *Genes Dev 26*: 1911–1925. 10.1101/gad.193169.112

Novikova O, Belfort M. 2017. Mobile group II introns as ancestral eukaryotic elements. *Trends Genet 33*: 773–783. 10.1016/j.tig.2017.07.009

Olson BL, Siliciano PG. 2003. A diverse set of nuclear RNAs transfer between nuclei of yeast heterokaryons. *Yeast* 20: 893–903. 10.1002/yea.1015

Pannone BK, Xue D, Wolin SL. 1998. A role for the yeast La protein in U6 snRNP assembly: evidence that the La protein is a molecular chaperone for RNA polymerase III transcripts. *EMBO J 17*: 7442–7453. 10.1093/emboj/17.24.7442

Pannone BK, Kim SD, Noe DA, Wolin SL. 2001. Multiple functional interactions between components of the Lsm2-Lsm8 complex, U6 RNA, and the yeast La protein. *Genetics 158*: 187–196. 10.1093/genetics/158.1.187

Patterson B, Guthrie C. 1987. An essential yeast snRNA with a U5-like domain is required for splicing in vivo. *Cell* 49: 613–624. 10.1016/0092-8674(87)90537-x

Pessa HKJ, Will CL, Meng X, Schneider C, Watkins NJ, Perälä N, Nymark M, Turunen JJ, Lührmann R, Frilander MJ. 2008. Minor spliceosome components are predominantly localized in the nucleus. *Proc Natl Acad Sci* 105: 8655–8660. 10.1073/pnas.0803646105

R Core Team. 2021. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Roth AJ, Shuman S, Schwer B. 2018. Defining essential elements and genetic interactions of the yeast Lsm2–8 ring and demonstration that essentiality of Lsm2–8 is bypassed via overexpression of U6 snRNA or the U6 snRNP subunit Prp24. *RNA* 24: 853–864. 10.1261/rna.066175.118

Salgado-Garrido J, Bragado-Nilsson E, Kandels-Lewis S, Séraphin B. 1999. Sm and Sm-like proteins assemble in two related complexes of deep evolutionary origin. *EMBO J* 18: 3451–3462. 10.1093/emboj/18.12.3451

Seraphin B, Rosbash M. 1989. Identification of functional U1 snRNA-pre-mRNA complexes committed to spliceosome assembly and splicing. *Cell* 59: 349–358. 10.1016/0092-8674(89)90296-1

Shchepachev V, Wischnewski H, Missiaglia E, Soneson C, Azzalin C. 2012. Mpn1, mutated in poikiloderma with neutropenia protein 1, is a conserved 3′-to-5′ RNA exonuclease processing U6 small nuclear RNA. *Cell Rep* 2: 855–865. 10.1016/j.celrep.2012.08.031

Sikorski RS, Hieter P. 1989. A system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in Saccharomyces cerevisiae. *Genetics* 122: 19-27. doi: 10.1093/genetics/122.1.19

Sikorski RS, Boeke JD. 1991. In vitro mutagenesis and plasmid shuffling: from cloned gene to mutant yeast. *Methods Enzymol* 194: 302–318. 10.1016/0076-6879(91)94023-6

Singh R, Reddy R. 1989. γ-Monomethyl phosphate: a cap structure in spliceosomal U6 small nuclear RNA. *Proc Natl Acad Sci* 86: 8280–8283. 10.1073/pnas.86.21.8280

Spiller MP, Boon KL, Reijns MAM, Beggs JD. 2007a. The Lsm2-8 complex determines nuclear localization of the spliceosomal U6 snRNA. *Nucleic Acids Res* 35: 923–929. 10.1093/nar/gkl1130

Spiller MP, Reijns MAM, Beggs JD. 2007b. Requirements for nuclear localization of the Lsm2-8p complex and competition between nuclear and cytoplasmic Lsm complexes. *J Cell Sci* 120: 4310–4320. 10.1242/jcs.019943

Steinmetz EJ, Conrad NK, Brow DA, Corden JL. 2001. RNA-binding protein Nrd1 directs poly(A)-independent 3′-end formation of RNA polymerase II transcripts. *Nature* 413: 327–331. 10.1038/35095090

Steinmetz EJ, Warren CL, Kuehner JN, Panbehi B, Ansari AZ, Brow DA. 2006. Genome-wide distribution of yeast RNA polymerase II and its control by Sen1 helicase. *Mol Cell* 24: 735–746. 10.1016/j.molcel.2006.10.023

Tang W, Kannan R, Blanchette M, Baumann P. 2012. Telomerase RNA biogenesis involves sequential binding by Sm and Lsm complexes. *Nature* 484: 260–264. 10.1038/nature10924

Townsend C, Leelaram MN, Agafonov DE, Dybkov O, Will CL, Bertram K, Urlab H, Kastner B, Stark H, Lührmann R. 2020. Mechanism of protein-guided folding of the active site U2/U6 RNA during spliceosome activation. *Science* 370: eabc3753. 10.1126/science.abc3753

Treco DA, Lundblad V. 1993. Preparation of yeast media. *Curr Protoc Mol Biol* 23: 13.11.11–13.11.17. 10.1002/0471142727.mb1301s23

Wise JA, Tollervey D, Maloney D, Swerdlow H, Dunn EJ, Guthrie C. 1983. Yeast contains small nuclear RNAs encoded by single copy genes. *Cell* 35: 743–751. 10.1016/0092-8674(83)90107-1

Younis I, Dittmar K, Wang W, Foley SW, Berg MG, Hu KY, Wei Z, Wan L, Dreyfuss G. 2013. Minor introns are embedded molecular switches regulated by highly unstable U6atac snRNA. *Elife* 2: 00780. 10.7554/eLife.00780

CHAPTER 3

In vivo fluorescent tagging of U4 snRNA for single molecule observation of spliceosome activation kinetics

Karli A. Lipinski, Xin Chen, Peter J. Unrau, Aaron A. Hoskins

AH and KAL conceptualized the project. KAL and XC designed, cloned, and transformed U4 Mango variants. KAL performed experiments and analyzed data. KAL wrote the chapter. See contributions (pg. xii).

CHAPTER 3: In vivo fluorescent tagging of U4 snRNA for single molecule observation of spliceosome activation kinetics

3.1 Abstract

Splicing involves the removal of introns and ligation of exons to form mature mRNA transcripts, an essential process for gene expression in eukaryotes. Spliceosomes are complex macromolecules that undergo large conformational and compositional changes as they assemble and catalyze splicing. While cryo-EM structures have provided insights into spliceosome composition and arrangement, the transition between intermediate states remains poorly understood. During the transition from B to B^{ACT} spliceosomes, ~50 factors are exchanged. The Brr2 helicase mediates release of the U4 snRNA and associated factors by disrupting U4/U6 base pairing, yet the exact mechanism remains unclear. The U4/U6 di-snRNA Stem II duplex is stabilized by bound proteins (Prp3/4) and supporting evidence for Brr2 translocation through Stem II is lacking. To examine the mechanism of U4 release during spliceosome activation, we employ endogenous fluorescent labeling strategies, such as incorporating MS2 or Mango sequences, for observation of the U4 snRNA in colocalization single molecule spectroscopy (CoSMoS). We find that incorporation of these tags is compatible with spliceosome function and has minimal effects on yeast cell growth and splicing activity. Further testing and optimization of endogenously labeled U4 fluorescent systems is warranted for single molecule studies of spliceosome activation.

3.2 Introduction

Splicing is an essential process in gene regulation where introns are removed, and exons are joined together to form mature mRNA transcripts. The spliceosome, a large macromolecular complex made of dozens of proteins and five small nuclear RNAs (snRNA; U1, U2, U4, U5, U6), catalyzes splicing. Spliceosome subunits (small nuclear ribonucleoproteins; snRNPs) composed of one snRNA and associated proteins assemble de novo on each intron through a series tightly regulated steps. Splicing occurs in two sequential transesterification reactions that result in intron excision and exon ligation to form an mRNA that can be translated. Prior to either reaction, spliceosomes undergo a large composition and conformation change where the active site is formed. Cryo-EM structures of spliceosome intermediates before activation (B complex) and after activation (B^{ACT} complex) have provided insights into composition and arrangements of protein and snRNA components (Shi, 2017; Townsend et al. 2020). However, the transition between B to B^{ACT} intermediates during activation remains poorly characterized in part due to the number of factors (~50) exchanged. Predominantly, the U4 snRNP components are lost and a protein only complex, the NTC (NineTeen Complex) with NTC-related factors

(NTR), arrives (Tarn et al. 1993 a and b). The resulting composition changes and structural rearrangements lead to large changes within the spliceosome active site.

U6 forms the spliceosome active site where the internal-stem loop (ISL) coordinates essential magnesium ions for catalysis (Karadumann et al. 2006; Montemayor et al. 2014; Yan et al. 2015). The ISL is present in the free U6 snRNP and in catalytic spliceosomes but is unwound and base paired to the U4 snRNA in di-snRNP, tri-snRNP, and spliceosome complexes prior to catalysis (Wahl et al. 2009; Hoskins et al. 2011; Staley and Guthrie, 1998). U4 maintains U6 in a catalytically inactive conformation and release of the U4 snRNP is therefore required for structural rearrangement of the catalytic core (Liu and Cheng, 2015). Helicase activity of Brr2 mediates release of the U4 snRNA, but it is unknown if Brr2 activity alone is enough to facilitate U4 release and whether conformational changes within U6 also contribute to release.

The U4/U6 di-snRNA is composed of two stem regions (I and II) separated by a three-way junction formed by U6 and the U4 5′ stem loop (SL) (Hardin et a. 2015; Cornilescu et al. 2016). Crosslinking localizes Brr2 within Stem I, but not Stem II of the U4/U6 duplex (Hahn et al. 2012). Stem II is the more thermodynamically stable stem and is additionally separated from Stem I by proteins that bind the U4 5' SL. Stem II is further stabilized by binding of Prp3/4 di-snRNP specific proteins along the Stem II duplex (Liu et al. 2015). Therefore, disruption of the Stem II U4/U6 duplex and release of U4 may occur through a second ATP-independent step, which requires release of the bound Prp3/4 proteins prior to activation.

Indeed, disruption of the Stem II duplex can be accomplished without the addition of Brr2 or ATP in vitro by the addition of complementary oligo nucleotides (Rodgers et al. 2016). Competing snRNA structures, such as the U6 telestem or the U2/U6 duplex, could facilitate release of U4 in combination with Brr2 activity (Rodgers et al. 2016; Theuser et al. 2016). Release of Prp3, which binds a portion of Stem II, may first be required before any competing structures can be formed (Liu et al. 2015; Wan et al. 2016; Nguyen et al, 2016). Addition of ATP to in vitro reconstituted di-snRNPs results in free U6 snRNA, free Brr2, U4/Prp31/Snu13, and Prp3/4 particles (Theuser et al. 2016. Therefore, it is known that Prp3/4 are released as a separate particle, however elucidating whether Prp3/4 release is required before release of the U4 snRNA will yield insight into the activation mechanism.

Our lab has previously used colocalization single molecule spectroscopy (CoSMoS) to study protein arrival and release during spliceosome activation (Hoskins et al, 2016; Fu et al. 2022). Characterization of spliceosome activation by CoSMoS revealed that association of the tri-snRNP is reversible while release of the U4 snRNP protein Prp3 is irreversible (Hoskins et al. 2016). Studies of spliceosome assembly and activation using CoSMoS have employed endogenous strategies for protein labeling in WCEs (whole cell extracts), such as fusion of a SNAP tag that reacts with benzyl guanine derivatives containing fluorophores (Hoskins et al. 2016). However, endogenous fluorescent labeling

of an snRNA component of the spliceosome for single molecule CoSMoS studies has not been attempted.

Biochemical studies of the U1 and U2 snRNAs in yWCE have been accomplished via depletion from yWCE and addition of in vitro prepared snRNA for U2, U5, and U6 in yeast (Fabrizio et al. 1989; McPheeters et al. 1989; O'Keefe et al.1996). In contrast, depletion of the U4 snRNA and addition of an in vitro prepared U4 has had limited success, requiring 300-fold excess of endogenous U4 levels to reconstitute functional di-snRNPs (Hayduk et al. 2012). Addition of high quantities of an in vitro fluorescently labeled U4 would produce high fluorescent background and lower signal to noise ratios in single molecule experiments. To overcome the challenges of labeling U4 snRNA in yeast whole cell extracts, endogenous approaches may be explored.

Addition of an RNA sequence within the U4 snRNA gene, such as the MS2 tag or RNA aptamers Mango, may yield successful endogenous fluorescent labeling. However, it is unknown if insertion of additional nucleotides within U4 will be compatible with the function and stability of in the spliceosome. We have designed several variant U4 snRNAs by incorporating either Mango or MS2 sequences and assessed the growth and splicing activity of yeast cells. Insertions within the 3' SL of the U4 snRNA have minimal effects on yeast cell growth and splicing activity. Preliminary single molecule experiments require further optimization to assess the suitability of single copies of tags inserted in the U4 snRNA for single molecule studies of spliceosome activation.

3.3 Results

Design of endogenously tagged U4 snRNA

Introduction of sequence motifs for endogenous fluorescent labeling of the U4 snRNA will ideally not disrupt base pairing or RNA-protein interactions present in the di-snRNP (**Figure 3.1**). Although presence of the U4 3′ SL is essential for viability (Bordonne et al. 1990), the primary sequence is mutation tolerant with only the deletion of nt 131-133 producing a phenotype in yeast (Hu et al. 1995). These results suggest that the overall structure of the stem loop is important for function. Using the U4 3′ SL as a scaffold, we designed three U4 constructs incorporating the MS2 SL for binding the MCP (MS2 coat protein) and three for incorporating Mango-II, -III, and iMango variants.

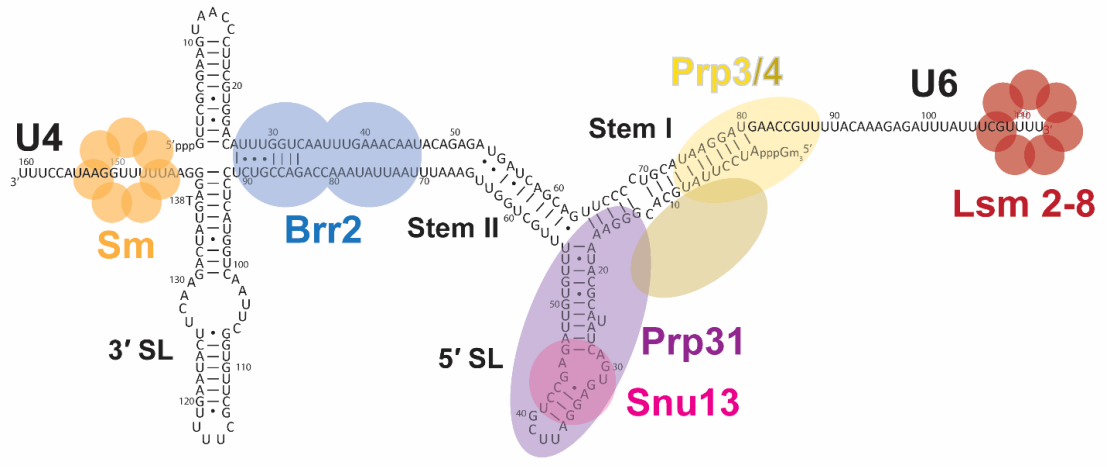

Figure 3.1 U4/U6 di-snRNA base pairing interactions in *S. cerevisiae*. Approximate locations of bound proteins are shown. Sm and Lsm 2-8 rings are bound to the 3′ end of the U4 and U6 snRNAs, respectively. Prp31 and Snu13 are bound to the U4 5 SL, which forms a three-way junction between Stem I and II of the di-snRNA. Brr2 is loaded onto the singly stranded U4 region adjacent to the U4 3′ SL.

The MS2-MCP system has been used previously for single molecule localization of nascent mRNAs within cells, albeit with multiple copies of the MS2 SL and endogenous expression of the MCP-GFP (Bertrand et al. 1998). MCP is not expressed endogenously in our system; instead, purified MCP-SNAP protein will be added to yWCE to reduce fluorescence background in single molecule experiments. We used the C-variant of the MS2 SL that has increased affinity for the MCP (K_d reduced from 10 nM to 1 nM) and decreased disassociation kinetics by ~90-fold (Lowary et al. 1987; Valegard et al. 1997). The MS2 SL was added to the U4 3′ SL (boxed, **Fig. 3A**) either replacing the lower stem resulting in U4-MS2 Short (**Fig 3.2B**) or stacked on the lower stem resulting in U4-MS2 Long (**Fig 3.2B**). U4-MS2 Short maintains the overall length of the U4 snRNA, but replacement of the lower stem with the MS2 SL results in replacement of several wobble pairs with more stable Watson-Crick pairs, hyper-stabilizing the lower half of the U4 3′ SL. Most wobble pairs in the 3′ SL are maintained in U4-MS2 Long; however, the insertion increases the overall length of U4 by 12 nt. Finally, as an alternative to incorporation of the MS2 SL in the 3′ SL of U4, we added a scaffold stem analogous to a human sequence absent in yeast after the Sm ring binding site. The MS2 SL sequence was added to the scaffold stem to create U4-MS2 3′ (**Figure 3.2D**). It is not known if incorporation of the human stem and the MS2 SL will interfere with 3′ end processing or if the tag will remain intact.

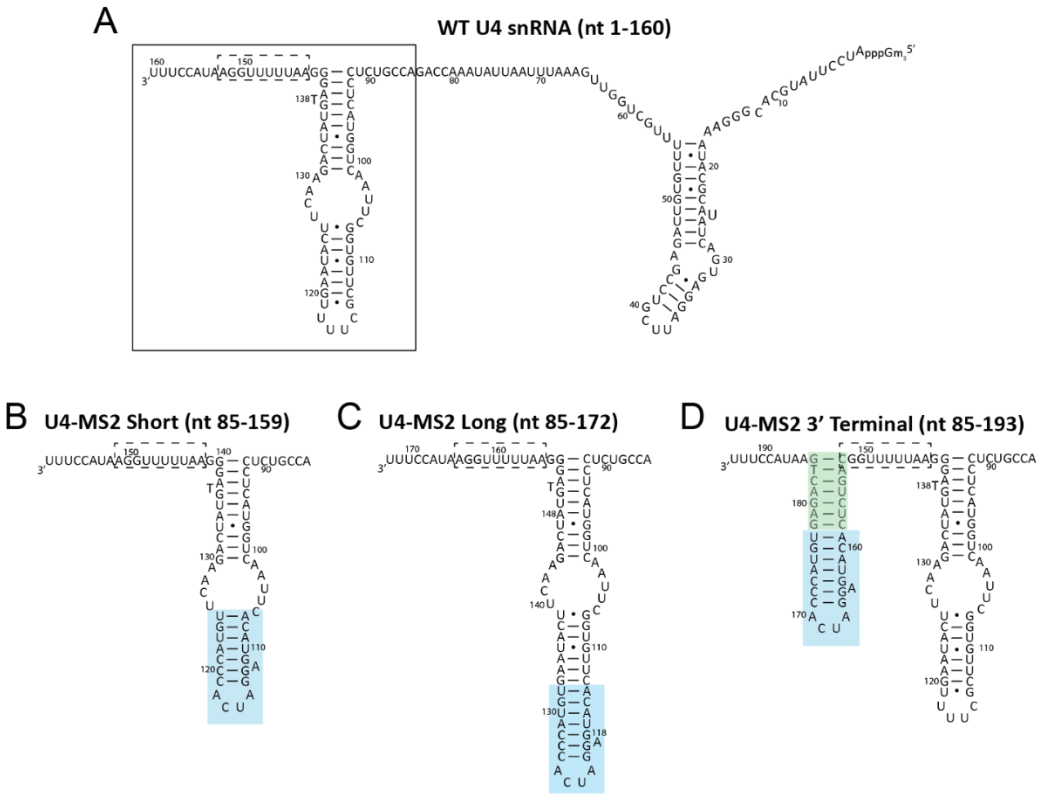

Figure 3.2 U4-MS2 constructs. A) U4 snRNA sequence and secondary structure in U4/U6 di-snRNP shown with a dotted box around the Sm ring binding site and a box around the sequence shown below containing the MS2 SL C-variant (blue) inserted either **B)** onto the lower half of the U4 3′ SL replacing endogenous sequence, **C)** stacked onto the existing SL after the final Watson-crick base pair, or **D)** after the Sm binding sequence stacked on an added stem (green) present in the human U4 snRNA secondary structure.

We also designed three U4 constructs incorporating the fluorogenic RNA aptamer Mango (Mango-II, Mango-III, and iMango). Fluorogenic RNA aptamers fold into a tertiary structure that binds a non-fluorescent small molecule in solution and induces fluorescence upon binding due to conformational constraint of the small molecule. One such RNA aptamer is Mango, which is advantageous for its small size (~30 nt) and bright fluorescence turn-on (Dolgosheina et al. 2014). Crystal structures of Mango bound to thiazole orange biotin (TO1-Biotin) show a G-quadruplex aptamer core with the entire TO1-biotin molecule bound on the face of the upper G-quartet (Trachman III et al. 2017). Single molecule studies of the spliceosome using CoSMoS require tethering of a pre-mRNA to the slide surface, most readily accomplished using binding of streptavidin to biotin (Hoskins et al. 2016). Therefore, presence of a biotin within Mango ligand binding pocket is less ideal for our single molecule set-up. We therefore opted to use reselected variants of Mango (II, III, and iMango) which do not show density in their crystal structures for the biotin molecule (Trachman III et al. 2018, 2019). For microscopy, we used a TO1

(thiazole orange) derivative, TO1-Cy5, using the TO1 moiety as a handle for Mango binding and a Cy5 fluorophore suitable for SM microscopy (**Figure 3.3A**). Either Mango-II, III, or iMango were inserted into the 3′ SL leaving most of the stem intact but replacing the lower half of the stem (**Figure 3.3B**). The wobble G-U base pair at position 110 was changed to a Watson-Crick base pair to stabilize the stem and promote aptamer folding and stability.

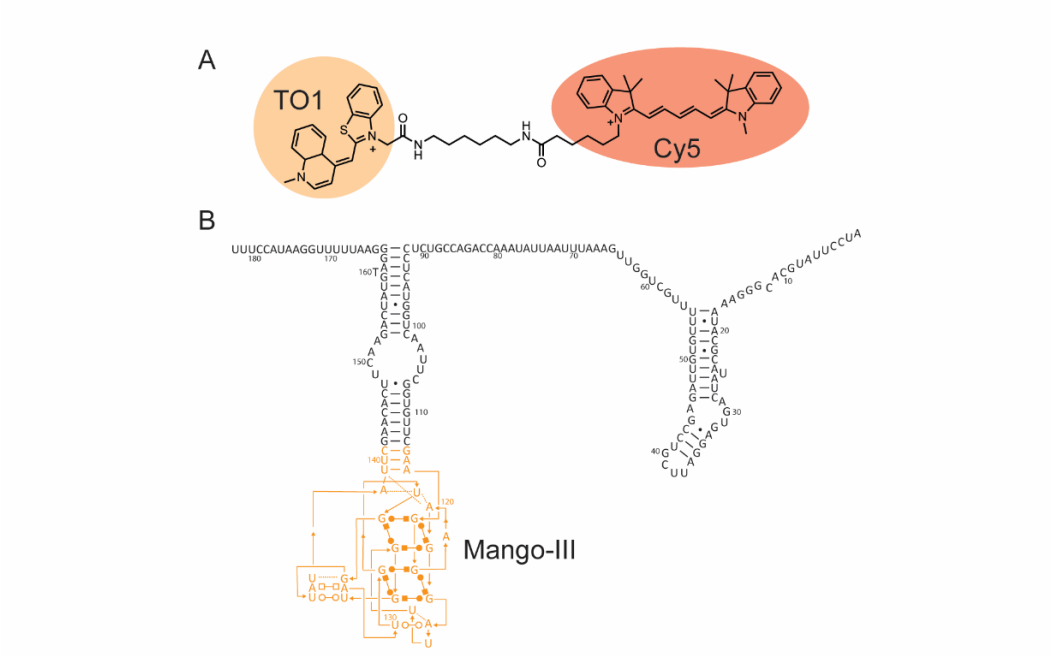

Figure 3.3 Integration of Mango into the U4 snRNA. A) Structure of the TO1-Cy5 ligand for single molecule microscopy. B) Secondary structure of the U4 with Mango-III inserted into the 3′ SL (orange).

Expression of U4 variants in yeast affect growth minimally

To examine whether our U4 constructs deleteriously affect yeast growth, we expressed each U4 construct as the sole source of U4 snRNA in yeast in strains deleted for both *SNR14* and *SNR6*. *SNR14* (U4) variants and WT *SNR6* (U6) were then supplied on a single plasmid bearing the *URA3* gene. *SNR14* is an essential gene and variants that are unable to be utilized in spliceosomes will result in lethality. We therefore employed a genetic screening method leveraging the indispensability of *SNR14* and the sensitivity of the *URA3* gene to 5-fluoroorotic acid (5FOA). Initially, yeast cells are transformed with plasmids bearing variant *SNR14* genes on a plasmid bearing a *TRP* marker. Transformed yeast cells are selected on tryptophan-deficient plates to isolate successful transformants. Supplementation of 5FOA in media serves to discriminate between cells harboring functional U4 snRNA variants capable of mediating yeast survival. Yeast cells expressing functional *SNR14* variants will lose the WT *SNR14* and *URA3* genes and be unable to metabolize 5FOA via the Ura3 activity. Conversely, yeast cells expressing nonfunctional

variants will retain the WT SNR14 and URA3 genes, leading to conversion of 5FOA to the cytotoxic compound 5-fluorouracil and lethality. All of our U4 constructs for either U4-MS2 or U4-Mango resulted in viable yeast colonies when selected on 5FOA. We also tested our U4 constructs in a strain background with genomically incorporated C-terminal SNAP and DHFR tags for fluorescent labeling of the U4-associated protein Prp3 and the tri-snRNP proteins Brr2 and Snu114, respectively. All U4 strains were viable in the triply tagged strain background except for U4-MS2 3′.

For surviving strains expressing our U4 constructs, we assessed for defects in yeast growth by plating the yeast at nonoptimal temperatures for growth. Compared to yeast expressing WT U4, yeast expressing all variants U4-MS2 grow less well at 30°C and display cold-sensitive growth defects when grown at 23 and 16°C (**Fig. 3.4A**). However, of the three U4-MS2 variants, only U4-MS2 Short displays growth defects at 37°C. In the triply tagged strain background, viable U4-MS2 variants become more cold sensitive compared to the untagged strain (**Fig. 3.4A**). Additionally, the triply tagged strain expressing WT U4 is heat sensitive, and the growth defect is still present when expressing U4-MS2 Long.

To determine which of the three tags produces the temperature-sensitive growth defect, we tested the growth of strains at different temperatures expressing one, two, or three tagged proteins. Fusion of either Prp3, Prp4, or Snu114 with DHFR or SNAP did not result in heat sensitive growth defects (**Fig 3.4B,C**). However, C-terminal tagging of Brr2 with a SNAP tag reproduces the heat sensitive phenotype.

We also measured growth of yeast expressing U4-MS2 variants in liquid culture to determine if there if growth at optimal temperatures is affected by insertion of the MS2 SL. Yeast expressing either U4-MS2 Short, Long, or 3′ have an increased doubling time from yeast expressing WT U4 by ~1 hr (**Fig. 3.4D,E**). In the triply tagged yeast background, the doubling time of yeast expressing WT U4 increases by ~30 min while the doubling time for yeast expressing U4-MS2 Short is similar to the untagged strain background. However, yeast expressing U4-MS2 Long have a greatly increased doubling time of about 1.5 h from the untagged yeast background. Differences in doubling time are consistent with the growth at 30°C of plated yeast in the temperature growth assays (**Fig. 3.4A**). Overall, expression of U4-MS2 in yeast results in viable colonies with minimal growth defects.

Purified tdMCP-SNAP binds MS2 with high affinity

In yeast expressing U4-MS2 variants, a recombinant fluorescently labeled MCP must be added to WCE for single molecule experiments. We expressed a tandem dimer of two fused copies of the MCP as it eliminates the need for dimerization before binding to a MS2 SL (Wu et al. 2012). Each MCP also contained a single point mutation (V29I) to decrease coat protein oligomerization (Peabody et al. 1992). To fluorescently label the tdMCP, we fused the protein to a SNAP tag which selectively reacts with benzyl-guanine

derivatives for covalent fluorescent labeling. We expressed tdMCP-SNAP in *E. coli* (**Fig. 3.5A, B**), performed fluorescent labeling with the fluorophore Dy649 for observation in the 633 nm laser channel, and removed excess fluorescent SNAP reagent (see Materials and Methods).

We then assessed the binding affinity of tdMCP-SNAP for a fluorescently labeled MS2 SL oligo (Cy3) using an electrophoretic mobility shift assay (EMSA). tdMCP-SNAP bound tightly to the MS2 oligo, as seen by the strong shift of the MS2 oligo at very low concentrations of the tdMCP-SNAP protein (**Figure 3.5C**). Due to the high affinity of tdMCP-SNAP for the MS2 SL, low enough concentrations of protein were not tested to accurately determine a K_d value. However, the EMSA showed that the purified tdMCP-SNAP protein is functional for binding the MS2 sequence in vitro.

U4-Mango constructs bind TO1-Biotin in vitro

The most important aspect for functionality of RNA aptamers is correct folding without which a ligand cannot bind. Arrays of Mango aptamers in vivo have been shown to have less actively fluorescent aptamers than the total amount of aptamers within the array (Cawte et al. 2022). We assessed if the U4-Mango construct can fold properly in vitro to bind one of its fluorescent ligands, TO1-Biotin, by running total RNA extracted under denaturing conditions from yeast expressing U4-Mango constructs on a denaturing gel and staining with TO1-Biotin according to established protocols (**Figure 3.6**; Yaseen et al. 2019). There is a faint band present in samples containing U4-Mango that is absent in samples containing WT U4. We therefore conclude that this band is U4-Mango and the U4-Mango snRNA is able to fold into a tertiary structure capable of binding its ligand in vitro.

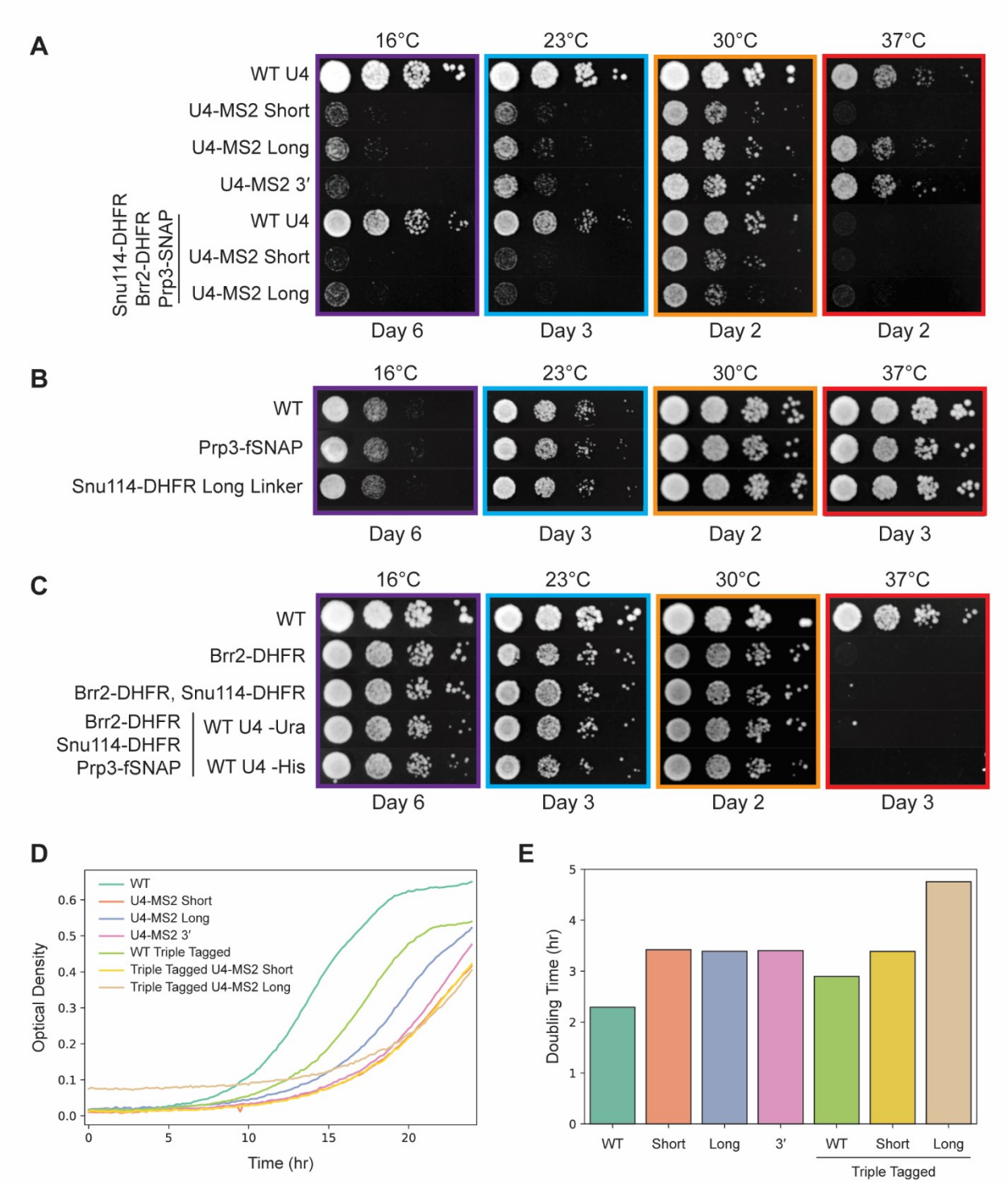

Figure 3.4 Growth of U4-MS2 variants expressed in yeast. Temperature growth assay with yeast spotted on YPD plates expressing either A) U4-MS2 variants, B) singly tagged Prp3 or Snu114, or C) successively tagged strains expressing tags present in triply tagged (Brr2, Snu114, Prp3) strain. placed at either 16, 23, 30, or 37°C. Plate images were taken on the indicated days. D) Growth curve of yeast cultured over 24 hours and E) calculated doubling times from yeast growth curves.

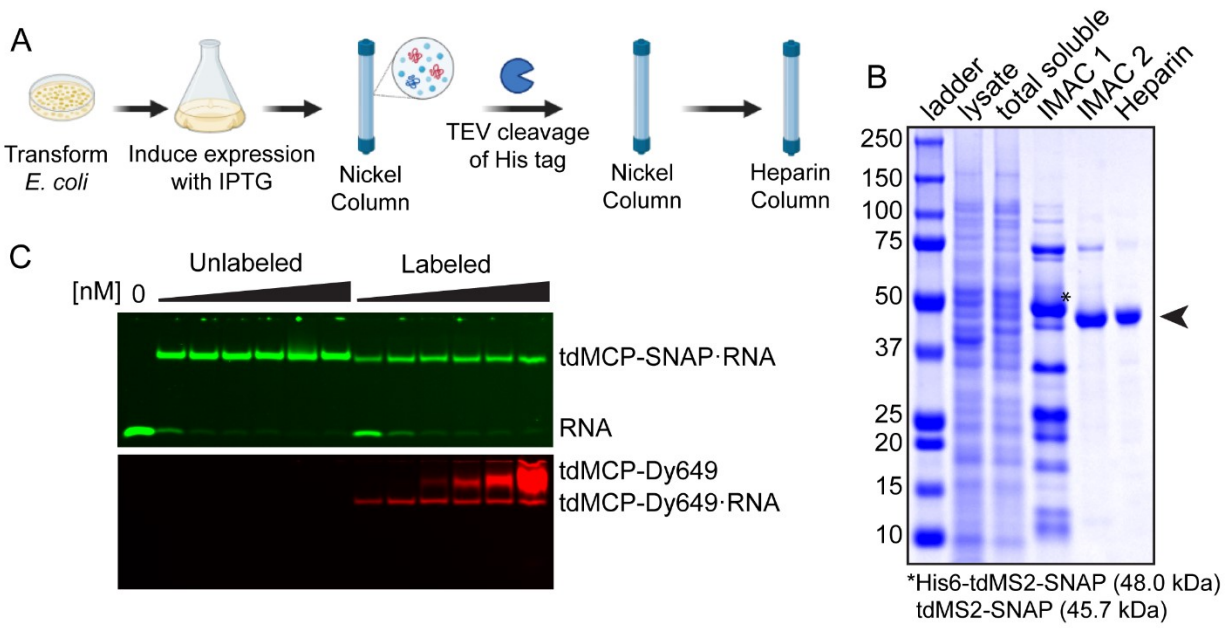

Figure 3.5 Purification of tdMCP-SNAP. A) Purification scheme made with BioRender (see Methods). **B)** Coomassie stained SDS-Page gel showing pure tdMCP-SNAP protein after heparin column step (arrow). **C)** Gel shift of purified tdMCP-SNAP with and without fluorescent labels (Dy649) and fluorescently labeled MS2-Cy3 oligo. Unlabeled and labeled tdMCP-SNAP was present at 0, 0.1, 1, 10, 100, 1000, 2500 nM concentrations.

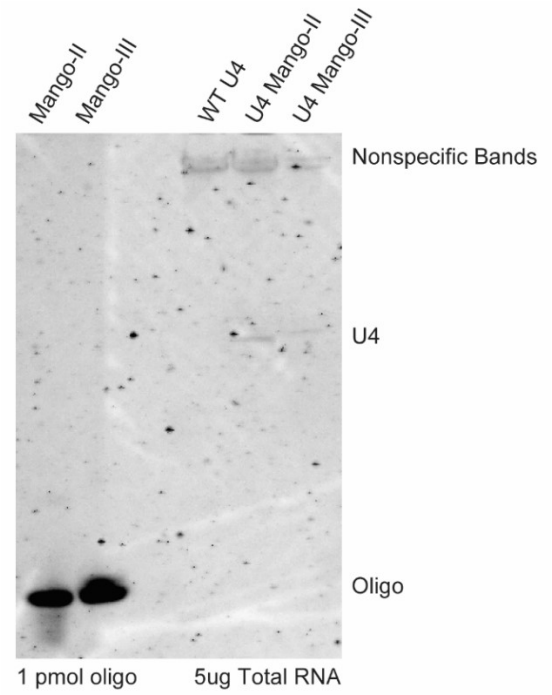

Figure 3.6 TO1-Biotin-stained gel of total RNA isolated from yeast cells expressing either WT U4, U4-Mango-II, or U4-Mango-III. Gel was post-stained with TO1-Biotin. Mango-II and Mango-III oligos were used as positive controls.

Splicing efficiency is not decreased by addition of RNA tags to U4

Although addition of MS2 or Mango tags to U4 impacts growth of yeast only minimally, the efficiency of splicing in in vitro WCE reactions could potentially be decreased. We therefore calculated efficiencies for production of first (lariat-intermediate and mRNA) and second (mRNA) step splicing products from in vitro splicing of a [^{32}P]-labeled model RNA substrate (RP51A). Splicing efficiencies for first or second steps are not lowered by incorporation of U4-Mango-III in spliceosomes (**Fig. 3.7A,B**). However, U4-iMango leads to a decrease in he second-step splicing efficiency (**Fig. 3.7C,D**). First and second step splicing efficiencies are not lowered by U4-MS2 constructs, except for U4-MS2 Short where the second step splicing efficiency is decreased (**Fig. 3.8A,B**). It is not surprising to see splicing defects in the U4-MS2 Short strain given that it had temperature sensitivity when the other strains did not (**Fig. 3.4A**).

To ensure that the presence of either TO1-Cy5 or tdMCP-SNAP for U4-Mango and U4-MS2 constructs, respectively, did not impact splicing efficiency we conducted in vitro splicing assays in the presence TO1-Cy5 or tdMCP-SNAP. The addition of TO1-Cy5 to yWCEs made from yeast expressing either U4-Mango-II or U4-iMango at increasing concentrations did not decrease splicing efficiency of either step of splicing for either extract (**Figure 3.7E**). Addition of tdMCP-SNAP at 20 nM to yWCE made from yeast with tagged Prp3, Brr2, and Snu114 expressing either U4-MS2 Short or U4-MS2 Long did not lower splicing efficiency of either step (**Figure 3.8A,B**). The decreased second step splicing efficiency seen with U4-MS2 Short was also present in splicing assays with tdMCP-SNAP present. A caveat of these assays is that it is unknown if TO1-Cy5 was bound to Mango or tdMCP-SNAP was bound to U4-MS2.

Preliminary single molecule microscopy shows little spot accumulation for TO1-Cy5 under low ATP conditions

We have shown that insertion of Mango and MS2 tags into the 3 SL of the U4 snRNA are generally well-tolerated by yeast; however, it remains unclear whether a single copy of either of these tags would be suitable for single molecule visualization of the U4 snRNA during spliceosome activation. Preliminary experiments using TO1-Cy5 in yWCE with U4 Mango-II yielded unexpected results. ATP was added to WCEs at various concentrations either at 2 mM (high), 50 μM (low), or depleted by endogenous hexokinase activity with the addition of 2 mM glucose (see Methods; Hoskins et al. 2016). Spliceosomes stall prior to activation at low ATP conditions whereas under ATP-depleted conditions spliceosomes are expected to be unable to complete initial ATP-dependent steps of spliceosome assembly (Hoskins et al. 2011; Hoskins et al. 2016). Spot accumulation on the slide surface for Cy5 was never seen under high or low ATP conditions; however, Cy5 spots did somewhat accumulate under ATP depleted conditions (**Fig. 5.9A,B**). Dy549 spots (Prp3) accumulated under low ATP conditions as expected,

but also under ATP depleted conditions like Cy5 (**Fig. 5.9A,B**). Accumulation of spots in either channel was not observed when RNA was not present (**Fig. 5.9C**).

Initial single molecule experiments with the U4-MS2 variants likewise yielded less than ideal results. Very few spots accumulated under low ATP conditions and observations were independent of whether RNA was present on the slide surface. The addition of fluorescently labeled MCP-SNAP protein at low concentrations produced high background, effectively lowering the signal to noise ratio in these experiments. MCP-SNAP proteins (data not shown), despite the introduction of destabilizing mutations, appeared to form large aggregates. Optimization of both the concentration of MCP-SNAP and dispersion of the aggregates will be necessary to collect high quality single molecule datasets in future experiments.

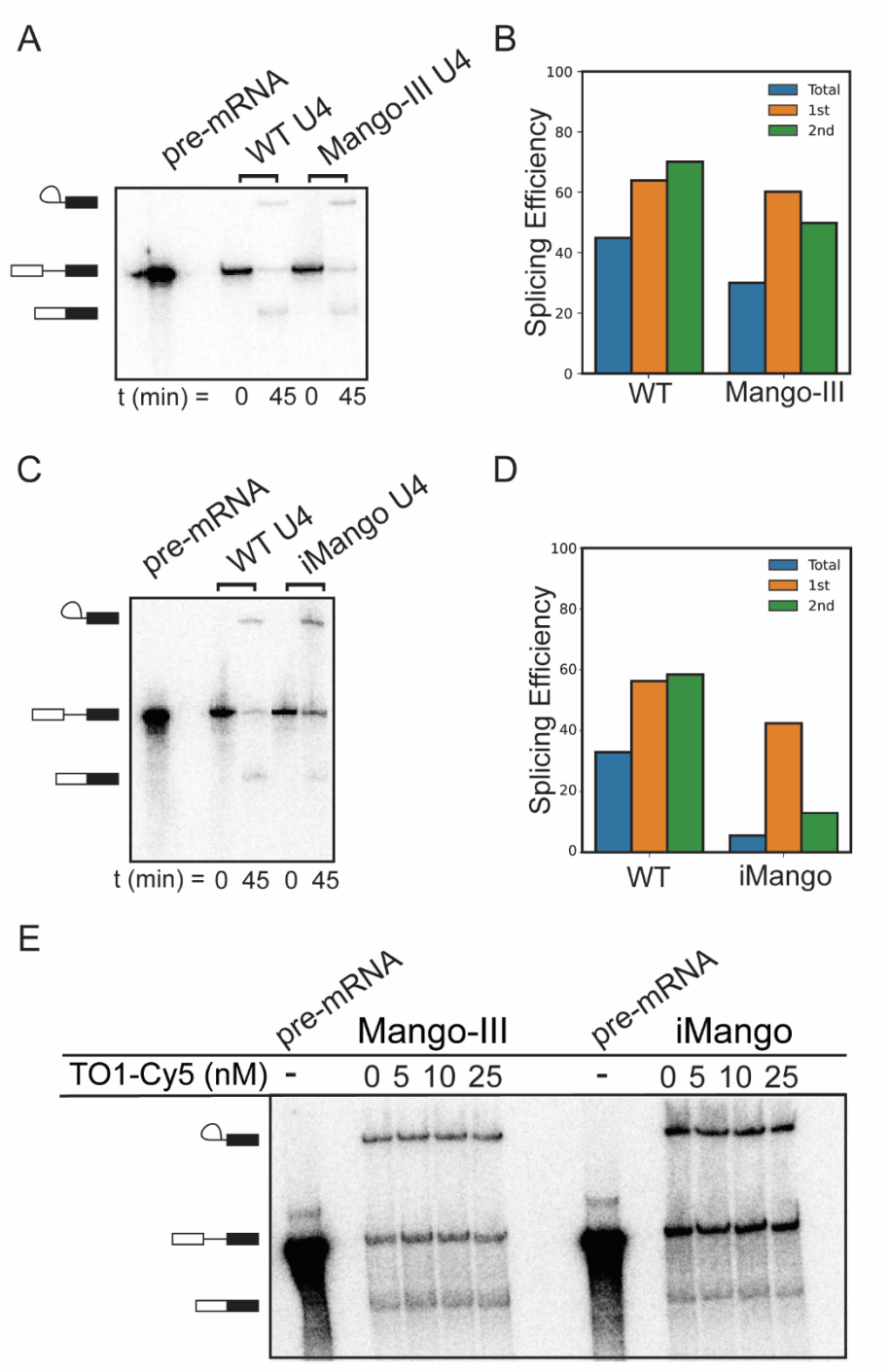

Figure 3.7 WCE made from yeast expressing U4-Mango splices well in the presence of the TO1-Cy5 ligand. A) Gel of splicing products after a 45 min splicing reaction using WCE made from yeast expressing U4-Mango-III variant. **B)** Quantitation of band intensities from gel in A. **C)** Gel of splicing products after a 45 min splicing reaction using WCE made from yeast expressing U4-Mango-III variant. **D)** Quantitation of band intensities from gel in B. **E)** Splicing assay using extracts made from yeast expressing either U4-Mango-III or U4-iMango with TO1-Cy5 present in 0, 5, 10, or 25 nM concentrations.

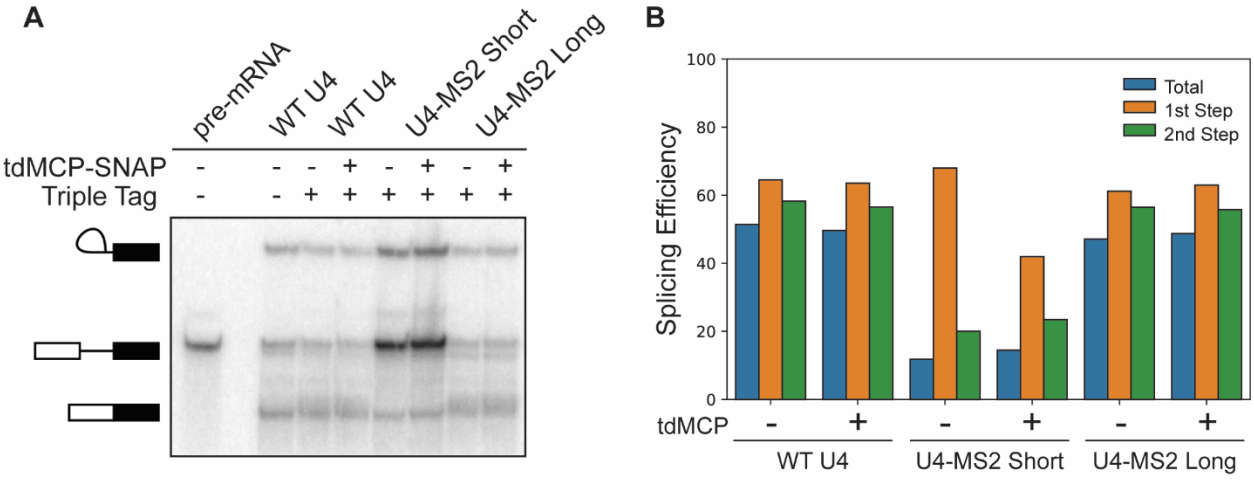

Figure 3.8 Presence of tdMCP in WCE does not affect splicing efficiency of extracts prepared with yeast expressing U4-MS2 Long. A) Gel of splicing products after 45 min reaction with WCE made from yeast expressing U4-MS2 variants. Viable transformants in the triply tagged strain background (Prp3-SNAP, Brr2-DHFR, Snu114-DHFR) were used. tdMCP-SNAP is either present or absent at a 20nM concentration in extracts during splicing. **B)** Quantitation of band intensities in splicing gel. See Materials and Methods.

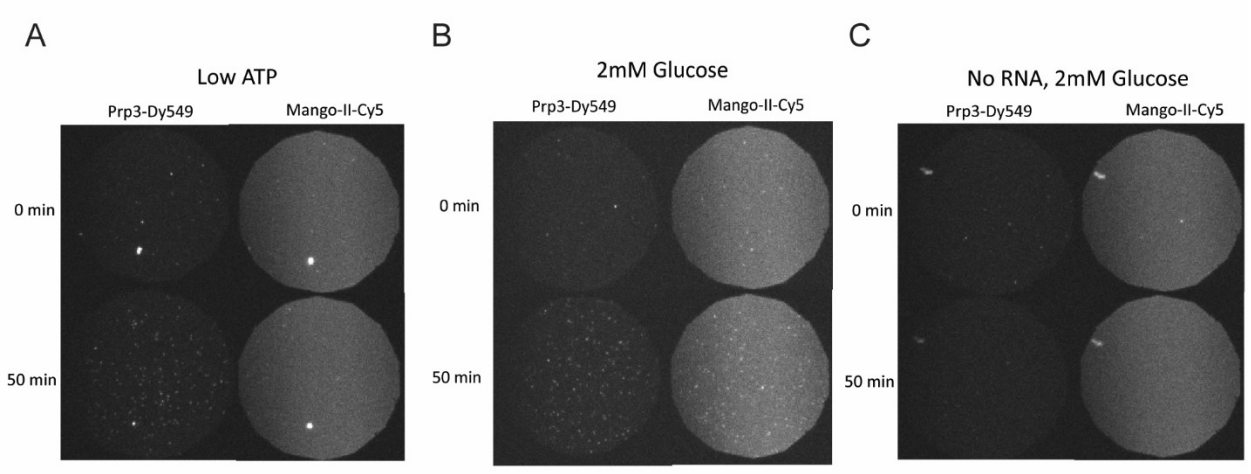

Figure 3.9 Spots only accumulate in the Cy5 channel under ATP depleted conditions. A) CoSMoS fields of view for green channel (Prp3-Dy549, left) and red channels (U4-Mango-II, right) at 0- and 50-min times points either under A) low ATP (50 uM) with RNA present, B) 2mM glucose (ATP depleted) with RNA present, or C) 2mM glucose (ATP depleted) with RNA absent. Experiments under high ATP conditions (2mM ATP) were not conducted at the same time as these data and thus, are not shown.

3.4 Discussion

We have shown that that U4 snRNA 3′ SL is amenable to insertion of two RNA tags, the MS2 SL and Mango RNA aptamer. Variants of both U4-MS2 and U4-Mango are viable in yeast and therefore can function in spliceosomes. However, yeast expressing U4-MS2 have cold sensitivities and, in the case of U4-MS2 Short, temperature sensitivities. Growth of U4-Mango strains was not examined. DHFR tagging of Brr2 causes temperature sensitivity and affects viability and growth of U4-MS2 variants. In strains where Brr2-DHFR, Snu114-DHFR, and Prp3-SNAP are expressed, expression of U4 MS2 3′ results in synthetic lethality. Additionally, U4-MS2 Short and Long are more cold sensitive compared to a WT strain background. Splicing efficiency of U4-Mango-II or U4-MS2 Long in the triply-tagged strain was not decreased from WT at either step. However, U4-MS2 Short and U4-iMango showed decreases in splicing efficiency from WT. U4-Mango-II or U4-MS2 3′ were not tested. The addition of ligands for either Mango or MS2 does not lower splicing efficiency at the conditions tested. It remains to be seen if the ligands are bound to their targets during splicing, although purified tdMCP-SNAP was shown to be functional for binding an MS2 fluorescent oligo in vitro.

Preliminary microscopy revealed several unusual properties of tagged U4 extracts. Spots for TO1-Cy5 ligands only accumulate when ATP is fully depleted, whereas spots for Dy549 labeled Prp3 accumulate under both low and depleted ATP conditions. An explanation for this could be that these extracts have low endogenous hexokinase activity, which depletes ATP by phosphorylation of glucose. Future experiments may benefit from the addition of recombinant hexokinase to extracts. When RNA is not present, neither set of spots accumulates. Further replicates and analysis of single molecule experiments are needed to evaluate the suitability of single RNA tags within the U4 3′ SL for single molecule experiments.

U4-MS2 Short increases growth defects and decreases splicing efficiency

U4-MS2 Short and Long vary by seven additional wobble and Watson-Crick base pairs within the lower stem of the 3′ SL. The MS2 SL is composed mainly of more-stable Watson-crick base pairs than the endogenous U4 stem and replacement likely results in hyperstability. Whether stabilization of this region deleteriously affects Brr2 loading onto the U4 snRNA is unknown. It would be interesting to further characterize defects in snRNP assembly within these mutants. Additionally, spliceosome activation kinetics may differ between spliceosomes with U4-MS2 Short and Long variants which could be compared in single molecules experiments. Ideally, single molecule experiments will be conducted from WCE made with strains expressing untagged Brr2 to remove confounding factors introduced by the tag.

Strains expressing U4-MS2 Short in the absence of other tagged proteins have an interesting growth phenotype of being both cold and heat sensitive. Additionally, splicing efficiency of the second step, and not the first, is mainly lowered. A decrease in the second

step of splicing, may result from more time needed to fully activate spliceosomes. Therefore, spliceosome activation kinetics may differ between spliceosomes with U4-MS2 Short and Long variants which could be compared in single molecules experiments.

Processing defects might cause synthetic lethality of U4-MS2 3′ with Brr2-DHFR

A stem loop is present in human U4 snRNAs after the Sm ring binding site that is absent in yeast. The introduction of this stem as a scaffold for MS2 in U4 MS2 3′ was viable in yeast that did not have tagged proteins. However, tagging of Prp3, Brr2, or Snu114 in combination with U4 MS2 3′ was synthetically lethal (data not shown). Perhaps loading of Brr2 onto the U4 snRNA is blocked by addition of this stem. Defects in 3′ end processing could also be present that, in combination with Brr2 loading defects, result in lethality. It remains unclear if the MS2 stem in U4-MS2 3′ remains intact after 3′ end trimming by the exosome. Further experiments are needed to clarify if the SL is present in mature snRNPs.

Properties of yWCE may result in unwinding of Mango when ATP is present

Accumulation of spots only under ATP depleted conditions, suggests that extracts could be high in ATP. Only under ATP depleted conditions could signal from TO1-Cy5 be seen (**Figure 3.9**). The ATP concentration could therefore be higher in these strains than anticipated and addition of a lower concentration of ATP or no ATP might be required to result in comparable conditions to data collected at low ATP concentrations (50μM) in previous publications (Hoskins et al. 2016). To determine if TO1-Cy5 signal corresponds to labeling of U4-Mango-II snRNA under ATP depleted conditions, analysis of Cy5 spot colocalization with pre-mRNA and/or Prp3-Dy549 spots will need to be conducted.

Conclusions

Strategies for endogenous fluorescent labeling of the U4 snRNA presented in this work may provide a method of characterizing Brr2 activity during activation and allow for discernment of sequential or simultaneous release of the U4 snRNA from the di-snRNP specific protein Prp3. Further optimization tdMCP-Dy649 concentrations added to U4-MS2 WCEs is needed before single molecule experiments can be conducted. The phenotypes seen within variants of the U4-MS2 provide an opportunity for interrogating snRNP formation in yeast, including loading of Brr2 onto the U4 snRNA.

3.5 Methods

Yeast transformation

Yeast strains and plasmids used in this study are described in Tables 3.1 and 3.2, respectively. Yeast transformation, plasmid shuffling (also 5-FOA selection), and growth were carried out using standard techniques and media (Treco and Lundblad 1993; Sikorski and Boeke 1991).

Yeast growth assays

For plated temperature growth assays, yeast strains were grown overnight in 5 mL YPD to stationary phase and subsequently diluted to 0.5 OD_{600}/mL in 10% (v/v) glycerol. Cultures were further diluted 1:10 1:100, and 1:100 in 10% glycerol before stamping diluted cultures on YPD plates and placing at 16, 23, 30, or 37°C. Plates were incubated for the number of days indicated in the figure legend following stamping.

For liquid culture growth assays, yeast were diluted from 5mL YPD overnight cultures to 0.02 OD/mL in YPD and grown in a 48-well plate for 48 h. Optical density measurements were taken every 10 min. Doubling times were calculated from 48 h data as described previously (Lipinski et al. 2022).

Plasmids

Plasmids used in these experiments are listed in Table 3.2. U4-Mango II, Mango-III, and iMango were synthesized by GeneWiz and cloned into pUC57 vectors. Plasmids containing U4 variants were digested with XhoI and SacI. The XhoI/SacI digested product was ligated into a XhoI/SacI-digested pRS313 backbone. For U4-MS2 variants, WT U4 sequences amplified by PCR with primers KL_036_U4_MS2Oper_F paired with either KL_034_U4_MS2Oper_R1 or KL_035_U4_MS2Oper_R2 for creation of U4-MS2 Short and U4-MS2 Long, respectively. PCR amplification of WT U4 sequences with KL_037_U4_MS2Oper_3end_F paired with KL_038_U4_MS2Oper_3end_R was used to make U4-MS2 3′. Plasmid sequences were verified by Plasmidsaurus.

SNAP and DHFR tagging

Prp3, Brr2, and Snu114 were tagged as described in Hoskins et al, 2016. Briefly, double strand DNAs were prepared by PCR using Herculase (Agilant). dsDNAs were purified and transformed into yeast as described above. Viable yeast colonies grown on selective media supplemented with either Hygromycin, CloNAT, or Phleomycin were checked for addition of the SNAP or DHFR tag by colony PCR.

MCP-SNAP purification and fluorescent labeling

The tdMCP sequence was cloned into a pET-28a vector with the addition of a 6x-His tag and TEV cleavage site (see **Table 3.2**) and transformed into BL21 Star (DE3) cells (Cat# C601003, ThermoFisher Scientific, Waltham, MA, USA). Cells were cultured at

37°C in 1L of 2xYT supplemented with Kan (35 µg/mL) until OD600 = 0.6 - 0.8 then expression was induced with 0.5 mM IPTG and cultures were incubated overnight at 16°C. Cell pellets were washed with buffer (50 mM Tris pH 7.4, 500 mM NaCl, 10% sucrose) before freezing at -80°C.

For cell lysis, pellets were resuspended in lysis buffer 20 mM HEPES•NaOH pH 7.5, 1 M NaCl, 1 M Urea, 5 mM TCEP, 1 mM PMSF, 0.1 mg/ml lysozyme) and incubated for 30 min at RT with stirring. DNase I (0.5 U/mL) and $MgCl_2$•6H2O (5 mM) were added and incubated for an additional 30 min. The lysis solution was sonicated for 5 min (3 s on, 12 s off). Lysates were clarified with two centrifugations at 30000 xg at 4°C for 30 min. Clarified lysates were diluted in IMAC Buffer A (20 mM HEPES•NaOH pH 7.5, 1 M NaCl, 1 M Urea, 1 mM TCEP), then loaded onto a HisTrap HP 5mL Ni-NTA column (Cytiva) and eluted with a gradient of IMAC Buffer B (20 mM HEPES•NaOH pH 7.5, 1 M NaCl, 1 M Urea, 300 mM Imidazole, 1 mM TCEP). Eluted fractions were pooled in dialysis cassettes and with MBP-SuperTEV protease and dialyzed overnight against Dialysis 1 Buffer (20 mM HEPES•NaOH pH 7.5, 0.25 M NaCl, 1 M Urea, 1 mM TCEP). TEV and cleaved 6x-His tag were removed with an additional IMAC step as described above. Pooled protein fractions were dialyzed against Dialysis 2 Buffer (20 mM HEPES•NaOH pH 7.5, 0.05 M NaCl, 1 mM TCEP) overnight.

Dialyzed protein was loaded onto a Heparin column (Cytiva) equilibrated in Heparin Buffer A (20 mM HEPES•NaOH pH 7.5, 0.05 M NaCl, 1 mM TCEP) and eluted with a gradient of Heparin Buffer B (20 mM HEPES•NaOH pH 7.5, 1 M NaCl, 1 mM TCEP). For SNAP tag fluorescent labeling, tdMCP-SNAP protein was incubated with excess SNAP Surface 649 reagent (NEB) according to manufacturer’s instructions. Excess dye was removed with Heparin column flow-through.

Primer extension

Yeast cultures were inoculated from stationary phase saturated cultures grown overnight in 5mL YPD media. When diluted cultures reached OD_{600} = 0.6 − 0.8, 10 OD_{600} units were collected by centrifugation. The MaterPure Yeast Total RNA Purification Kit (Epicentre, Madison, WI) protocol with minor changes as previously described was was used for isolation of total RNA and depletion of contaminating DNA (Carrocci et al., 2017). IR700 dye conjugated probes (Integrated DNA Technologies, Skokie, IL) were used for primer extension of the snRNAs (2 pmol each U2RTALL, U4RTALL, U6D) (Dobbyn et al 2004; Carrocci et al., 2017; van der Feltz et al., 2021). Primer extension products were visualized on a 7% (w/v) 19:1 acrylamide:bis-acrylamide denaturing gel (8M urea; 1X TBE: 0.13 M tris (pH 7.6), 45 mM boric acid, 2.5 mM EDTA; 42 cm × 22 cm × 0.75 mm). Products were separated using 35W for 80 min at RT. An Amersham Typhoon NIR laser scanner (Cytiva) was used for gel imaging, and band intensities were quantified with Image J (version 1.53v, 2022).

Splicing assays

Yeast splicing extracts and RP51A [^{32}P]-labeled pre-mRNA substrates were prepared as previously described (Crawford et al. 2008). Protease inhibitors (0.5 mM PMSF, 1µg/mL pepstatin, 1µg/mL leupeptin) were added prior to resuspension of cells in buffer before flash freezing and ball milling. Splicing assays were conducted at room temperature using 40% (v/v) WCE and 0.2 nM pre-mRNA substrate (Crawford et al. 2008). [^{32}P]-labeled RNAs were isolated as previously described and separated on a 12% (w/v) denaturing PAGE gel. Gels were exposed overnight to phosphor imaging screen and imaged with an Amersham Typhoon laser scanner (Cytiva). Band intensities were quantitated in ImageJ. Splicing efficiencies of the first and second steps were calculated using the quantitated band density values in Equations 1 and 2 as previously described (Mayerle and Guthrie 2016).

$$1st\ Step\ Efficiency = \frac{\text{lariat intermediate}}{\text{pre−mRNA+lariat interemdiate+mRNA}} \quad (1)$$

$$2nd\ Step\ Efficiency = \frac{\text{mRNA}}{\text{lariat interemdiate+mRNA}} \quad (2)$$

$$Total\ Splicing\ Efficiency = \frac{\text{mRNA}}{\text{pre−mRNA+ lariat interemdiate+mRNA}} \quad (3)$$

TO1-Biotin gel staining

Total RNA was isolated from yeast cells as described for primer extension. TO1-Biotin post-staining protocol was performed as described in Yaseen et al. 2019, briefly total RNA was mixed 1:1 with clear formamide loading dye and heated to 95°C for 5 min to denature RNAs. Samples were removed from the PCR machine and allowed to air-cool at RT for several min. Samples were loaded onto a denaturing 8% PAGE 6M urea gel and separated using 28W for 30 min at RT. The gel was placed in a container with 100mL of the TO1-Biotin post-staining solution (140 mM KCl, 10 mM phosphate buffer (pH 7.2), 0.0005% Tween 20 (v/v), and 20 nM TO1-Biotin) and rocked for 30 min. Post-staining solution was replaced halfway through incubation. Gels were rinsed briefly with water before imaging using an Amersham Typhoon laser scanner (Cytiva).

CoSMoS

Fluorescent pre-mRNA molecules were prepared by splinted ligation of a [^{32}P]-labeled, capped RP51A transcript to a 2'-*O*-methyl oligonoucleotide labeled with Alexa488 (IDT) as previously described (Crawford et al., 2008). Slides and buffers including oxygen scavengers, but not triplet quenchers were prepared as previously described (Hoskins et al., 2016). Fluorescent beads were included for stage drift correction as fiducial markers (TransFluoSpheres, streptavidin-labeled). Images were collected on a micromirror TIRF microscope at room temperature as previously described (Hoskins et al. 2016). Laser

powers were typically set to 1-1.4 mW for 488 nm, 600 uW for 532 nm, and 440 uW for 633 nm. Data were collected for three-color experiments with a 1s simultaneous exposure of 532 and 633 nm lasers followed by a 4s spacing for the duration of the experiment.

Table 3.1 Yeast Strains

Strain	Parent	Genotype	Reference
CJM000	ANK640	MATa, snr6::LEU2, snr14::trp1::ADE2, trp1, ura3, lys2, his3, ade2, [pRS316-U4wt-U6mini]	McManus et al. 2007
yAAH0431	CJM000	MATa, snr6::LEU2, snr14::trp1::ADE2, trp1, ura3, lys2, his3, ade2, + [pRS313 U4 WT] + [pRS314 WT U6]	Lipinski et al. 2022
yAAH1361	CJM000	MATa, snr6::LEU2, snr14::trp1::ADE2, trp1, ura3, lys2, his3, ade2, + [pRS313 U4 WT] + [pRS314 U6 PolII Sm]	Lipinski et al. 2022
yAAH1391	CJM000	MATa, snr6::LEU2, snr14::trp1::ADE2, trp1, ura3, lys2, his3, ade2, [pRS316-U4wt-U6mini] + BRR2::DHFR-hyg + SNU114::DHFR-phleo + PRP3::fSNAP-ClonNAT	This study
yAAH1442	CJM000	MATa, snr6::LEU2, snr14::trp1::ADE2, trp1, ura3, lys2, his3, ade2, + [pRS313 U4 WT] + [pRS314 U6 PolII Lsm]	Lipinski et al. 2022
yAAH1550	CJM000	MATa, snr6::LEU2, snr14::trp1::ADE2, trp1, ura3, lys2, his3, ade2, + [pRS413-U4-Mango V20] + [pRS314 WT U6]	This study
yAAH1551	CJM000	MATa, snr6::LEU2, snr14::trp1::ADE2, trp1, ura3, lys2, his3, ade2, BRR2::DHFR-hyg, SNU114::DHFR-phleo, PRP3::fSNAP-ClonNAT, + [pRS413-U4-Mango V20] + [pRS314 WT U6]	This study
yAAH2279	CJM000	MATa, snr6::LEU2, snr14::trp1::ADE2, trp1, ura3, lys2, his3, ade2, [pRS413 U4Mango-III] + [pRS314 WT U6]	This study
yAAH2280	CJM000	MATa, snr6::LEU2, snr14::trp1::ADE2, trp1, ura3, lys2, his3, ade2, [pRS413 U4 iMango] + [pRS314 WT U6]	This study
yAAH2516	CJM000	MATa, snr6::LEU2, snr14::trp1::ADE2, trp1, ura3, lys2, his3, ade2, BRR2::DHFR-hyg, SNU114::DHFR-phleo, PRP3::fSNAP-ClonNAT, + [pRS413 U4 Mango-III] + [pRS314 WT U6]	This study
yAAH2517	CJM000	MATa, snr6::LEU2, snr14::trp1::ADE2, trp1, ura3, lys2, his3, ade2, BRR2::DHFR-hyg, SNU114::DHFR-phleo, PRP3::fSNAP-ClonNAT, + [pRS413 U4 iMango] + [pRS314 WT U6]	This study
yAAH3329	CJM000	MATa, snr6::LEU2, snr14::trp1::ADE2, trp1, ura3, lys2, his3, ade2, BRR2::DHFR-hyg,	This study

		SNU114::DHFR-phleo, PRP3::fSNAP-ClonNAT + [pRS313 U4-MS2 Short]+ [pRS314 U6 WT]	
yAAH3330	CJM000	MATa, snr6::LEU2, snr14::trp1::ADE2, trp1, ura3, lys2, his3, ade2, BRR2::DHFR-hyg, SNU114::DHFR-phleo, PRP3::fSNAP-ClonNAT + [pRS313 U4-MS2 Long] + [pRS413 U6 WT]	This study
yAAH3340	CJM000	MATa, snr6::LEU2, snr14::trp1::ADE2, trp1, ura3, lys2, his3, ade2 + [pRS313 U4-MS2 Short] + [pRS314-U6wt]	This study
yAAH3341	CJM000	MATa, snr6::LEU2, snr14::trp1::ADE2, trp1, ura3, lys2, his3, ade2 + [pRS313 U4-MS2 Long] + [pRS314-U6wt]	This study
yAAH3342	CJM000	MATa, snr6::LEU2, snr14::trp1::ADE2, trp1, ura3, lys2, his3, ade2 + [pRS313 U4-MS2 3′] + [pRS314-U6wt]	This study

Table 3.2 Plasmids

Plasmid	Description	Reference
pAAH0412	pRS314 WT U6/TRP/AMP	McManus et al. 2007
pAAH0413	pRS313 WT U4/HIS/AMP	McManus et al. 2007
pAAH0667	pRS413 U4 Mango-II/HIS/AMP	This work
pAAH1124	pUC57 U4 Mango-III +/-300 bp/KAN	This work
pAAH1125	pUC57 U4 iMango +/- 300 bp/KAN	This work
pAAH1131	pRS413 U4 Mango-III +/-300 bp/HIS/AMP	This work
pAAH1132	pRS413 U4 iMango +/- 300 bp/HIS/AMP	This work
pAAH1525	pET28a His6_TEV_tdMS2_SNAPf/KAN	This work
pAAH1526	pRS313 U4-MS2 Short/HIS/AMP	This work
pAAH1527	pRS313 U4-MS2 Long/HIS/AMP	This work
pAAH1528	pRS313 U4-MS2 3'/HIS/AMP	This work

Table 3.3 RNA and DNA Oligonucleotides

Name	Sequence (5′ to 3′)
KL_0031_MangoII	rGrGrCrArCrGrUrArCrGrArArGrGrArGrArGrGrArGr ArGrGrArArGrArGrGrArGrArGrUrArCrGrUrGrC
KL_0032_MangoIII	rGrGrCrArCrGrUrArCrGrArArGrGrArArGrGrArUrUr GrGrUrArUrGrUrGrGrUrArUrArUrUrCrGrUrArCrGrU rGrCrC
KL_033_tdMS2_SNAP_F	GCTATCGCTGCGAATAGTGGTATCTATGGTTCTG GCGGATCAGGGATGGACAAAGACTGC
KL_034_U4_MS2Oper_R1	GAAGTATTCACATGGGTGATCCTCATGTGAACAC CGAATTGACCATGAGGAGAC
KL_035_U4_MS2Oper_R2	GAAACATGGGTGATCCTCATGTCGAATTGACCAT GAGGAGAC
KL_036_U4_MS2Oper_F	AAGACTATGTAGGGAATTTTTGGAATACCTTTTTT CATTACCG
KL_037_U4_MS2Oper_3end_F	ATGAGGATCACCCATGTGAGACTGAATACCTTTT TTCATTACCG
KL_038_U4_MS2Oper_3end_R	GTGAGACTGCCAAAAATTCCCTACATAGTCTTG
KL_039_U4_F	CCGAATTCGATATCAAGCTTATCG
KL_040_U4_R	GGGGATCCACTAGTTCTAGAG
KL_0041_MS2	/5Cy3/rCrArC rArUrG rArGrG rArUrC rArCrC aCrArU rGrUrU
KL_0042_Cy3_MS2_Bio	/5Cy3/rCrArC rArUrG rArGrG rArUrC rArCrC aCrArU rGrUrU /3Bio/
KL_0043_Cy3_MS2a_Bio	/5Cy3/rCrArC rArUrG rGrGrA UrCrA rCrCrC aArUrG UrU /3Bio/
U4RTALL	/5IRD700/ GGT ATT CCA AAA ATT CCC TAC ATA GTC
U6D	/5IRD700/ AAA ACG AAA TAA ATC TCT TTG
U2RTALL124	/5IRD700/TTTGGGTGCCAAAAAATGTGTATT GTAAC

3.6 References

Bordonne R, Banroques J, Abelson J, Guthrie C. 1990. Domains of yeast U4 spliceosomal RNA required for PRP4 protein binding, snRNP-snRNP interactions and pre-mRNA splicing in vivo. *Genes Dev*. **4**:1185– 1196. doi: 10.1101/gad.4.7.1185

Crawford DJ, Hoskins AA, Friedman LJ, Gelles J, Moore MJ. 2008. Visualizing the splicing of single pre-mRNA molecules in whole cell extract. *RNA* **14**: 170–179. doi: 10.1261/rna.794808

Carrocci TJ, Zoerner DM, Paulson JC, Hoskins AA. 2017. SF3b1 mutations associated with myelodysplastic syndromes alter the fidelity of branchsite selection in yeast. *Nucleic Acids Res* **45**: 4837–4852. doi: 10.1093/nar/gkw1349

Cawte, AD, Unrau PJ, Rueda DS. 2020. Live cell imaging of single RNA molecules with fluorogenic Mango II arrays. *Nat Commun* **11:**1283. doi: 10.1038/s41467-020-14932-7

Cornilescu G, Didychuk AL, Rodgers ML, Michael LA, Burke JE, Montemayor EJ, Hoskins AA, Butcher SE. 2016. Structural analysis of multi-helical RNAs by NMR-SAXS/WAXS: application to the U4/U6 di-snRNA. *J. Mol. Biol.* **428**:777–789. doi: 10.1016/j.jmb.2015.11.026

Dobbyn HC, O'Keefe RT. 2004. Analysis of Snu13p mutations reveals differential interactions with the U4 snRNA and U3 snoRNA. *RNA* **10**:308–320. doi: 10.1261/rna.5970404

Dolgosheina EV, Jeng SCY, Panchapakesan SSS, Cojocaru R, Chen PSK, Wilson PD, Hawkins N, Wiggins PA, and Unrau PJ. 2014. RNA Mango Aptamer-Fluorophore: A Bright, High-Affinity Complex for RNA Labeling and Tracking. *ACS Chem. Biol.* 9:(10) 2412–2420. doi: 10.1021/cb500499x

Fabrizio P, McPheeters DS, Abelson J. 1989. In vitro assembly of yeast U6 snRNP: a functional assay. Genes Dev 3: 2137–2150. doi: 10.1101/gad.3.12b.2137

Fu X, Kaur H, Rodgers ML, Montemayor EJ, Butcher SE, Hoskins AA. 2022. Identification of transient intermediates during spliceosome activation by single molecule fluorescence microscopy. *PNAS* **119**(48): e2206815119. doi: 10.1073/pnas.2206815119

Hahn D, Kudla G, Tollervey D, Beggs JD. 2012. Brr2p-mediated conformational rearrangements in the spliceosome during activation and substrate repositioning. *Genes Dev*. **26**:2408–2421. doi: 10.1101/gad.199307.112

Hardin JW, Warnasooriya C, Kondo Y, Nagai K, Rueda D. 2015. Assembly and dynamics of the U4/U6 di-snRNP by single-molecule FRET. *Nucleic Acids Res* **43**:10963–10974. doi: 10.1093/nar/gkv1011

Hayduk AJ, Stark MR, Radar SD. 2012. In vitro reconstitution of yeast splicing with U4 snRNA reveals multiple roles for the 3′ stem–loop. *RNA* **18**: 1075-1090. doi: 10.1261/rna.031757.111

Hoskins AA, Rodgers ML, Friedman LJ, Gelles J, Moore MJ. 2016. Single molecule analysis reveals reversible and irreversible steps during spliceosome activation. *eLife* **5**: e14166. doi: 10.7554/eLife.14166

Hu J, Xu D, Schappert K, Xy Y, Friesen JD. 1995. Mutational Analysis of Saccharomyces cerevisiae U4 Small Nuclear RNA Identifies Functionally Important Domains. *Molecular and Cellular Biology* **15**(3): 1274–1285. doi: 10.1128/MCB.15.3.1274

Karaduman RR, Fabrizio PR, Hartmuth KR, Urlaub HR, Lührmann RR. 2006. RNA structure and RNA-protein interactions in purified yeast U6 snRNPs. *J. Mol. Biol.* **356**:15–15. doi: 10.1016/j.jmb.2005.12.013

Lipinski KA, Chi J, Chen X, Hoskins AA, Brow DA. 2022. Yeast U6 snRNA made by RNA polymerase II is less stable but functional. *RNA* **28**(12):1606-1620. doi: 10.1261/rna.079328.122

Lipinski KA, Senn KA, Zeps NJ, Hoskins AA. 2023. Biochemical and genetic evidence supports Fyv6 as a second-step splicing factor in Saccharomyces cerevisiae. *RNA* **29**: 1792-1802. doi: 10.1261/rna.079607.123

Liu S, Mozaffari-Jovin S, Wollenhaupt J, Santos KF, Theuser M, Dunin-Horkawicz S, Fabrizio P, Bujnicki JM, Luhrmann R, Wahl MC. 2015. A composite double-/single-stranded RNA-binding region in protein Prp3 supports tri-snRNP stability and splicing. *Elife* **4**:e07320. doi: 10.7554/eLife.07320

Lowary PT, Uhlenbeck, OC. 1987. An RNA mutation that increases the affinity of an RNA–protein interaction. *Nucleic Acids Res.* **15**, 10483–10493. doi: 10.1093/nar/15.24.10483

McManus CJ, Schwartz ML, Butcher SE, Brow DA. 2007. A dynamic bulge in the U6 RNA internal stem-loop functions in spliceosome assembly and activation. *RNA* **13**: 2252-2265. doi: 10.1261/rna.699907.

McPheeters DS, Fabrizio P, Abelson J. 1989. In vitro reconstitution of functional yeast U2 snRNPs. *Genes Dev* **3**: 2124–2136. doi: 10.1101/gad.3.12b.2124

Montemayor EJ, Curran EC, Liao HC, Andrews KL, Treba CN, Butcher SE, Brow DA. 2014. Core structure of the U6 small nuclear ribonucleoprotein at 1.7-Å resolution. *Nat. Struct. Mol. Biol.* **21**:544–551. doi: 10.1038/nsmb.2832

Nguyen THD, Galej WP, Bai X-C, Oubridge C, Newman AJ Scheres SHW, Nagai K. 2016. Cryo-EM structure of the yeast U4/U6.U5 tri-snRNP at 3.7 Å resolution. *Nature* **530**:298–302. doi: 10.1038/nature16940

O'Keefe RT, Norman C, Newman AJ. 1996. The invariant U5 snRNA loop 1 sequence is dispensable for the first catalytic step of pre-mRNA splicing in yeast. *Cell* **86**: 679–689. doi: 10.1016/S0092-8674(00)80140-3

Peabody DS Ely KR. 1992. Control of translational repression by protein-protein interactions. *Nucleic Acids Res.* **20**:1649-1655. doi: 10.1093/nar/20.7.1649

Rodgers ML, Didychuk AL, Butcher SE, Brow DA, Hoskins AA. 2016. A multi-step model for facilitated unwinding of the yeast U4/U6 RNA duplex, *Nucleic Acids Research* **44**(22):10912–10928. doi:10.1093/nar/gkw686

Shi Y. 2017. Mechanistic insights into precursor messenger RNA splicing by the spliceosome. *Nat. Rev. Mol. Cell. Biol.* **18**: 655–670. doi: 10.1038/nrm.2017.86

Sikorski RS, Boeke JD. 1991. In vitro mutagenesis and plasmid shuffling: from cloned gene to mutant yeast. *Methods Enzymol* **194**: 302–318. doi: 10.1016/0076-6879(91)94023-6

Tarn WY, Lee KR, Cheng SC. 1993a. Yeast precursor mRNA processing protein PRP19 associates with the spliceosome concomitant with or just after dissociation of U4 small nuclear RNA. *Proc. Natl. Acad. Sci.* **90**: 10821–10825. doi: 10.1073/pnas.90.22.10821

Tarn WY, Lee KR, Cheng SC. 1993b. The yeast PRP19 protein is not tightly associated with small nuclear RNAs, but appears to associate with the spliceosome after binding of U2 to the pre-mRNA and prior to formation of the functional spliceosome. *Mol. Cell. Biol.* **13**: 1883–1891. doi: 10.1128/mcb.13.3.1883-1891.1993

Theuser M, Höbartner C, Wahl MC, Santos KF. 2016. Substrate-assisted mechanism of RNP disruption by the spliceosomal Brr2 RNA helicase. *Proc. Natl. Acad. Sci. U.S.A.* **113**:7798–7803. doi: 10.1073/pnas.1524616113

Townsend C et al. 2020. Mechanism of protein-guided folding of the active site U2/U6 RNA during spliceosome activation. *Science* **370**: eabc3753. doi: 10.1126/science.abc3753

Trachman III RJ, Demeshkina N, Lau M, et al. 2017. Structural basis for high-affinity fluorophore binding and activation by RNA Mango. *Nat Chem Biol* **13**: 807–813. doi: 10.1038/nchembio.2392

Trachman III RJ, Abdolahzadeh A, Andreoni A, Cojocaru R, Knutson JR, Rycklynck M, Unrau PJ, Ferré-D'Amaré AR. 2018. Crystal Structures of the Mango-II RNA Aptamer Reveal Heterogeneous Fluorophore Binding and Guide Engineering of Variants with Improved Selectivity and Brightness. *Biochemistry* **57**(26):3544–3548. doi: 10.1021/acs.biochem.8b00399

Trachman III RJ, Autour A, Jeng SCY, Abdolahzadeh A, Andreoni A, Cojocaru R, Garipov R, Dolgoshenia EV, Knutson JR, Ryckelynck M, et al. 2019. Structure and functional reselection of the Mango-III fluorogenic RNA aptamer. *Nature Chemical Biology* **15**:472–479. doi: 10.1038/s41589-019-0267-9

Treco DA, Lundblad V. 1993. Preparation of Yeast Media. *Current Protocols in Molecular Biology* **23**: 13.1.1–13.1.7. doi: 10.1002/0471142727.mb1301s23

Valegârd K, Murray JB, Stonehouse NJ, van den Worm S, Stockley PG, Liljas L. 1997. The three-dimensional structures of two complexes between recombinant MS2 capsids and RNA operator fragments reveal sequence-specific protein-RNA interactions. *J Mol Biol.* **270**(5):724-38. doi: 10.1006/jmbi.1997.1144.

van der Feltz C, Nikolai B, Schneider C, Paulson JC, Fu X, Hoskins AA. 2021. Saccharomyces cerevisiae Ecm2 modulates the catalytic steps of pre-mRNA splicing. *RNA* **27**: 591–603.doi:10.1261/rna.077727.120

Wan R, Yan C, Bai R, Wang L, Huang M, Wong CCL, Shi Y. 2016. The 3.8 Å structure of the U4/U6.U5 tri-snRNP: insights into spliceosome assembly and catalysis. *Science* **351**:466–475. doi: 10.1126/science.aad6466

Wu B, Chao JA, Singer RH. 2012. Fluorescence fluctuation spectroscopy enables quantitative imaging of single mRNAs in living cells. *Biophys J.* **102**(12):2936-44. doi: 10.1016/j.bpj.2012.05.017.

Yan C, Hang J, Wan R, Huang M, Wong CCL., Shi Y. 2015. Structure of a yeast spliceosome at 3.6-angstrom resolution. *Science* **349**:1182–1191. doi: 10.1126/science.aac7629

Yaseen IM, Ang QR, Unrau PJ. 2019. Fluorescent Visualization of Mango-tagged RNA in Polyacrylamide Gels via a Poststaining Method. *J. Vis. Exp.* **148**: e59112, doi:10.3791/59112.

CHAPTER 4

Biochemical and genetic evidence supports Fyv6 as a second-step splicing factor in *Saccharomyces cerevisiae*

A version of this chapter is published as: Karli A. Lipinski*, Katherine A. Senn*, Natalie J. Zeps, Aaron A. Hoskins. 2023. Biochemical and Genetic Evidence Supports Fyv6 as a Second-Step Splicing Factor in Saccharomyces cerevisiae. *RNA* **29**:1792-1802. doi: 10.1261/rna.079607.123

*Indicates co-first authors

AAH, KAL, and KAS conceptualized the project. KAL, KAS, and NJZ performed experiments and analyzed data. KAL, KAS, and AAH wrote the manuscript. See contributions (pg. xii).

CHAPTER 4: Biochemical and genetic evidence supports Fyv6 as a second-step splicing factor in *Saccharomyces cerevisiae*

4.1 Abstract

Precursor mRNA (pre-mRNA) splicing is an essential process for gene expression in eukaryotes catalyzed by the spliceosome in two transesterification steps. The spliceosome is a large, highly dynamic complex composed of five small nuclear RNAs and dozens of proteins, some of which are needed throughout the splicing reaction while others only act during specific stages. The human protein FAM192A was recently proposed to be a splicing factor that functions during the second transesterification step, exon ligation, based on analysis of cryo-electron microscopy (cryo-EM) density. It was also proposed that Fyv6 might be the *Saccharomyces cerevisiae* functional and structural homolog of FAM192A; however, no biochemical or genetic data has been reported to support this hypothesis. Herein, we show that Fyv6 is a splicing factor and acts during exon ligation. Deletion of *FYV6* results in genetic interactions with the essential splicing factors Prp8, Prp16, and Prp22 and decreases splicing in vivo of reporter genes harboring intron substitutions that limit the rate of exon ligation. When splicing is assayed in vitro, whole-cell extracts lacking Fyv6 accumulate first-step products and exhibit a defect in exon ligation. Moreover, loss of Fyv6 causes a change in 3′ splice site (SS) selection in both a reporter gene and the endogenous SUS1 transcript in vivo. Together, these data suggest that Fyv6 is a component of the yeast spliceosome that influences 3′ SS usage and the potential homolog of human FAM192A.

4.2 Introduction

The removal of introns from precursor mRNA (pre-mRNA) molecules is carried out by the spliceosome, a large macromolecular complex made up of five small nuclear RNAs (snRNAs) and dozens of proteins which assemble de novo on each pre-mRNA substrate. Splicing consists of two, stepwise transesterification reactions in which the 5′ splice site (5′ SS; the boundary between the 5′ exon and the intron) is first cleaved by the formation of a lariat intron (first step) and then the intron is released concomitant with exon ligation by attack of the 5′ exon at the 3′ SS (second step). Spliceosome composition changes dramatically throughout the course of splicing due to the sequential arrivals and departures of different components as well as large-scale conformational changes (Plaschka et al. 2019). This results in the formation of several intermediate complexes with distinct architectures during the reaction, many of which have now been visualized by cryo-electron microscopy (cryo-EM) (Plaschka et al. 2019). Cleavage of the 5′ SS is completed during the transition from the spliceosome B* to the C complex, and exon ligation occurs during the transition between the C* and P (product) complexes. While

some components of the spliceosome remain part of the machine throughout the reaction, others transiently associate, dissociate, or rearrange to interact with the catalytic site only at specific times. Just prior to 5′ SS cleavage, the first-step factors (Cwc25, Isy1, and Yju2) function to juxtapose the 5′ SS and branch site (BS) (Villa and Guthrie 2005; Liu et al. 2007a; Chiu et al. 2009; Wan et al. 2019; Wilkinson et al. 2021). Cwc25 and Isy1 are then released after 5′ SS cleavage, and second-step factors bind (Slu7, Prp18) or are repositioned (Prp17) to facilitate exon ligation (James et al. 2002; Ohrt et al. 2013; Fica et al. 2017; Yan et al. 2017; Plaschka et al. 2019). Proper progression through splicing requires the coordinated association and dissociation of these first- and second-step factors with the active site and these transitions are enabled, in part, by ATP-dependent DExD/H-box helicases. The ATPase Prp16 promotes rearrangement of the spliceosome active site and splicing factor release between the first and second step of splicing (Schwer and Guthrie 1992; Semlow et al. 2016), while Prp22 promotes release of the mRNA product from the spliceosome after exon ligation (Wagner et al. 1998; Schwer 2008).

Recently, a putative new second-step factor (FAM192A or PIP30) was identified for the human spliceosome. The protein was found by fitting unassigned density present in cryo-EM maps of spliceosomes transitioning between conformations competent for the first or second steps (Zhan et al. 2022). Depletion of FAM192A from human nuclear extracts reduced in vitro splicing, but adding purified protein back did not restore this activity, potentially due to the simultaneous depletion of other, unidentified splicing factors (Zhan et al. 2022). Consequently, its role in splicing is still poorly defined.

Interestingly, Zhan et al. also identified a potential FAM192A homolog, Fyv6 (function required for yeast viability 6), in *Saccharomyces cerevisiae* (hereafter, yeast) despite having <20% sequence identity (**Fig. 4.1A**). (It should be noted, however, that this level of sequence identity is similar to that between yeast and human homologs of the other second-step factors Slu7 and Prp18.) The predicted AlphaFold structure of Fyv6 (Jumper et al. 2021) was able to be modeled as three α-helices into unassigned EM density from yeast C* spliceosome complexes previously labeled as unknown protein X (**Fig. 4.1B**; α1–3 labeled as in Zhan et al. 2022). Prior to this work, Fyv6 had been detected by mass spectrometry analysis of purified B^{act} and C complex spliceosomes (Fabrizio et al. 2009; Warkocki et al. 2009) as well as postulated to be responsible for unassigned density in a cryo-EM structure of a yeast P complex spliceosome (referred to as UNK in that structure, **Fig. 4.1C**; Liu et al. 2017). In both the C* and P complex spliceosomes, the unassigned EM density is located in a position that could significantly impact splicing chemistry: in C* the density contacts core splicing factors Cef1, Syf1, and Prp8, while in P complex it contacts these factors in addition to the lariat intron, Slu7, and Prp22 (**Fig. 4.1**; **Fig. 4.2**). The unassigned densities in the yeast cryo-EM structures occupy positions on the spliceosome that are very similar to the position of FAM192A in the human pre-C* structures, which contacts the human homologs of these splicing

factors. Together, the combined cryo-EM and mass spectrometry data hint at Fyv6 functioning during splicing; however, no genetic or biochemical evidence for this has been reported.

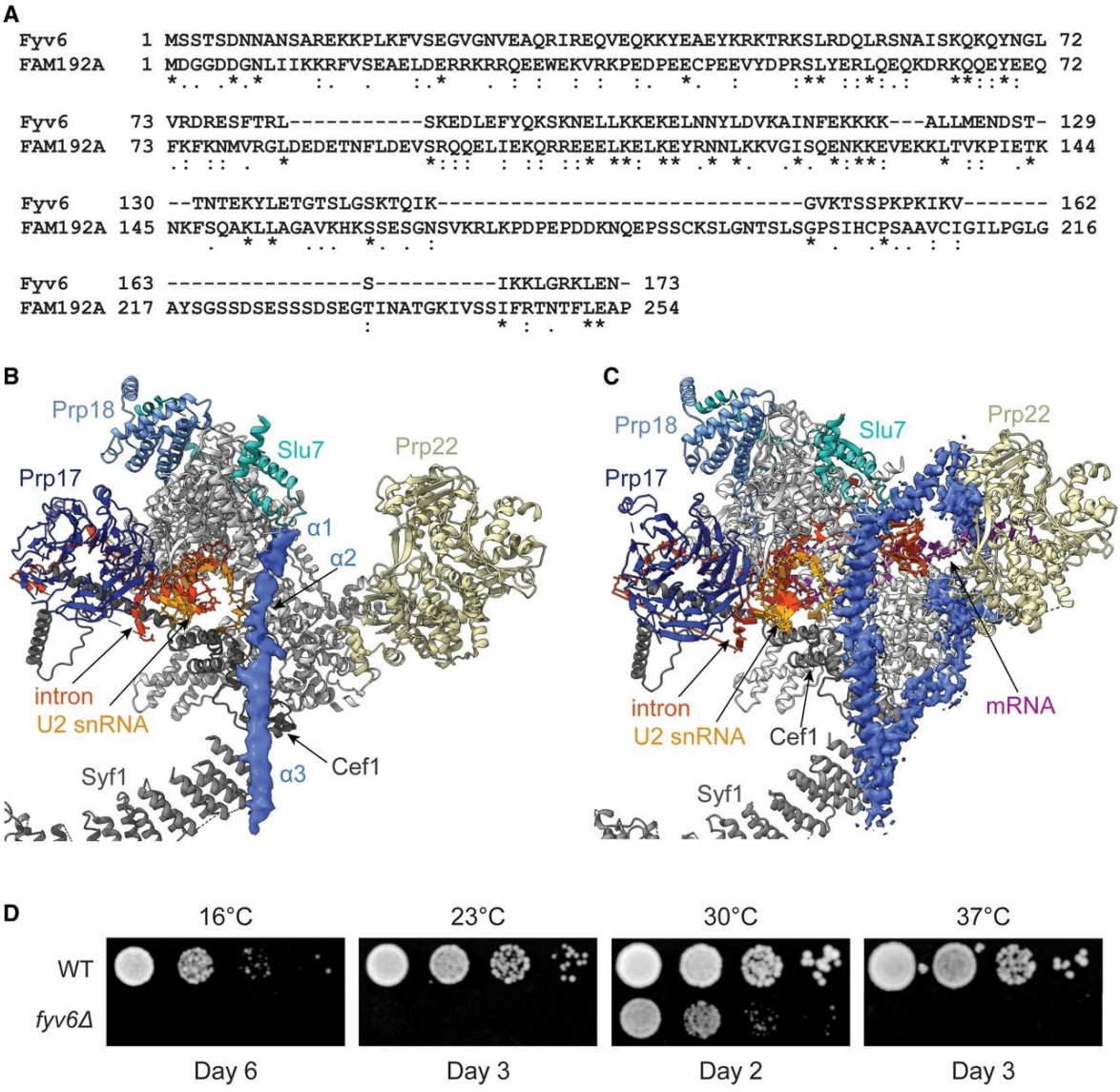

Figure 4.1 Sequence alignment of Fyv6 with FAM192 and unassigned EM density in yeast spliceosome structures. (A) Sequence-based alignment of *S. cerevisiae* Fyv6 and human FAM192A using EMBOSS Needle (Needleman and Wunsch 1970). **(B,C)** Superposition of the atomic models for the spliceosome C* (panel B, 5MQ0) and P (panel C, 6BK8) complexes with the unassigned EM density shown in blue spacefill. The three putative Fyv6 α helices identified by Zhan and coworkers are annotated next to the corresponding EM density in panel B. Images were prepared using ChimeraX (Pettersen et al. 2021). **(D)** Impact of *fyv6Δ* on yeast growth at various temperatures. Plates were imaged on the days shown.

Fyv6 is a poorly studied and nonessential yeast protein originally identified in a screen for mutants sensitive to K1 killer toxin (Pagé et al. 2003). Since identification, Fyv6 has appeared in genetic screens for mutants with sensitivity to heat (Auesukaree et al. 2009), calcineurin inhibitor FK506 (Viladevall et al. 2004), and changes to cell size (Maitra et al. 2019). Fyv6 is localized to the nucleus and has previously been proposed to play a role in nonhomologous end joining, but little is known about its function or interacting partners (Wilson 2002; Huh et al. 2003). Here, we studied the function of Fyv6 during splicing by probing genetic interactions between Fyv6 and the splicing factors Prp8, Prp16, and Prp22. In addition, we assayed splicing in vivo and in vitro by deleting *FYV6* and showed that its loss inhibits exon ligation. Finally, we used both a reporter gene and an endogenous transcript to examine how the loss of *FYV6* impacts 3′ SS selection. Together, these data are consistent with Fyv6 functioning as a second-step splicing factor in yeast.

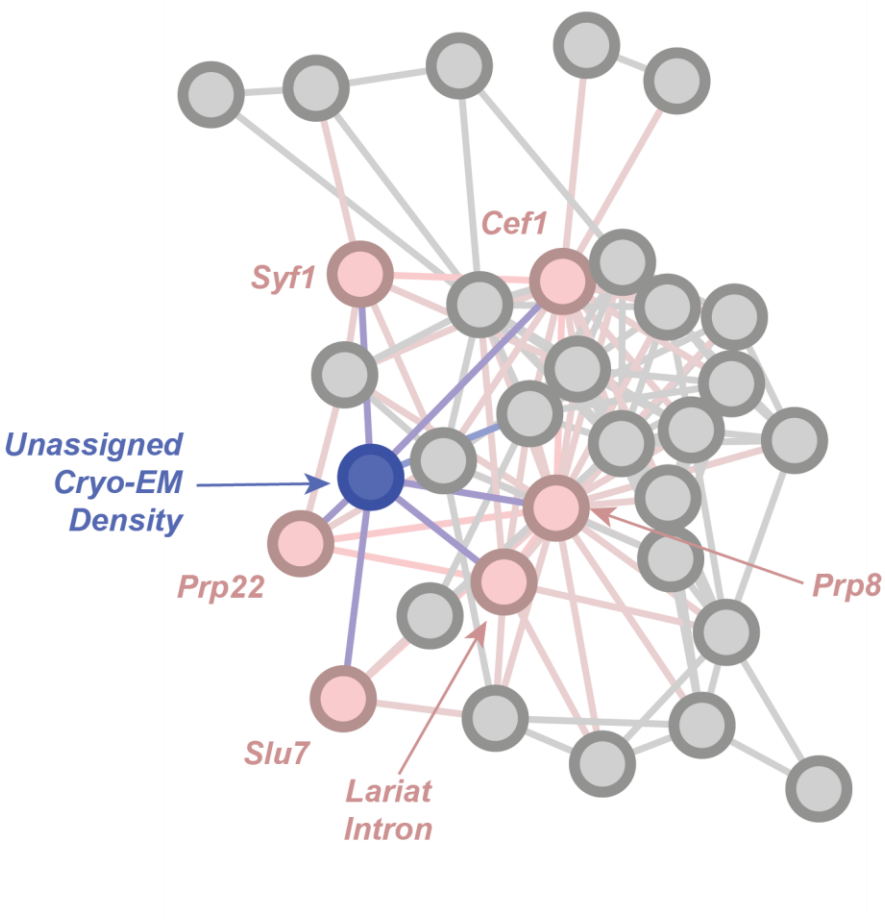

Figure 4.2 Network Analysis of the Spliceosome P Complex. Shown are direct contacts between the unassigned EM density (blue) and core splicing factors and the excised lariat-intron (pink). The unassigned density is not peripherally connected to the network but is, instead, well-integrated with the splicing machinery. Network derived from 6BK8 as described in the Methods section.

4.3 Results

Genetic interactions between Fyv6 and Prp8 first or second-step alleles

To examine a potential role for Fyv6 in splicing, we deleted *FYV6* from the yeast genome, confirmed the deletion by PCR (**Fig. 4.3**), and assayed for genetic interactions with known splicing factors. While *FYV6* is nonessential for yeast viability, its deletion does cause a slow growth defect at 30°C and both cold and temperature sensitivity (*cs* and *ts*) phenotypes at other temperatures (**Fig. 4.1D**). We first tested for genetic interactions with the essential spliceosome component Prp8. Prp8 is a central protein in the spliceosome that scaffolds the active site RNAs and can impact equilibria between the intron branching and exon ligation reactions through structural rearrangement (Query and Konarska 2004; Fica and Nagai 2017). As such, multiple alleles of Prp8 stabilize the spliceosome in either the first- or second-step conformation at the expense of the other state (**Fig. 4.4A**; Umen and Guthrie 1995a; Schneider et al. 2004; Query and Konarska 2004; Liu et al. 2007b). Moreover, alleles of second-step factors Prp18 and Slu7 (*prp18-1*, *slu7-1*, *slu7-ccss*) are synthetically lethal with a first-step allele of Prp8 (*prp8-101* or Prp8^{E1960K}) (Umen and Guthrie 1995b), presumably since both alleles work in concert to promote the first step or inhibit proper progression to the second step. Since Fyv6, like Slu7 and Prp18, is predicted to interact with Prp8 (**Fig. 4.2**), we expected that genetic interactions should occur between Fyv6 and Prp8 if the former is also involved in splicing.

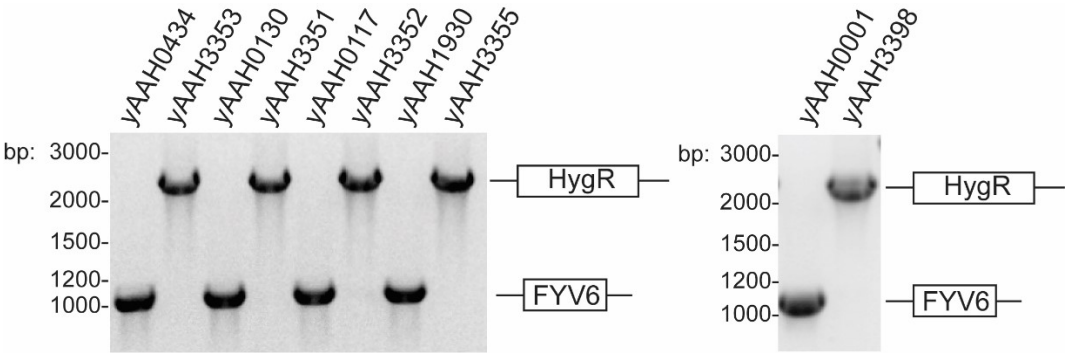

Figure 4.3 Evidence for genomic deletion of *FYV6*. Agarose gel showing products of PCR from genomic DNA isolated from the indicated strains using primers flanking the *FYV6* genomic locus (see Methods). Expected sizes of PCR products are 1036 bp when *FYV6* is present and 2190 bp when *FYV6* has been replaced by a hygromycin resistance cassette (HygR).

Plasmid shuffle of a gene expressing Prp8^{E1960K} into a *fyv6Δ* strain resulted in synthetic lethality even at the normally permissive temperature of 30°C (**Fig. 4.4B**). Synthetic lethality was also observed for another first-step allele of Prp8, *prp8-R1753K* (**Fig. 4.4B**). In contrast, when we shuffled in a second-step allele, *prp8-161* (Prp8^{P986T}), we observed suppression of the growth defect caused by *fyv6Δ* at both

30°C and 37°C (**Fig. 4.4B,C**). When the P986T and R1753K mutations were combined, synthetic lethality was still observed with *fyv6Δ* (**Fig. 4.4B**). These results are consistent with Fyv6 acting to promote the second step of splicing and its deletion, *fyv6Δ*, promoting the first step. Combining *fyv6Δ* with a first-step Prp8 allele can be synthetically lethal due to failure to properly transition to the second step, while combining *fyv6Δ* with a second-step allele may improve yeast growth by facilitating proper first-step/second-step equilibrium.

Genetic interactions between Fyv6 and Prp16 or Prp22

Prp8 first and second-step alleles, as well as first and second-step alleles in U6 snRNA and Cef1, can also genetically interact with mutants of the Prp16 or Prp22 ATPases that promote the first to second-step transition or exit out of the second step by product release, respectively (**Fig. 4.4A**; Query and Konarska 2006, 2012; Eysmont et al. 2019). Based on these observations, we next tested genetic interactions between *fyv6Δ* and Prp16 or Prp22 mutants that may slow these conformational changes. Prp16^{R686I} likely slows the first to second-step transition, leads to a *cs* phenotype at 16°C (Hotz and Schwer 1998), and is synthetically lethal with first-step Prp8 alleles (Query and Konarska 2006). A yeast strain with *fyv6Δ* plus Prp16^{R686I} exacerbates the cold sensitivity, resulting in almost no growth at 16°C and reduced growth at 23°C compared to strains with either allele alone (**Fig. 4.4D**). The combination of Prp16^{R686I} with *fyv6Δ* also results in reduced growth at 30°C and is synthetic lethal at 37°C. Both *fyv6Δ* and first-step Prp8 alleles interact negatively with the Prp16^{R686I} ATPase.

The Prp22^{T637A} mutant uncouples ATP hydrolysis from RNA unwinding (Schwer and Meszaros 2000), likely perturbing the transition out of the second-step conformation and product release. Prp22^{T637A} is also a *cs* allele and does not grow at 16°C or 23°C (**Fig. 4.4E**; Schwer and Meszaros 2000) and is synthetically lethal with second-step alleles of Cef1 (Query and Konarska 2012). When Prp22^{T637A} and *fyv6Δ* were combined, we did not observe any suppression of the *cs* phenotype of Prp22^{T637A} or the *cs*/*ts* phenotype of *fyv6Δ* (**Fig. 4.4E**). Prp22^{T637A}/*fyv6Δ* yeast were viable at 30°C but grew more slowly than strains containing only a single allele. The genetic interactions with Prp22^{T637A} are difficult to interpret as one could expect destabilization of the second-step conformation by *fyv6Δ* to suppress defects in mRNA release from Prp22^{T637A}. Alternatively, slowed product release by Prp22^{T637A} could favor a longer lifetime of the exon ligation conformation and suppression of defects caused by *fyv6Δ*. The genetic and potential physical interactions between Fyv6 and Prp22 need further study to disentangle how the proteins are influencing these steps collaboratively or not.

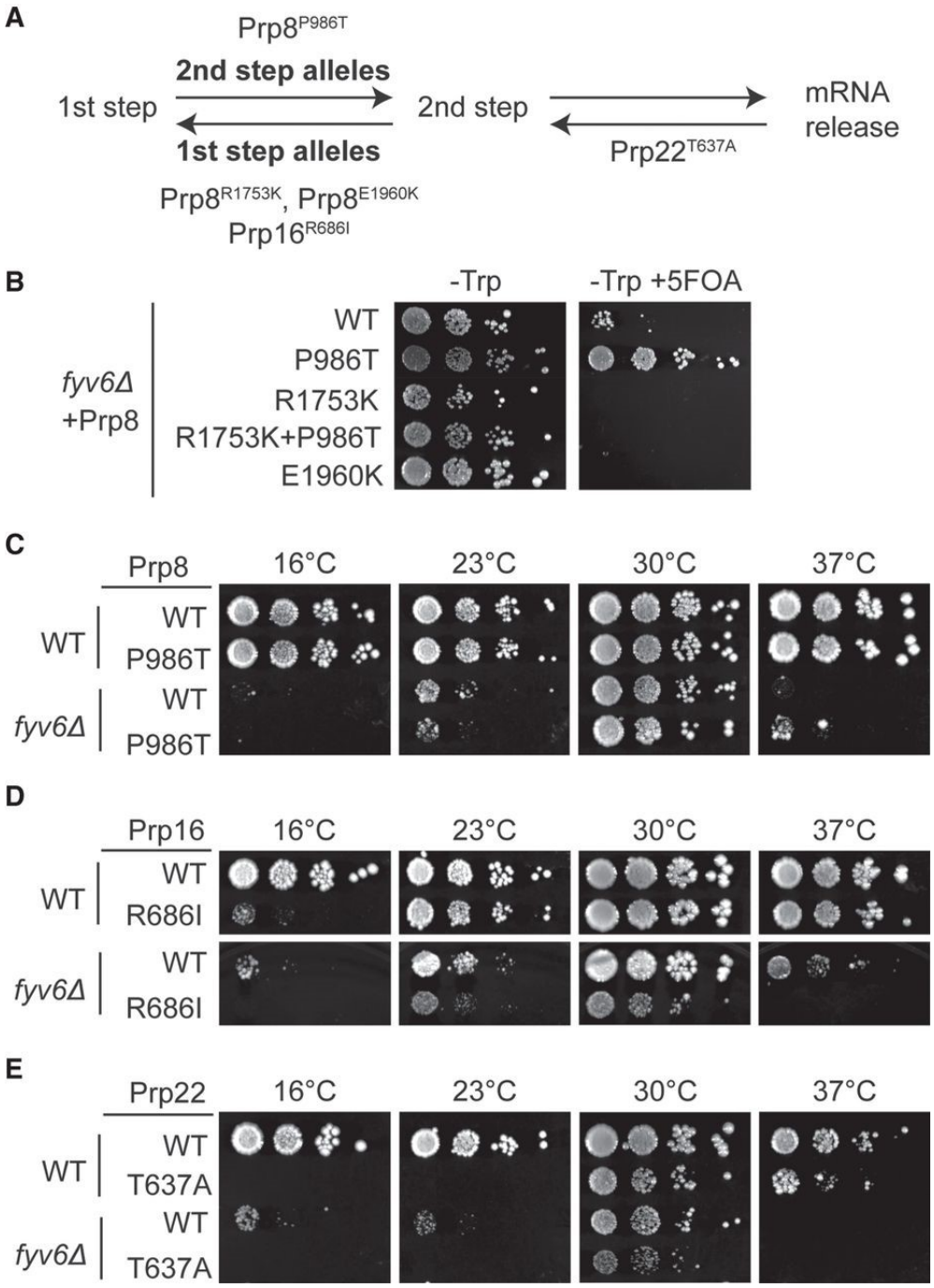

Figure 4.4 Genetic interactions between Fyv6 and Prp8, Prp16, or Prp22. (A) Diagram of how Prp8, Prp16, and Prp22 alleles impact the first and second steps of splicing. **(B)** Alleles of Prp8 were combined with *fyv6Δ* in Prp8 shuffle strains and grown on −Trp or −Trp + 5-FOA plates. Yeast growth was imaged after 3 d at 30°C. **(C)** Prp8^{P986T}/fyv6Δ strains were tested for suppression or exacerbation of temperature-dependent growth phenotypes. **(D,E)** Alleles of Prp16 and Prp22 were combined with *fyv6Δ* and tested for suppression or exacerbation of temperature-dependent growth phenotypes. For panels C–E, yeast were plated on YPD and imaged after 3 (30°C), 4 (23°C and 37°C), or 10 (16°C) days.

Impact of *fyv6Δ* on yeast growth using the ACT1–CUP1 splicing reporter assay

If Fyv6 is a component of the splicing machinery as the genetic interactions suggest, we would also predict changes in in vivo splicing in the absence of the protein. To test this, we used the ACT1–CUP1 reporter assay in which changes in the splicing of the reporter pre-mRNA (**Fig. 4.5A**) confer proportional changes in the copper tolerance of a sensitized yeast strain with increased splicing efficiency leading to growth at higher copper concentrations (Lesser and Guthrie 1993). Since yeast lacking Fvy6 grow more slowly than WT even under optimal growth conditions (**Fig. 4.1D**, for example) we scored WT yeast growth on Cu^{2+}-containing plates after 48 h but *fyv6Δ* yeast were scored after 72 h. Consistent with the slow growth and with a function of Fyv6 during splicing, we observed slightly reduced copper tolerance for even the WT ACT1–CUP1 reporter in the *fyv6Δ* strain (**Fig. 4.5B,C**).

We also tested several ACT1–CUP1 reporters containing substitutions in the 5′ (U2A, A3C, and A3U) or 3′ SS (UuG and gAG) or the BS (A258U, BSG, and BSC) (**Fig. 4.5A**) to determine if loss of Fyv6 is especially detrimental or beneficial to introns with nonconsensus sequences. This set of reporters includes those that are limiting for the first step (A258U), those that accumulate lariat intermediates due to being limiting for the second step (U2A, A3C, BSG, UuG, and gAG), and those that are limiting for both steps of splicing (BSC) (Lesser and Guthrie 1993; Liu et al. 2007b). Previous work has shown that first-step alleles of Prp8 result in reduced copper tolerance and fewer ligated mRNA products with the U2A, A3C, BSC, BSG, UuG, and gAG reporters and that second-step alleles of Prp8 or Cef1 improve copper tolerance of yeast with the U2A, BSC, BSG, UuG, and gAG reporters (Liu et al. 2007b; Query and Konarska 2012). We would predict, therefore, that if *fyv6Δ* is a first-step allele we should see reduced copper tolerance of reporters limiting for the second step of splicing, similar to those in Prp8.

When we tested this prediction, we found that copper tolerance was similar between WT and *fyv6Δ* yeast containing reporters with the A3C, UuG, A258U, and BSC substitutions. However, strains containing A3U, BSG, and gAG reporters exhibited less tolerance to copper than WT (**Fig. 4.5B,C**), suggesting poorer splicing of pre-mRNAs with these substitutions. We conclude that loss of Fyv6 results in a subset of changes in copper tolerance akin to first-step alleles of Prp8 and is consistent with Fyv6 supporting the second step when present.

While the strongest effects in the ACT1–CUP1 assay support the involvement of Fyv6 in the second step, it should be noted that we also observed a slight increase in copper tolerance in the presence of U2A (**Fig. 4.5B,C**). The U2A reporter is a 5′ SS mutant but is not limiting for the first step as it can readily accumulate lariat intermediates that can be discarded from the spliceosome (Liu et al. 2007b). This suggests that it is defective for the transition to the second step or the second step itself. As described in the preceding paragraph, second-step alleles of Prp8 also increase the copper tolerance of

yeast with the U2A reporter (Liu et al. 2007b). This could mean that *fyv6Δ* can, at least in this case of U2A, provide some support for the second step as well.

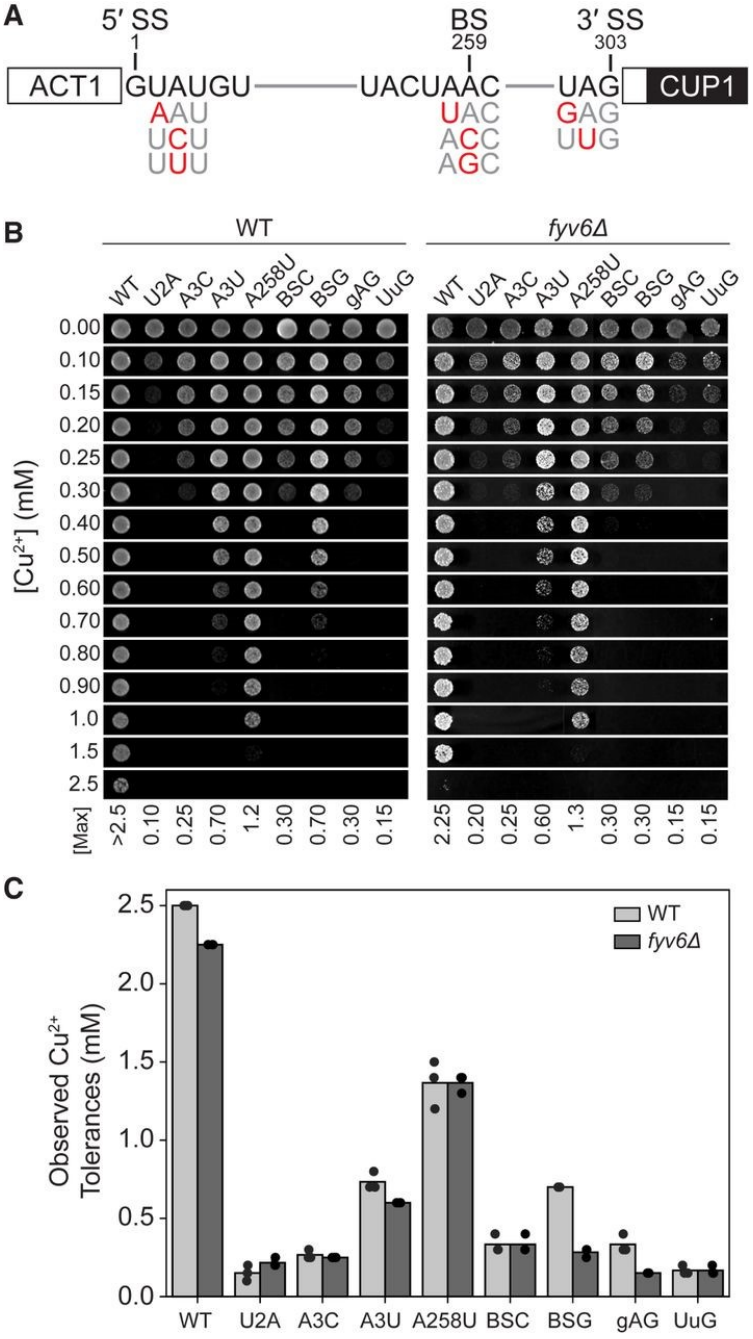

Figure 4.5 Impact of *FYV6* deletion on yeast copper tolerance using the ACT1–CUP1 assay. (A) Schematic of the WT ACT1–CUP1 reporter along with intronic substitutions. **(B)** Images of representative yeast growth on copper-containing media shown after 48 (WT) or 72 h (*fyv6Δ*) for strains containing the indicated ACT1–CUP1 reporters. **(C)** Maximum copper tolerances observed for each strain for N = 3 replicates (dots). Bars represent the average values.

First-step products accumulate in the absence of Fyv6 in vitro

We next tested if Fyv6 plays a role in exon ligation using in vitro splicing assays with whole-cell extracts (WCEs) made from either WT or *fyv6Δ* yeast. Splicing extracts made from the same yeast strains used in the ACT1–CUP1 assays show accumulation of first-step products and decreased second-step efficiency in the absence of Fyv6 (**Fig. 4.6**). Interestingly, this effect was less pronounced when the BJ2168 protease-deficient yeast strain commonly used to prepare WCEs was used (**Fig. 4.7**). This strain also displayed less cold sensitivity at 23°C (close to the temperature at which in vitro splicing assays are conducted) than the *cup1Δ* strains when *FYV6* was deleted (**Fig. 4.7**). This suggests that the strength of some *fyv6Δ* phenotypes may be strain dependent.

Loss of Fyv6 changes 3′ SS selection

Since both Prp18 and Slu7 can change 3′ SS selection (Frank and Guthrie 1992; Kawashima et al. 2009, 2014) and Slu7 contacts the unassigned density attributed to Fyv6 (**Fig. 4.1B,C**), we tested whether or not loss of Fyv6 can also change splicing outcomes. We utilized an ACT1–CUP1 reporter containing an additional, alternative 3′ SS proximal to the BS which results in a frameshift when used instead of the distal 3′ SS (**Fig. 4.8A**). Previous studies with this reporter showed that use of the proximal 3′ SS greatly increases in the presence of the *slu7-1* allele with an ∼20-fold change in the ratio of mRNAs produced using the proximal versus distal sites (Frank and Guthrie 1992). Indeed, when this reporter was used, we observed the largest differences in copper tolerance (**Fig. 4.8B**). To confirm that this change in survival was due to a change in 3′ SS usage and use of the proximal SS, we isolated RNAs from the yeast strains and quantified SS usage by primer extension. These results showed an increase in the use of the proximal SS and an approximately fivefold increase in the ratio of mRNAs produced using the proximal versus distal sites (**Fig. 4.8C,D**). Moreover, these data suggest that Fyv6, like Slu7, helps to enforce a preference for BS distal 3′ SS. However, one limitation of our studies is that we have not mapped the BS used in the absence of Fyv6 and cannot completely exclude the possibility of an upstream, cryptic BS in the ACT1 intron (Kao et al. 2021) driving usage of the proximal 3′ SS. Nonetheless, we believe that this is less likely than a change in the second-step reaction itself given our in vitro data (**Fig. 4.6**).

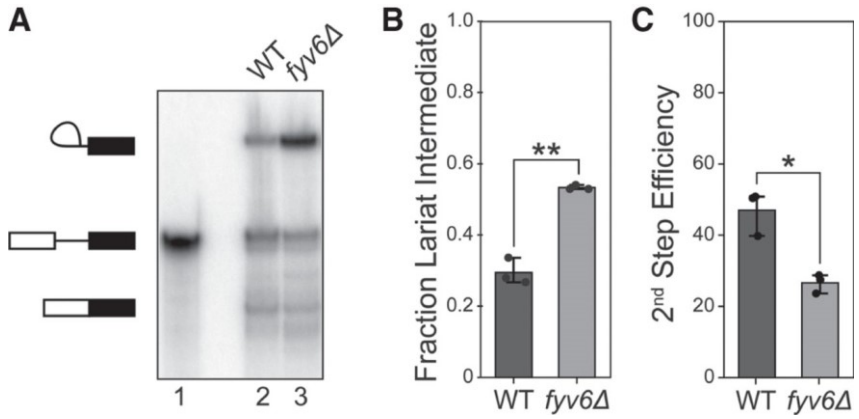

Figure 4.6 Accumulation of splicing intermediates occurs in the absence of Fyv6 in in vitro splicing assays. (A) Products of the first and second steps after incubation of a radioactively labeled RP51A pre-mRNA (lane 1) in WCE from *cup1Δ* strains (yAAH0434 and yAAH3353, **Table 4.1**) for 45 min (lanes 2,3). Quantitation of **(B)** lariat intermediates or **(C)** second-step splicing efficiency from three technical replicates. Statistical significance was determined by unpaired Welch's two-tailed t-test (P = 0.006415 and 0.01788 for fraction of lariat intermediate and second-step efficiency, respectively). The fraction of the lariat intermediate represents the fraction of that species relative to the substrate and all splicing products while the second-step efficiency represents the fraction of mRNA relative to the sum of first- and second-step products (lariat intermediate plus mRNA). Details for calculations can be found in the Materials and Methods.

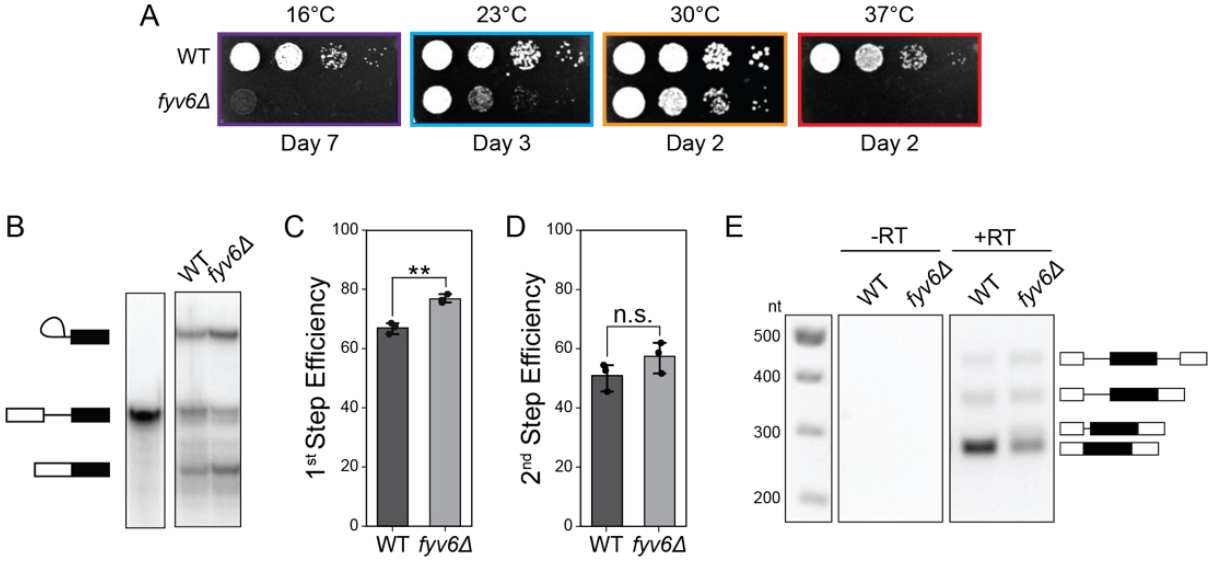

Figure 4.7 Consequences of Fyv6 deletion in the yeast strain BJ2168 background. A) Impact of *fyv6Δ* on yeast growth at various temperatures from BJ2168 strains yAAH0001 (WT) and yAAH3398 (*fyv6Δ*). Plates were imaged on the days shown. **B)** Products of the first and second steps of splicing after incubation of a radioactively labeled RP51A pre-mRNA (lane 1) in WCE from the BJ2168 strains for 45 min. Quantitation of the splicing products from three technical replicates that completed the **C)** 1st step or **D)** 2nd step of splicing. **E)** Representative gel image of RT-PCR of SUS1 in strains with (WT) and without (*fyv6Δ*) Fyv6 in a BJ2168 background. (+RT reactions contain reverse transcriptase; -RT control reactions do not contain reverse transcriptase). Statistical significance was determined by a student's two-tailed t-test (p = 0.005048). Details for how the 1st and 2nd step efficiencies were calculated can be found in the Methods.

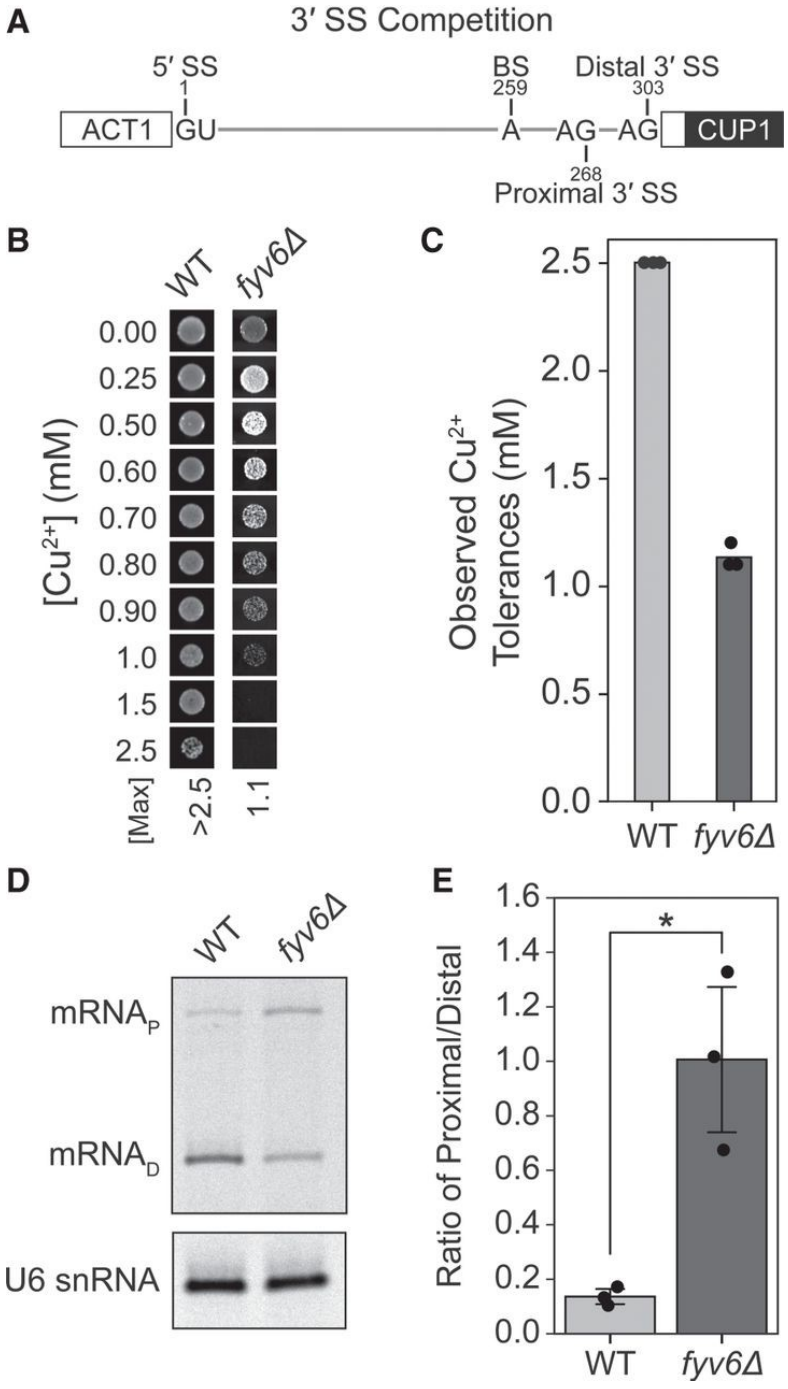

Figure 4.8 Loss of Fyv6 changes 3′ SS selection in yeast. (A) Schematic of the 3′ SS competition reporter (3′ SS comp) showing relative locations of the proximal and distal sites. **(B)** Images of representative yeast growth on copper-containing media shown after 48 (WT) or 72 h (*fyv6Δ*) for strains containing the 3′ SS comp ACT1–CUP1 reporter. **(C)** Maximum copper tolerances observed for each strain for N = 3 replicates (dots). Bars represent the average values. **(D)** Representative primer extension analysis of mRNAs generated by yeast using the distal ($mRNA_D$) or proximal ($mRNA_P$) 3′ SS in the presence (WT) or absence of Fyv6 (*fyv6Δ*). U6 snRNA was analyzed as a loading control. **(E)** Quantification of the primer extension results from N = 3 replicates (dots) expressed as a ratio of $mRNA_P/mRNA_D$. Bars represent the average of the replicates ±SD. Means between the two experimental groups were compared with an unpaired Welch's two-tailed t-test (P = 0.04262).

Finally, we examined if the loss of Fyv6 impacts the splicing of endogenous yeast transcripts in vivo. Given its sensitivity to perturbations in the splicing machinery (Hossain et al. 2009; Cuenca-Bono et al. 2011; Hossain et al. 2011), we examined *SUS1* splicing—one of the rare genes in yeast with two introns. We performed RT-PCR on RNA extracted from strains that either contained (WT) or lacked *FYV6* and additionally were *upf1Δ* to prevent nonsense-mediated decay (NMD) of unspliced or alternatively spliced transcripts (Sayani et al. 2008). In the case of the WT *upf1Δ* yeast, we observed products for unspliced, partially spliced (one of two introns removed), and fully spliced mRNA. The *fyv6Δ* strain has these same products present in the WT strain as well as a product that is slightly larger than the fully spliced mRNA (**Fig. 4.9A**). This larger product was also seen upon RT-PCR of *SUS1* in the BJ2168 *fyv6Δ* strain (**Fig. 4.7E**). The appearance of the new band was the biggest change in splicing of *SUS1* upon *FYV6* deletion (**Fig. 4.9B**). We sequenced this product and determined that it results from the use of an alternative 3′ SS in the first intron of *SUS1* that, to our knowledge, has not previously been reported (**Fig. 4.9C,D**). In this isoform, the last twenty nucleotides of intron 1 are retained, and an intronic CAG located proximal to the BS is used as the alternate 3′ SS. This new splice site is only 8 nt downstream from the BS adenosine (**Fig. 4.9D**). It is possible that this transcript is normally degraded by NMD since it includes a premature termination codon 35 amino acids into the coding sequence, and the corresponding isoform was less prevalent by RT-PCR from the Upf1-containing BJ2168 strain (**Fig. 4.7E**).

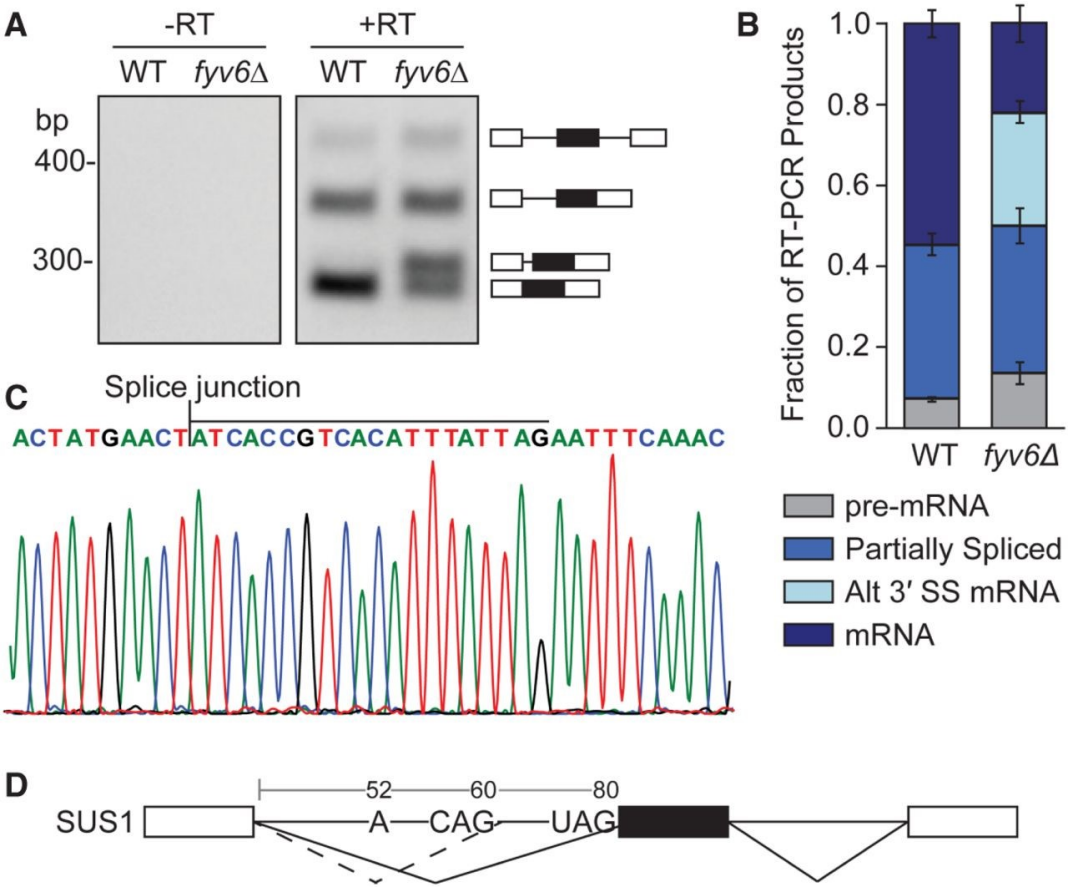

Figure 4.9 Loss of Fyv6 results in the use of an alternative 3′ SS in SUS1. (A) Representative gel image of RT-PCR of SUS1 in strains with (WT) and without (*fyv6Δ*) Fyv6 in a *upf1Δ* background. (+RT reactions contain reverse transcriptase; −RT control reactions do not contain reverse transcriptase). **(B)** Quantification of band intensities of each isoform as a fraction of the total SUS1 product in a lane. The bars indicate the average of three experiments with standard deviation. **(C)** Portion of the Sanger sequencing chromatogram of the SUS1 splice variant identified as an RT-PCR product in the *fyv6Δ* strain. The bar above the nucleotides indicates those from intron 1 included in the splice variant due to the use of an alternative 3′ SS. **(D)** Diagram of the SUS1 gene structure with the BS adenosine and the two, alternative 3′ SS of intron 1 indicated. The numbering of the nucleotides begins at the first nucleotide of intron 1. The newly identified 3′ SS is at position 60.

4.4 Discussion

Together, the genetic and biochemical data presented here as well as the mass spectrometry and cryo-EM work of others, indicate that yeast Fyv6 is a second-step splicing factor and likely is a component of the spliceosome. Our results do not, however, confirm that the unassigned EM density present in yeast C* and P complex spliceosome cryo-EM maps is due to Fyv6. Further work will be needed to verify that this is indeed the case either through obtaining higher resolution cryo-EM data or experimental approaches that probe protein–protein interactions to map the Fyv6 binding site.

The results we obtained for the deletion of *FYV6* are similar to those reported for perturbations of other second-step splicing factors. Different alleles of Fyv6, Slu7, or Prp18 result in synthetic lethality with first-step alleles of Prp8 (Umen and Guthrie 1995b). Like Fyv6, the second-step factors Slu7, Prp18, and Prp22 also affect 3′ SS choice (Frank and Guthrie 1992; Crotti et al. 2007; Kawashima et al. 2014; Semlow et al. 2016). Interestingly, in strains lacking Fyv6, 3′ SS that were within nine nucleotides of the BS were used more frequently in ACT1–CUP1 and *SUS1* (**Figs. 4.8, 4.9**). This resembles observations made with Slu7, which is necessary for the use of any splice sites that are >9 nt downstream from the BS (Brys and Schwer 1996). Structurally, Slu7 is positioned within the spliceosome in a way that suggests that it could guide distal 3′ SS into the correct location (Fica et al. 2017; Wilkinson et al. 2017). In addition, Prp22 is important for splicing when the 3′ SS is >20 nt away from the BS (Schwer and Gross 1998). In the human spliceosome, FAM192A is proximal to Prp22, Slu7, and the intron. This suggests that, as a potential FAM192 homolog and in agreement with the unassigned density in yeast spliceosomes (**Fig. 4.1B,C**), Fyv6 could also impact the second-step spliceosome conformation and/or positioning of the intron either directly through contacts with the intron or via interactions with Slu7 and Prp22.

Many outstanding questions remain about Fyv6 function. We do not know when Fyv6 associates or dissociates from the spliceosome, whether it can function at stages other than exon ligation, or the structural mechanism behind its influence on 3′ SS usage. Finally, since it seems likely that Fyv6 is the functional yeast homolog of human FAM192A in terms of pre-mRNA splicing, it will be worth investigating if other functions of FAM192A/Fyv6 are conserved in yeast. FAM192A also associates with the 20S proteasome via interaction with PA28γ, a 20S proteasome regulator, for ubiquitin-independent protein degradation within the nucleus (Jonik-Nowak et al. 2018). While yeast lack an apparent homolog for PA28γ (Jonik-Nowak et al. 2018), it may be interesting to determine if Fyv6 is involved in protein degradation and if there are any Fyv6-dependent links between proteostasis and pre-mRNA splicing.

4.5 Materials and Methods

Yeast strains and plasmids used in this study are described in Tables 4.1 and 4.2. Yeast transformation, plasmid shuffling/5-FOA selection, and growth were carried out using standard techniques and media (Sikorski and Boeke 1991; Treco and Lundblad 1993).

Network analysis of potential Fyv6 interactions

Protein–protein and protein–RNA interactions found in the atomic model for the P complex spliceosome (6BK8) were identified using LouiseNET, and the resulting nodes and edges were plotted as an undirected network model using GEPHI, as previously described (Bastian et al. 2009; Kaur et al. 2022).

Deletion strain creation

The *FYV6* and *UPF1* genes were deleted through replacement with a hygromycin or kanamycin resistance cassette, respectively, by homologous recombination (see **Table 4.1**; Goldstein and McCusker 1999). Gene deletion was confirmed by genomic DNA extraction from the strains and PCR amplification of the corresponding genomic locus. Primers for *FYV6* were Fyv6-check-fwd 5′-TGGATCGAACACAGGACCTC-3′ and Fyv6-check-rev 5′-GTGGAACGAGCAATCAATGTGATC-3′. Primers for UPF1 were Upf1-check-up 5′-CAGCCAACAAACGTTGAAGATTTCATC-3′ and Upf1-check-down 5′-TTGCAGCGCTCATTTCACGGTTGAGC-3′.

ACT1–CUP1 copper tolerance assays

Yeast strains expressing ACT1–CUP1 reporters were grown to stationary phase in −Leu DO media to maintain selection for plasmids, diluted to OD_{600} = 0.5 in 10% (v/v) glycerol, and spotted onto −Leu DO plates containing 0–2.5 mM $CuSO_4$ (Lesser and Guthrie 1993; Carrocci et al. 2018). Plates were scored and imaged after 48 h of growth at 30°C for WT strains, and after 72 h of growth at 30°C for *fyv6Δ* strains due to differential growth between strains.

Growth assays

For temperature-dependent growth assays, yeast strains were grown overnight to stationary phase in YPD media, diluted to OD_{600} = 0.5 in 10% (v/v) glycerol, and stamped onto YPD plates. The plates were incubated at 16°C, 23°C, 30°C, or 37°C for the number of days indicated in each figure before imaging.

For growth assays in the presence of 5-FOA, yeast strains were grown overnight to stationary phase in −Trp DO media, diluted to OD_{600} = 0.5 in 10% (v/v) glycerol, and stamped onto −Trp and −Trp +5-FOA plates. The plates were incubated at 30°C for 3 d before imaging.

Primer extension

Cell cultures were inoculated from stationary phase saturated cultures grown overnight in −Leu DO media. The cultures were then grown until OD_{600} = 0.6–0.8, and 10 OD_{600} units were collected by centrifugation. Total RNA was isolated from yeast, and contaminating DNA was depleted using the MasterPure Yeast RNA Purification Kit (Epicentre) protocol, with minor changes as previously described (Carrocci et al. 2017). IR700 dye conjugated probes (Integrated DNA Technologies) were used for primer extension of the ACT1–CUP1 reporter (10 pmol yAC6: /5IRD700/GGCACTCATGACCTTC) and U6 snRNA (2 pmol yU6: /5IRD700/GAACTGCTGATCATGTCTG) (Carrocci et al. 2017; van der Feltz et al. 2021). Primer extension products were visualized on a 7% (w/v) denaturing polyacrylamide gel (42 cm × 22 cm × 0.75 mm) run at 35W for 80 min at RT. Gels were imaged with an Amersham Typhoon NIR laser scanner (Cytiva), and band intensities were quantified with ImageJ (v1.53, 2022).

RT-PCR

Cell cultures were inoculated from stationary phase saturated cultures grown overnight in YPD media. The cultures were then grown until OD_{600} = 0.7–0.9, and 1 mL of cells was collected by centrifugation. Total RNA was isolated from yeast, and contaminating DNA was depleted using the MasterPure Yeast RNA Purification Kit (Epicentre) protocol with minor changes as previously described (Carrocci et al. 2017). RT-PCR reactions were set up using the Access RT-PCR System (Promega Corporation) protocol with 75 ng RNA per 25 µL reaction. Primers used were SUS1-exon1 5′-TGGATACTGCGCAATTAAAGAGTC-3′ and SUS1-exon3 5′-TCATTGTGTATCTACAATCTCTTCAAG-3′. Products of the reaction were run on 2% (w/v) agarose-TBE gels and imaged. Band intensities were quantified with ImageJ Software (v1.53, 2022) (Schindelin et al. 2015). In order to identify the sequences of the reaction products, TOPO cloning (TOPO TA Cloning Kit, Thermo Fisher) was used to insert the RT-PCR products into a vector, and the inserts were then Sanger sequenced.

Splicing assays

Splicing WCEs and [^{32}P]-labeled RP51A substrate pre-mRNAs were prepared as previously described (Crawford et al. 2008). Splicing assays were conducted at room temperature using 40% (v/v) WCE and 0.2 nM pre-mRNA substrate (Crawford et al. 2008). [^{32}P]-labeled RNAs were then isolated and separated on a 12% (w/v) denaturing PAGE gel followed by phosphor imaging after exposing gels overnight to the phosphor imaging screen. Data were analyzed using ImageJ. The fractions of the lariat intermediate and efficiencies of the second step were calculated using Equations 1 and 2 and the corresponding band intensities as previously described (Mayerle and Guthrie 2016).

$$Fraction\ Lariat\ Intermediate = \frac{\text{lariat intermediate}}{\text{pre}-\text{mRNA}+\text{lariat interemdiate}+\text{mRNA}} \quad (1)$$

$$Second\ Step\ Efficiency = \frac{\text{mRNA}}{\text{lariat interemdiate}+\text{mRNA}} \quad (2)$$

Table 4.1 Yeast strains

Name	Genotype	Description	Source or reference
46a (yAAH0434)	MATalpha cup1D::ura3 ura3-52 his3-D200 trp1-D63 lys2-801 ade2-1 leu2-D1	Cu^{2+} sensitive strain	Lesser and Guthrie, 1993
yAAH0520	yAAH0434+pAAH0470	Cu^{2+} sensitive strain with WT ACT1-CUP1	Lesser and Guthrie, 1993.
yAAH3343	yAAH0434+pAAH0513	Cu^{2+} sensitive strain with U2A ACT1-CUP1	This study
yAAH3344	yAAH0434+pAAH1032	Cu^{2+} sensitive strain with A3C ACT1-CUP1	This study
yAAH3345	yAAH0434+pAAH0514	Cu^{2+} sensitive strain with A3U ACT1-CUP1	This study
yAAH3346	yAAH0434+pAAH0524	Cu^{2+} sensitive strain with A258U ACT1-CUP1	This study
yAAH0519	yAAH0434+pAAH0441	Cu^{2+} sensitive strain with BSC (A259C) ACT1-CUP1	Lesser and Guthrie, 1993.
yAAH3347	yAAH0434+pAAH0880	Cu^{2+} sensitive strain with BSG (A259G) ACT1-CUP1	This study
yAAH3348	yAAH0434+pAAH0526	Cu^{2+} sensitive strain with gAG (U301G) ACT1-CUP1	This study
yAAH3349	yAAH0434+pAAH0527	Cu^{2+} sensitive strain with UuG (A302U) ACT1-CUP1	This study
yAAH3350	yAAH0434+pAAH0734	Cu^{2+} sensitive strain with 3′ss competition ACT1-CUP1	This study
yAAH3353	yAAH0434+*fyv6Δ*::hphMX	Cu^{2+} sensitive *fyv6Δ* strain	This study

yAAH3356	yAAH3353+pAAH0470	Cu^{2+} sensitive *fyv6Δ* strain with WT ACT1-CUP1	This study
yAAH3357	yAAH3353+pAAH0513	Cu^{2+} sensitive *fyv6Δ* strain with U2A ACT1-CUP1	This study
yAAH3358	yAAH3353+pAAH1032	Cu^{2+} sensitive *fyv6Δ* strain with A3C ACT1-CUP1	This study
yAAH3359	yAAH3353+pAAH0514	Cu^{2+} sensitive *fyv6Δ* strain with A3U ACT1-CUP1	This study
yAAH3360	yAAH3353+pAAH0524	Cu^{2+} sensitive *fyv6Δ* strain with A258U ACT1-CUP1	This study
yAAH3361	yAAH3353+pAAH0441	Cu^{2+} sensitive *fyv6Δ* strain with BSC ACT1-CUP1	This study
yAAH3362	yAAH3353+pAAH0880	Cu^{2+} sensitive *fyv6Δ* strain with BSG ACT1-CUP1	This study
yAAH3363	yAAH3353+pAAH0526	Cu^{2+} sensitive *fyv6Δ* strain with gAG ACT1-CUP1	This study
yAAH3364	yAAH3353+pAAH0527	Cu^{2+} sensitive *fyv6Δ* strain with UuG ACT1-CUP1	This study
yAAH3365	yAAH3353+pAAH0734	Cu^{2+} sensitive *fyv6Δ* strain with 3′ss competition ACT1-CUP1	This study
yAAH0130	MATa ltrp1 ura3 his lys2 leu2 ade2 *prp16Δ*::lys2, pSB2-Prp16 (Prp16 URA)	shuffle strain for Prp16, yS78	Wang and Guthrie, 1998.
yAAH1946	yAAH0130 + pAAH1040	$Prp16^{WT}$	van der Feltz et al. 2021.
yAAH1947	yAAH0130 + pAAH1039	$Prp16^{R686I}$	van der Feltz et al. 2021.
yAAH3351	yAAH0130 + fyv6Δ::hphMX	*fyv6Δ* shuffle strain for Prp16	This study
yAAH3371	yAAH3351+pAAH1040	*fyv6Δ*, $Prp16^{WT}$	This study

yAAH3372	yAAH3351+pAAH1039	*fyv6Δ*, Prp16^{R686I}	This study
yAAH1930	MATa ade2 cup1Δ::ura3 his3 leu2 lys2 trp1 ura3 GAL+ prp22Δ::loxP p360-22 (Prp22 URA CEN ARS)	Cu^{2+} sensitive Prp22 shuffle strain	Charles Query/Magda Konarska
yAAH1931	yAAH1930+pAAH1042	Prp22^{WT}	Charles Query/Magda Konarska
yAAH1932	yAAH1930+pAAH1043	Prp22^{T637A}	Charles Query/Magda Konarska
yAAH3355	yAAH1930 + *fyv6Δ*::hphMX	Cu^{2+} sensitive *fyv6Δ* shuffle strain for Prp22	This study
yAAH3379	yAAH3355+pAAH1042	*fyv6Δ*, Prp22^{WT}	This study
yAAH3380	yAAH1043	*fyv6Δ*, Prp22^{T637A}	This study
yAAH0117	ade2 cup1Δ ura3 his3 leu2 lys2 prp8Δ:lys2 trp1 pJU169 (PRP8 URA3)	Cu^{2+} sensitive Prp8 shuffle strain	Christine Guthrie; Umen and Guthrie, 1996.
yAAH2435	yAAH0117+pAAH0997	Prp8^{WT}	van der Feltz et al. 2021.
yAAH2437	yAAH0117+pAAH1001	Prp8^{P986T} (Prp8-161)	van der Feltz et al. 2021.
yAAH3352	yAAH0117 + *fyv6Δ*::hphMX	Cu^{2+} sensitive *fyv6Δ* shuffle strain for Prp8	This study
yAAH3368	yAAH3352+pAAH0997	Prp8^{WT}/URA plus Prp8^{WT} merodiploid	This study
yAAH3369	yAAH3352+pAAH1001	Prp8^{WT}/URA plus Prp8^{P986T} merodiploid	This study
yAAH3370	yAAH3352+pAAH1006	Prp8^{WT}/URA plus Prp8$^{R1753K+P986T}$ merodiploid	This study

yAAH3375	yAAH3352+pAAH1004	$Prp8^{WT}$/URA plus $Prp8^{R1753K}$ merodiploid	This study
yAAH3376	yAAH3352+pAAH1003	$Prp8^{WT}$/URA plus Prp8 8-101 merodiploid	This study
yAAH3373	yAAH3352+pAAH0997 (selected from yAAH3368)	$Prp8^{WT}$	This study
yAAH3374	yAAH3352+pAAH1001 (selected from yAAH3369)	$Prp8^{P986T}$	This study
yAAH0001	MATa prc1–407 prb1–1122 pep4–3 leu2 trp1 ura3–52 gal2	BJ2168, Protease-deficient strain	Hoskins et al. 2011.
yAAH3398	yAAH0001+ *fyv6Δ*::hphMX	*fyv6Δ* strain	This study
yAAH3399	yAAH0434+*upf1Δ*::Kan MX	Cu^{2+} sensitive *upf1Δ* strain	This study
yAAH3400	yAAH3353+*upf1Δ*::Kan MX	Cu^{2+} sensitive *fyv6Δ upf1Δ* strain	This study

Table 4.2 Plasmids

Plasmid ID	Plasmid name	Description	Source or reference
pAAH0060	pAG32-hphMX4	Contains hphMX4 cassette used to replace *FYV6* in deletion strains	Goldstein and McCusker, 1999.
pAAH1040	$Prp16^{WT}$	Plasmid used to generate $Prp16^{WT}$ strains (PRP16 TRP1 CEN ARS).	Gift from Charles Query.
pAAH1039	$Prp16^{R686I}$	Plasmid used to generate $Prp16^{R686I}$ strains (PRP16-R686I TRP1 CEN ARS).	Gift from Charles Query.
pAAH1042	$Prp22^{WT}$	Plasmid used to generate $Prp22^{WT}$ strains (PRP22 TRP1 CEN ARS).	Gift from Charles Query.
pAAH1043	$Prp22^{T637A}$	Plasmid used to generate $Prp22^{T637A}$ strains (PRP22-T637A TRP1 CEN ARS).	Gift from Charles Query.
pAAH0997	$Prp8^{WT}$	Plasmid used to generate $Prp8^{WT}$ strains (PRP8 2micron TRP1), pJU225-4	Gift from Dave Brow.
pAAH1003	Prp8 8-101	Plasmid used to generate Prp8 8-101 strains (Prp8 8-101 2micron TRP1), pMK8-20	Gift from Magda Konarska; Liu et al. 2007b.
pAAH1004	$Prp8^{R1753K}$	Plasmid used to generate $Prp8^{R1753K}$ strains (Prp8-R1753K 2micron TRP1), pMK8-14	Gift from Magda Konarska. Liu et al. 2007b.
pAAH1001	$Prp8^{P986T}$	Plasmid used to generate $Prp8^{P986T}$ strains (Prp8-161 2micron TRP1) prp8-161 allele, pMK8-18	Gift from Magda Konarska. Liu et al. 2007b.
pAAH1006	$Prp8^{R1753K/P986T}$	Plasmid used to generate $Prp8^{R1753K/P986T}$ strains (PRP8-R1753K/P986T 2micron TRP1). pMK8T-229	Gift from Magda Konarska. Liu et al. 2007b.

pAAH0470	ACT1-CUP1 WT	WT reporter used for ACT1-CUP1 assays. ACT1-CUP1 expressed from a GPD promoter on a Leu2-marked plasmid.	Gift from Charles Query.
pAAH0513	ACT1-CUP1 U2A	5′ss U2A reporter used for ACT1-CUP1 assays. ACT1-CUP1 expressed from a GPD promoter on a Leu2-marked plasmid.	Gift from Charles Query.
pAAH1032	ACT1-CUP1 A3C	5′ss A3C reporter used for ACT1-CUP1 assays. ACT1-CUP1 expressed from a GPD promoter on a Leu2-marked plasmid.	Gift from Charles Query.
pAAH0514	ACT1-CUP1 A3U	5′ss A3U reporter used for ACT1-CUP1 assays. ACT1-CUP1 expressed from a GPD promoter on a Leu2-marked plasmid.	Gift from Charles Query.
pAAH0524	ACT1-CUP1 A258U	BS A258U reporter used for ACT1-CUP1 assays. ACT1-CUP1 expressed from a GPD promoter on a Leu2-marked plasmid.	Gift from Charles Query.
pAAH0441	ACT1-CUP1 BSC (A259C)	BS A259C reporter used for ACT1-CUP1 assays. ACT1-CUP1 expressed from a GPD promoter on a Leu2-marked plasmid.	Gift from Charles Query.
pAAH0880	ACT1-CUP1 BSG (A259G)	BS A259C reporter used for ACT1-CUP1 assays. ACT1-CUP1 expressed from a GPD promoter on a Leu2-marked plasmid.	Gift from Charles Query.
pAAH0526	ACT1-CUP1 U301G	3′ss gAG reporter used for ACT1-CUP1 assays. ACT1-CUP1 expressed from a GPD promoter on a Leu2-marked plasmid.	Gift from Charles Query.

pAAH0527	ACT1-CUP1 A302U	3′ss UuG reporter used for ACT1-CUP1 assays. ACT1-CUP1 expressed from a GPD promoter on a Leu2-marked plasmid.	Gift from Charles Query.
pAAH0734	ACT1-CUP1 3′ss competition	3′ss competition reporter used for ACT1-CUP1 assays. ACT1-CUP1 expressed from a GPD promoter on a Leu2-marked plasmid.	Gift from Charles Query.

4.6 References

Auesukaree C, Damnernsawad A, Kruatrachue M, Pokethitiyook P, Boonchird C, Kaneko Y, Harashima S. 2009. Genome-wide identification of genes involved in tolerance to various environmental stresses in Saccharomyces cerevisiae. *J Appl Genet* **50**: 301–310.doi:10.1007/BF03195688

Bastian M, Heymann S, Jacomy M. 2009. Gephi: an open source software for exploring and manipulating networks. *Proc Int AAAI Conf Web Soc Media* **3**: 361–362.doi:10.1609/icwsm.v3i1.13937

Brys A, Schwer B. 1996. Requirement for SLU7 in yeast pre-mRNA splicing is dictated by the distance between the branchpoint and the 3′ splice site. *RNA* **2**: 707–717. Abstract

Carrocci TJ, Zoerner DM, Paulson JC, Hoskins AA. 2017. SF3b1 mutations associated with myelodysplastic syndromes alter the fidelity of branchsite selection in yeast. *Nucleic Acids Res* **45**: 4837–4852.doi:10.1093/nar/gkw1349

Carrocci TJ, Paulson JC, Hoskins AA. 2018. Functional analysis of Hsh155/SF3b1 interactions with the U2 snRNA/branch site duplex. *RNA* **24**: 1028–1040.doi:10.1261/rna.065664.118

Chiu Y-F, Liu Y-C, Chiang T-W, Yeh T-C, Tseng C-K, Wu N-Y, Cheng S-C. 2009. Cwc25 is a novel splicing factor required after Prp2 and Yju2 to facilitate the first catalytic reaction. *Mol Cell Biol* **29**: 5671–5678.doi:10.1128/MCB.00773-09

Crawford DJ, Hoskins AA, Friedman LJ, Gelles J, Moore MJ. 2008. Visualizing the splicing of single pre-mRNA molecules in whole cell extract. *RNA* **14**: 170–179.doi:10.1261/rna.794808

Crotti LB, Bačíková D, Horowitz DS. 2007. The Prp18 protein stabilizes the interaction of both exons with the U5 snRNA during the second step of pre-mRNA splicing. *Genes Dev* **21**: 1204–1216.doi:10.1101/gad.1538207

Cuenca-Bono B, García-Molinero V, Pascual-García P, Dopazo H, Llopis A, Vilardell J, Rodríguez-Navarro S. 2011. SUS1 introns are required for efficient mRNA nuclear export in yeast. *Nucleic Acids Res* **39**: 8599–8611.doi:10.1093/nar/gkr496

Eysmont K, Matylla-Kulińska K, Jaskulska A, Magnus M, Konarska MM. 2019. Rearrangements within the U6 snRNA core during the transition between the two catalytic steps of splicing. *Mol Cell* **75**: 538–548.e3.doi:10.1016/j.molcel.2019.05.018

Fabrizio P, Dannenberg J, Dube P, Kastner B, Stark H, Urlaub H, Lührmann R. 2009. The evolutionarily conserved core design of the catalytic activation step of the yeast spliceosome. *Mol Cell* **36**: 593–608.doi:10.1016/j.molcel.2009.09.040

Fica SM, Nagai K. 2017. Cryo-electron microscopy snapshots of the spliceosome: structural insights into a dynamic ribonucleoprotein machine. *Nat Struct Mol Biol* **24**: 791–799.doi:10.1038/nsmb.3463

Fica SM, Oubridge C, Galej WP, Wilkinson ME, Bai X-C, Newman AJ, Nagai K. 2017. Structure of a spliceosome remodelled for exon ligation. *Nature* **542**: 377–380.doi:10.1038/nature21078

Frank D, Guthrie C. 1992. An essential splicing factor, SLU7, mediates 3′ splice site choice in yeast. *Genes Dev* **6**: 2112–2124.doi:10.1101/gad.6.11.2112

Goldstein AL, McCusker JH. 1999. Three new dominant drug resistance cassettes for gene disruption in Saccharomyces cerevisiae. *Yeast* **15**: 1541–1553.doi:10.1002/(SICI)1097-0061(199910)15:14<1541::AID-YEA476>3.0.CO;2-K

Hoskins AA, Friedman LJ, Gallagher SS, Crawford DJ, Anderson EG, Wombacher R, Ramirez N, Cornish VW, Gelles J, Moore MJ. 2011. Ordered and dynamic assembly of single spliceosomes. *Science* **331**:1289-95. doi: 10.1126/science.1198830

Hossain MA, Claggett JM, Nguyen T, Johnson TL. 2009. The cap binding complex influences H2B ubiquitination by facilitating splicing of the SUS1 pre-mRNA. *RNA* **15**: 1515–1527.doi:10.1261/rna.1540409

Hossain MA, Rodriguez CM, Johnson TL. 2011. Key features of the two-intron Saccharomyces cerevisiae gene SUS1 contribute to its alternative splicing. *Nucleic Acids Res* **39**: 8612–8627.doi:10.1093/nar/gkr497

Hotz H-R, Schwer B. 1998. Mutational analysis of the yeast DEAH-box splicing factor Prp16. *Genetics* **149**: 807–815.doi:10.1093/genetics/149.2.807

Huh W-K, Falvo JV, Gerke LC, Carroll AS, Howson RW, Weissman JS, O'Shea EK. 2003. Global analysis of protein localization in budding yeast. *Nature* **425**: 686–691.doi:10.1038/nature02026

James S-A, Turner W, Schwer B. 2002. How Slu7 and Prp18 cooperate in the second step of yeast pre-mRNA splicing. *RNA* **8**: 1068–1077.doi:10.1017/s1355838202022033

Jonik-Nowak B, Menneteau T, Fesquet D, Baldin V, Bonne-Andrea C, Méchali F, Fabre B, Boisguerin P, de Rossi S, Henriquet C, et al. 2018. PIP30/FAM192A is a novel

regulator of the nuclear proteasome activator PA28γ. *Proc Natl Acad Sci* **115**: E6477–E6486.doi:10.1073/pnas.1722299115

Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, Tunyasuvunakool K, Bates R, Žídek A, Potapenko A, et al. 2021. Highly accurate protein structure prediction with AlphaFold. *Nature* **596**: 583–589.doi:10.1038/s41586-021-03819-2

Kao CY, Cao EC, Wai HL, Cheng SC. 2021. Evidence for complex dynamics during U2 SnRNP selection of the intron branchpoint. *Nucleic Acids Res* **49**: 9965–9977.doi:10.1093/nar/gkab695

Kaur H, van der Feltz C, Sun Y, Hoskins AA. 2022. Network theory reveals principles of spliceosome structure and dynamics. *Structure* **30**: 190–200.e2.doi:10.1016/j.str.2021.09.003

Kawashima T, Pellegrini M, Chanfreau GF. 2009. Nonsense-mediated mRNA decay mutes the splicing defects of spliceosome component mutations. *RNA* **15**: 2236–2247.doi:10.1261/rna.1736809

Kawashima T, Douglass S, Gabunilas J, Pellegrini M, Chanfreau GF. 2014. Widespread use of non-productive alternative splice sites in Saccharomyces cerevisiae. *PLoS Genet* **10**: e1004249.doi:10.1371/journal.pgen.1004249

Lesser CF, Guthrie C. 1993. Mutational analysis of pre-mRNA splicing in Saccharomyces cerevisiae using a sensitive new reporter gene, CUP1. *Genetics* **133**: 851–863.doi:10.1093/genetics/133.4.851

Liu Y-C, Chen H-C, Wu N-Y, Cheng S-C. 2007a. A novel splicing factor, Yju2, is associated with NTC and acts after Prp2 in promoting the first catalytic reaction of pre-mRNA splicing. *Mol Cell Biol* **27**: 5403–5413.doi:10.1128/MCB.00346-07

Liu L, Query CC, Konarska MM. 2007b. Opposing classes of prp8 alleles modulate the transition between the catalytic steps of pre-mRNA splicing. *Nat Struct Mol Biol* **14**: 519–526.doi:10.1038/nsmb1240

Liu S, Li X, Zhang L, Jiang J, Hill RC, Cui Y, Hansen KC, Zhou ZH, Zhao R. 2017. Structure of the yeast spliceosomal postcatalytic P complex. *Science* **358**: 1278–1283.doi:10.1126/science.aar3462

Maitra N, Anandhakumar J, Blank HM, Kaplan CD, Polymenis M. 2019. Perturbations of transcription and gene expression-associated processes alter distribution of cell size values in Saccharomyces cerevisiae. *G3 (Bethesda)* **9**: 239–250.doi:10.1534/g3.118.200854

Mayerle M, Guthrie C. 2016. Prp8 retinitis pigmentosa mutants cause defects in the transition between the catalytic steps of splicing. *RNA* **22**: 793–809.doi:10.1261/rna.055459.115

Needleman SB, Wunsch CD. 1970. A general method applicable to the search for similarities in the amino acid sequence of two proteins. *J Mol Biol* **48**: 443–453.doi:10.1016/0022-2836(70)90057-4

Ohrt T, Odenwälder P, Dannenberg J, Prior M, Warkocki Z, Schmitzová J, Karaduman R, Gregor I, Enderlein J, Fabrizio P, et al. 2013. Molecular dissection of step 2 catalysis of yeast pre-mRNA splicing investigated in a purified system. *RNA* **19**: 902–915.doi:10.1261/rna.039024.113

Pagé N, Gérard-Vincent M, Ménard P, Beaulieu M, Azuma M, Dijkgraaf GJP, Li H, Marcoux J, Nguyen T, Dowse T, et al. 2003. A Saccharomyces cerevisiae genome-wide mutant screen for altered sensitivity to K1 killer toxin. *Genetics* **163**: 875–894.doi:10.1093/genetics/163.3.875

Pettersen EF, Goddard TD, Huang CC, Meng EC, Couch GS, Croll TI, Morris JH, Ferrin TE. 2021. UCSF ChimeraX: structure visualization for researchers, educators, and developers. *Protein Sci* **30**: 70–82.doi:10.1002/pro.3943

Plaschka C, Newman AJ, Nagai K. 2019. Structural basis of nuclear pre-mRNA splicing: lessons from yeast. *Cold Spring Harb Perspect Biol* **11**: a032391.doi:10.1101/cshperspect.a032391

Query CC, Konarska MM. 2004. Suppression of multiple substrate mutations by spliceosomal prp8 alleles suggests functional correlations with ribosomal ambiguity mutants. *Mol Cell* **14**: 343–354.doi:10.1016/S1097-2765(04)00217-5

Query CC, Konarska MM. 2006. Splicing fidelity revisited. *Nat Struct Mol Biol* **13**: 472–474.doi:10.1038/nsmb0606-472

Query CC, Konarska MM. 2012. CEF1/CDC5 alleles modulate transitions between catalytic conformations of the spliceosome. *RNA* **18**: 1001–1013.doi:10.1261/rna.029421.111

Sayani S, Janis M, Lee CY, Toesca I, Chanfreau GF. 2008. Widespread impact of nonsense-mediated mRNA decay on the yeast intronome. *Mol Cell* **31**: 360–370.doi:10.1016/j.molcel.2008.07.005

Schindelin J, Rueden CT, Hiner MC, Eliceiri KW. 2015. The ImageJ ecosystem: an open platform for biomedical image analysis. *Mol Reprod Dev* **82**: 518–529.doi:10.1002/mrd.22489

Schneider S, Campodonico E, Schwer B. 2004. Motifs IV and V in the DEAH box splicing factor Prp22 are important for RNA unwinding, and helicase-defective Prp22 mutants are suppressed by Prp8. *J Biol Chem* **279**: 8617–8626.doi:10.1074/jbc.M312715200

Schwer B. 2008. A conformational rearrangement in the spliceosome sets the stage for Prp22-dependent mRNA release. *Mol Cell* **30**: 743–754.doi:10.1016/j.molcel.2008.05.003

Schwer B, Gross CH. 1998. Prp22, a DExH-box RNA helicase, plays two distinct roles in yeast pre-mRNA splicing. *EMBO J* **17**: 2086–2094.doi:10.1093/emboj/17.7.2086

Schwer B, Guthrie C. 1992. A conformational rearrangement in the spliceosome is dependent on PRP16 and ATP hydrolysis. *EMBO J* **11**: 5033–5039.doi:10.1002/j.1460-2075.1992.tb05610.x

Schwer B, Meszaros T. 2000. RNA helicase dynamics in pre-mRNA splicing. *EMBO J* **19**: 6582–6591.doi:10.1093/emboj/19.23.6582

Semlow DR, Blanco MR, Walter NG, Staley JP. 2016. Spliceosomal DEAH-Box ATPases remodel pre-mRNA to activate alternative splice sites. *Cell* **164**: 985–998.doi:10.1016/j.cell.2016.01.025

Sikorski RS, Boeke JD. 1991. In vitro mutagenesis and plasmid shuffling: from cloned gene to mutant yeast. *Methods Enzymol* **194**: 302–318.doi:10.1016/0076-6879(91)94023-6

Treco DA, Lundblad V. 1993. Preparation of yeast media. *Curr Protoc Mol Biol* **23**: 13.1.1–13.1.7.doi:10.1002/0471142727.mb1301s23

Umen JG, Guthrie C. 1995a. A novel role for a U5 snRNP protein in 3′ splice site selection. *Genes Dev* **9**: 855–868.doi:10.1101/gad.9.7.855

Umen JG, Guthrie C. 1995b. Prp16p, Slu7p, and Prp8p interact with the 3′ splice site in two distinct stages during the second catalytic step of pre-mRNA splicing. *RNA* **1**: 584–597.

Umen JG, Guthrie C. 1996. Mutagenesis of the yeast gene PRP8 reveals domains governing the specificity and fidelity of 3′ splice site selection. *Genetics* **143**: 723-739. doi: 10.1093/genetics/143.2.723

van der Feltz C, Nikolai B, Schneider C, Paulson JC, Fu X, Hoskins AA. 2021. Saccharomyces cerevisiae Ecm2 modulates the catalytic steps of pre-mRNA splicing. *RNA* **27**: 591–603.doi:10.1261/rna.077727.120

Viladevall L, Serrano R, Ruiz A, Domenech G, Giraldo J, Barceló A, Ariño J. 2004. Characterization of the calcium-mediated response to alkaline stress in Saccharomyces cerevisiae. *J Biol Chem* **279**: 43614–43624.doi:10.1074/jbc.M403606200

Villa T, Guthrie C. 2005. The Isy1p component of the NineTeen complex interacts with the ATPase Prp16p to regulate the fidelity of pre-mRNA splicing. *Genes Dev* **19**: 1894–1904.doi:10.1101/gad.1336305

Wagner JDO, Jankowsky E, Company M, Pyle AM, Abelson JN. 1998. The DEAH-box protein PRP22 is an ATPase that mediates ATP-dependent mRNA release from the spliceosome and unwinds RNA duplexes. *EMBO J* **17**: 2926–2937.doi:10.1093/emboj/17.10.2926

Wan R, Bai R, Yan C, Lei J, Shi Y. 2019. Structures of the catalytically activated yeast spliceosome reveal the mechanism of branching. *Cell* **177**: 339–351.e13.doi:10.1016/j.cell.2019.02.006

Wang Y, Guthrie C. 1998. PRP16, a DEAH-box RNA helicase, is recruited to the spliceosome primarily via its nonconserved N-terminal domain. *RNA* **4**: 1216–1229. doi: 10.1017/S1355838298980992

Warkocki Z, Odenwälder P, Schmitzová J, Platzmann F, Stark H, Urlaub H, Ficner R, Fabrizio P, Lührmann R. 2009. Reconstitution of both steps of Saccharomyces cerevisiae splicing with purified spliceosomal components. *Nat Struct Mol Biol* **16**: 1237–1243.doi:10.1038/nsmb.1729

Wilkinson ME, Fica SM, Galej WP, Norman CM, Newman AJ, Nagai K. 2017. Post-catalytic spliceosome structure reveals mechanism of 3′-splice site selection. *Science* **358**: 1283–1288.doi:10.1126/science.aar3729

Wilkinson ME, Fica SM, Galej WP, Nagai K. 2021. Structural basis for conformational equilibrium of the catalytic spliceosome. *Mol Cell* **81**: 1439–1452.e9.doi:10.1016/j.molcel.2021.02.021

Wilson TE. 2002. A genomics-based screen for yeast mutants with an altered recombination/end-joining repair ratio. *Genetics* **162**: 677–688.doi:10.1093/genetics/162.2.677

Yan C, Wan R, Bai R, Huang G, Shi Y. 2017. Structure of a yeast step II catalytically activated spliceosome. *Science* **355**: 149–155.doi:10.1126/science.aak9979

Zhan X, Lu Y, Zhang X, Yan C, Shi Y. 2022. Mechanism of exon ligation by human spliceosome. *Mol Cell* **82**: 2769–2778.e4.doi:10.1016/j.molcel.2022.05.021

CHAPTER 5

Control of 3' splice site selection by the yeast splicing factor Fyv6

A version of this chapter is available as a preprint and will be submitted for publication: Katherine A. Senn*, Karli A. Lipinski*, Natalie J. Zeps, Amory F. Griffin, Max E. Wilkinson, Aaron A. Hoskins. Control of 3' splice site selection by the yeast splicing factor Fyv6. *bioRxiv* 2024.05.04.592262. doi: 10.1101/2024.05.04.592262
*Indicates co-first authors

AAH, KAS, KAL, and MEW conceptualized the project; MEW collected and analyzed the cryo-EM data; KAL and NJZ prepared RNA for sequencing; KAL and AG analyzed the RNA-sequencing data; AAH, KAS, NJZ, and KAL created yeast strains; KAS, KAL, NJZ, and AG performed genetic and biochemical assays and analyzed data; KAS, KAL, MEW, and AAH wrote the manuscript. See contributions (pg. xii).

CHAPTER 5: Control of 3' splice site selection by the yeast splicing factor Fyv6

5.1 Abstract

Pre-mRNA splicing is catalyzed in two steps: 5' splice site (SS) cleavage and exon ligation. A number of proteins transiently associate with spliceosomes to specifically impact these steps (1st and 2nd step factors). We recently identified Fyv6 (FAM192A in humans) as a 2nd step factor in *S. cerevisiae*; however, we did not determine how widespread Fyv6's impact is on the transcriptome. To answer this question, we have used RNA-Seq to analyze changes in splicing. These results show that loss of Fyv6 results in activation of non-consensus, branch point (BP) proximal 3' SS transcriptome-wide. To identify the molecular basis of these observations, we determined a high-resolution cryo-EM structure of a yeast product complex spliceosome containing Fyv6 at 2.3 Å. The structure reveals that Fyv6 is the only 2nd step factor that contacts the Prp22 ATPase and that Fyv6 binding is mutually exclusive with that of the 1st step factor Yju2. We then use this structure to dissect Fyv6 functional domains and interpret results of a genetic screen for *fyv6Δ* suppressor mutations. The combined transcriptomic, structural, and genetic studies allow us to propose a model in which Yju2/Fyv6 exchange facilitates exon ligation and Fyv6 promotes usage of Prp22-dependent, BP distal 3' SS.

5.2 Introduction

Precursor messenger RNA (pre-mRNA) splicing is catalyzed by a large macromolecular complex called the spliceosome. Spliceosomes are composed of 5 small nuclear ribonucleoproteins (snRNPs), each composed of a small nuclear RNA (snRNA; U1, U2, U4, U5, and U6) and several different protein splicing factors. The snRNPs and dozens of other proteins assemble *de novo* at each intron to form spliceosomes. Spliceosomes are highly dynamic and form different complexes as proteins and snRNAs join and leave. Many of these complexes have been characterized biochemically, genetically, and by cryo-electron microscopy (cryo-EM) (Plaschka et al., 2019). The splicing reaction itself is carried out in two sequential transesterification reactions (**Fig. 5.1A**). First, the 5' splice site (5' SS) is cleaved by formation of an intron lariat (1st step). Second, the intron lariat is removed simultaneously with exon ligation at the 3' SS (2nd step).

The integrity of the genetic information contained within a pre-mRNA depends on correct identification of the 5' and 3' SS by the splicing machinery since a single nucleotide shift in either site could destroy a protein reading frame. In yeast, several DExD/H-box ATPases function to limit usage of suboptimal SS by enhancing splicing fidelity (Chung et al., 2023; Semlow and Staley, 2012). Two of these ATPases, Prp16 and Prp22, impact the fidelity of the 1st and 2nd catalytic steps, respectively. In addition to these roles, Prp16 is also required for remodeling of the spliceosome to permit the 1st-to-2nd step transition

(Schwer and Guthrie, 1992), and Prp22 is essential for releasing the mRNA product (Company et al., 1991). While Prp16 does not need to be present during the 1st step (Chung et al., 2023), Prp22 is required to be present for exon ligation if the branch point (BP) to 3' SS distance is ≥ 21 nucleotides (nt) (Schwer and Gross, 1998). Despite many biochemical, single molecule, and structural studies, how spliceosomes promote usage of BP distal, Prp22-dependent 3' SS has remained elusive.

Identification of the productive 3' SS is an especially challenging problem given that the consensus sequence in *Saccharomyces cerevisiae* (yeast) and humans is just three nucleotides (YAG, Y=U or C). Cryo-EM structures of spliceosome product (P) complexes have revealed how these nucleotides can be recognized within the spliceosome active site (Bai et al., 2017; Fica et al., 2019; Liu et al., 2017; Wilkinson et al., 2017; Zhang et al., 2019). The AG 3' SS dinucleotide is recognized by formation of non-Watson Crick base-pairing interactions with the 5' SS +1G and BP adenosine, while the pyrimidine at the -3 position of the 3' SS is recognized by the Prp8 protein. Given this short consensus sequence, multiple different splicing factors also help to ensure that the productive 3' SS is utilized, including Cwc21 and the 2nd step factors Slu7, Prp18, and Prp22 (Crotti et al., 2007; Frank and Guthrie, 1992; Gautam et al., 2015; Kawashima et al., 2014; Roy et al., 2023; Semlow et al., 2016). It is believed that Slu7 and Prp18 help to stabilize 3' SS docking to the active site and that this contributes to 2nd step efficiency. It has been proposed that Prp22 antagonizes 3' SS docking to permit sampling of different 3' SS to ensure that optimal sequences are used as a proofreading mechanism (Mayas et al., 2006; Semlow et al., 2016). Importantly, identification of the productive 3' SS depends not just on identification of the YAG sequence but also involves choosing which YAG sequence to use. Prp18 appears to aid in selection of BP distal 3' SS and avoidance of BP proximal (and often non-consensus) 3' SS. This observation could be due to Prp18 imposing a BP to 3' SS distance constraint, enforcing use of canonical YAG 3' SS sequences, or a combination of the two activities.

Recently, Fyv6 was identified as a novel 2nd step factor in yeast (Lipinski et al., 2023). Fyv6 is a homolog of the human protein FAM192A, which was identified in pre-C* structures of the human spliceosome (Zhan et al., 2022). While the impact of FAM192A on human spliceosome activity has not yet been well-characterized, loss of Fyv6 in yeast decreases 2nd step splicing efficiency *in vitro* and results in use of an alternative 3' SS in the *SUS1* pre-mRNA *in vivo*. It is likely that unassigned cryo-EM densities in yeast C* (just prior to exon ligation) and P complexes correspond to Fyv6 since these densities are in analogous locations to that for FAM192A in the human pre-C* structure. However, the resolutions of the yeast C* and P complex structures make unambiguous assignment of Fyv6 amino acid side chains and interactions difficult.

Here, we use a combination of transcriptomic, structural, biochemical, and genetic assays to elucidate how Fyv6 controls 3' SS usage in yeast. RNA-seq analysis reveals widespread activation of alternative, BP proximal nonconsensus 3' SS in the absence of

Fyv6 consistent with biochemical assays that show Fyv6 facilitates splicing of 3' SS located ≥ 21 nt from the BP. To further elucidate Fyv6 function, we determined a 2.3 Å cryo-EM structure of the yeast P complex spliceosome in which Fyv6 can be modeled. The structure reveals interactions between Fyv6 and Prp22 as well as a domain in Syf1 that interacts with either Fyv6 or the 1st step factor Yju2. 3D classification allowed us to identify two additional conformational states of P complex and provides insights into the coupling of 3' SS active site docking and the presence of 2nd step factors. We then used this structural data to dissect the functional domains of Fyv6, interpret the results of a genetic screen for suppressors of *fyv6Δ*, and to probe Fyv6 and Prp22 interactions. Combined, we propose a model in which Fyv6 is recruited to the spliceosome via interactions with the NTC component Syf1 in order to promote usage of BP distal, and Prp22-dependent, 3' SS.

5.3 Results

Deletion of *FYV6* Results in Widespread Use of Alternative 3′ Splice Sites

Our previous work showed that loss of *FYV6* causes changes in 3' SS usage in the *SUS1* pre-mRNA. To probe for changes transcriptome-wide, we used nonsense mediated decay (NMD)-deficient (*upf1Δ*, WT) yeast cells expressing or lacking Fyv6. By suppressing NMD, we hoped to limit degradation of alternatively spliced mRNA isoforms generated in the absence of Fyv6 (Sayani et al., 2008). Consistent with results from *fyv6Δ* strains, the *fyv6Δ upf1Δ* double mutant strain also showed reduced growth at 30°C and cold-sensitive (*cs*) and temperature-sensitive (*ts*) phenotypes relative to a *upf1Δ* control (**Fig. 5.1B**). We then used RNA-Seq to analyze the isolated RNAs from these strains. As expected, many more changes in splicing were observed in the *fyv6Δ upf1Δ* strain relative to the *fyv6Δ* strain (**Fig. 5.2A, B; Table 5.1**). This is consistent with many mRNA isoforms generated due to loss of Fyv6 being substrates for NMD. Loss of Fyv6 resulted in a number of changes in gene expression with many non-intron containing genes also being up- or down-regulated (**Fig. 5.2C**).

We compared and classified changes in splicing between strains containing and lacking *FYV6* under permissive growth conditions (30°C) or after the yeast had been shifted to a non-permissive temperature (16 or 37°C) for 1 h (**Fig. 5.1C**). In this case, we only considered events resulting in at least a 10% change in the percent spliced in (PSI) value when Fyv6 is deleted. At all temperatures, we see changes in alternative 3' SS usage in the *fyv6Δ* strains relative to those with *FYV6*. For example, we detected use of the alternative 3' SS in the first intron of the *SUS1* transcript as we previously reported (not shown; Lipinski et al., 2023). We also saw increased use of a cryptic, nonconsensus GAG 3' SS in the first, but not second, intron of the *YOS1* transcript in the *fyv6Δ* data sets and confirmed this result by RT-PCR (**Fig. 5.2D, E**). We observed several cases of

alternative 5' SS. However, the majority of these alternative 5' SS were used with equal efficiencies in the WT and *fyv6Δ* strains (**Fig. 5.2F**).

The strains temperature shifted to 16°C had the largest number of changes between strains with and without Fyv6 and the strains shifted to 37°C the least. The temperature shift to 16° caused an increase in the numbers of both alternative 3' SS and retained intron events detected (**Fig. 5.1C**). The changes in alternative 3' SS use at each temperature were mostly unique with few events detected under all three conditions (**Fig. 5.1D**). Collectively, we observed that 61 different RNAs changed splicing patterns due to loss of Fyv6. For many of these mRNAs, we detected multiple different 3' SS being used.

To investigate these splicing events more closely, we analyzed changes in usage of alterative splice junctions by calculating the Fraction of Annotated Splicing (FAnS) in the presence and absence of Fyv6 (Roy et al., 2023). FAnS ratios report the relative abundance of an alternative splicing event in relation to the main spliced isoform. We calculated FAnS values for all exon junction reads corresponding to usage of a canonical 5' SS and an alternative 3' SS for WT and *fyv6Δ* strains at each temperature. We then calculated the ratios of the *fyv6Δ*/WT FAnS values for each transcript. These results confirm changes in alternative 3' SS usage for both non-ribosomal protein gene (**Fig. 5.1E**) and RPG transcripts (**Fig. 5.1F**) at all conditions. We could not detect any significant differences between RPG and non-RPG splicing with this analysis, suggesting that differences in gene expression and splicing efficiencies between these two classes are not correlated with alternative 3' SS use. Combined these results show that Fyv6 plays a critical role in defining mRNA isoform production in yeast, particularly under non-optimal growth conditions such as temperature stress.

Table 5.1 RNA-seq datasets used for analysis.

Datasets	Description	Figures
WT-16-12-8, WT-16-1-5, WT-30-12-8, WT-30-1-5, Fyv6-16-12-8, Fyv6-16-1-5, Fyv6-30-12-8, Fyv6-30-1-5, Fyv6-37-12-8, Fyv6-37-1-5	Temperature shifted RNA-seq datasets for WT *upf1Δ* and *fyv6Δ upf1Δ* yeast strains	Fig. 1
WT-0409, WT-0417, Fyv6-0409, Fyv6-0417	RNA-seq datasets for WT and fyv6Δ yeast strains	Fig. S1C
WT0803, WT0811, Del0803, Del0811	RNA-seq datasets for WT upf1Δ and fyv6Δ upf1Δ yeast strains	Fig. S1A,D; Fig. 2

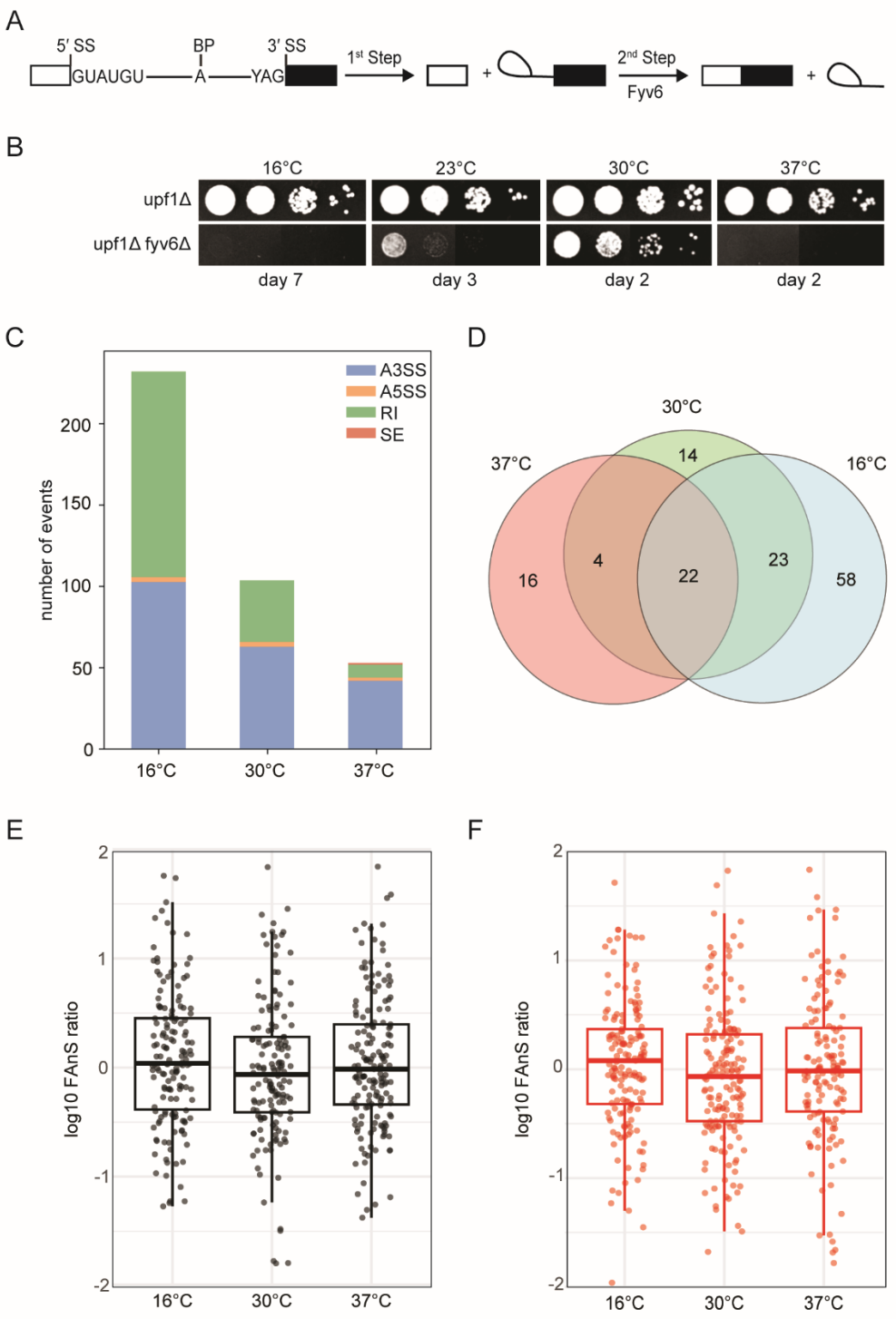

Figure 5.1 Splicing and changes in WT, *fyv6Δ*, *upf1Δ*, and *fyv6Δ upf1Δ* strains. A) Splicing reaction steps and products. **B)** Growth of yeast at 16, 23, 30, and 37°C. Images were taken at the indicated days. **C)** Splicing efficiencies for RNA-seq datasets collected from *upf1Δ* yeast shifted to either 16 (red), 30 (blue), or 37°C (green) for 1 hr. **D)** Number of alternative splicing events determined by SpliceWiz at 16, 30, and 37°C. **E)** Venn diagram of shared alternative splicing events at 16 (blue), 30 (green), and 37°C (blue). **F)** Plot of ratios of Fyv6 to WT FAnS at **G)** 16 and 30°C and **H)** 37 and 30°C for alternative 3′ SS with an annotated 5′ SS and nonzero FAnS for both WT and Fyv6 at both plotted temperatures.

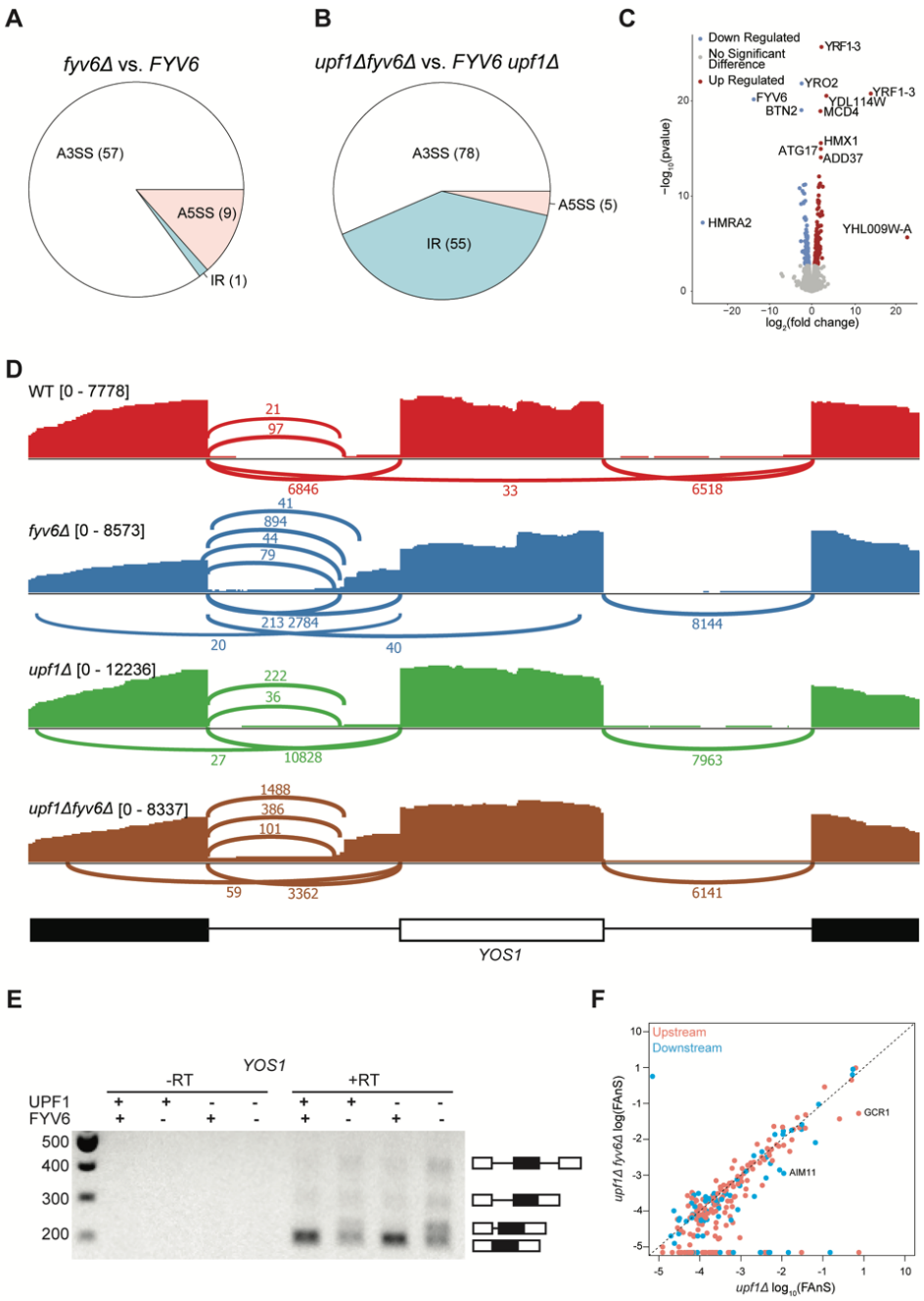

Figure 5.2 Gene expression analysis based on the RNA-Seq results and an example of Fyv6-dependent splicing changes in *YOS1*. A, B) Pie charts comparing alternative splicing events discovered by SpliceWiz at 10 PSI in (A) *fyv6Δ* and **(**B) *upf1Δ fyv6Δ* strain backgrounds relative to when Fyv6 is present**.** A3SS = alternate 3' SS; A5SS = alternate 5' SS; IR = intron retention. Note that these data sets were collected at higher read depth than those shown in Fig. 1. As a result, the number of detected changes in splicing cannot be directly compared between the two. **C)** Differential expression analysis in a *upf1Δ* background with presence or absence of Fyv6 Colored points are significant with $p \leq 0.05$ with blue points being downregulated gene expression and red being upregulated. **D)** Sashimi plots showing coverage across splice junctions of *YOS1* in WT (red), *fyv6Δ* (blue), *upf1Δ* (green), and *fyv6Δ upf1Δ* (brown). **E)** RT-PCR of YOS1 mRNA showing a higher molecular weight band in strains lacking *FYV6* corresponding to use of an alternative 3' SS in the first intron. **F**) FAnS plot for alternative 5' SS junctions with canonical 3' SS in *upf1Δ* compared to *fyv6Δ upf1Δ.*

Fyv6 Facilitates Usage of Branch Point Distal 3′ Splice Sites

We next analyzed the features of the 3' SS impacted by Fyv6 loss. We used RNAs collected from non-temperature shifted yeast grown at 30°C and sequenced these RNAs at higher depth than those used for the temperature-shift experiment described in **Fig. 5.1** (~400 million vs. 200 million reads; **Table 5.1**). We annotated each alternative 3' SS as being upstream (5') or downstream (3') of the canonical 3' SS (Grate and Ares, 2002). This analysis shows that the majority of the alternative 3' SS arising from Fyv6 deletion are upstream (**Fig. 5.3A**, pink points above the diagonal). When these alternative 3' SS are mapped according to their nucleotide distance from the canonical site, most of these are found within ~40 nt upstream (**Fig. 5.3B**, purple violin). Plotting these alternative 3' SS sites relative to their distance from the annotated BP reveals that the highest density of sites is $\leq$ 20 nt from the BP for transcripts with the highest FAnS ratios (**Fig. 5.3C**, purple violin). It is important to note that we did not map BP in Fyv6 deletion strains, and it is possible that in some cases a shift in 3' SS usage could also coincide with a shift in the BP. Indeed, a small number of alternative 3' SS detected in the Fyv6 data sets are located upstream of the annotated BP, suggesting activation of an alternative BP and 3' SS. However, we believe that the changes are predominantly due to usage of alternative 3' SS located between the annotated BP and canonical 3' SS.

The above results suggest that Fyv6 is important for splicing of 3' SS located distant from the BP. To test this directly and systematically, we created a series of ACT1-CUP1 reporters with various BP to 3' SS distances (9-50 nt, **Fig. 5.3D**, **Table 5.2**) based on sequences used in previous studies of 3' SS selection (Brys and Schwer, 1996; Frank and Guthrie, 1992; Schwer and Gross, 1998). We transformed these reporters into strains containing or lacking Fyv6 and without debranchase (Dbr1) to limit degradation of lariat intermediates. We then detected RNA products by primer extension and calculated the 2nd step (exon ligation) efficiency for each reporter (**Fig. 5.3E, F**). While loss of Fyv6 has minimal impact on exon ligation when BP to 3' SS distances are short (9-15 nt), exon ligation efficiency decreases significantly at distances of 21 nt or greater. In addition, we do not see evidence of cryptic BP usage in these assays. This result is consistent with the RNA-seq analysis and supports a function for Fyv6 in facilitating splicing at BP distal 3' SS.

Finally, we analyzed the sequence features of the alternative 3' SS activated by loss of Fyv6 (**Fig. 5.3G, H**). For sites located upstream of the canonical 3' SS, we observed much greater sequence variability relative to the canonical site. We detected increased use of highly variable 3' SS including those with atypical HAU (H=A,C,U) and BG (B=C,U,G) motifs (**Fig. 5.3H**). Increased use of variable 3' SS was also observed when the 2nd step factor Prp18 was deleted (Roy et al., 2023). However, the subset of RNAs most impacted by splicing factor deletion appear to be different. We were not able to detect changes in 3' SS usage for the *UBC12*, *MAF1*, *MUD1*, *PHO85*, *SPT14*, or *YCL002C* transcripts in the *fyv6*D data sets as was reported for *prp18*D. Prp18 and Fyv6

likely have transcript-specific effects on splicing outcomes. Combined our results strongly support Fyv6 as a 2nd step splicing factor that facilitates usage of consensus sites located distal to the BP *in vivo*.

Table 5.2 Sequences of introns between the branch point and 3' SS in ACT1-CUP1 reporters.

BP-3' SS distance	Intron sequence between branch site and 3' SS*
9 nt	**UACUAACA**UCGAUUAUA**UAG**
12 nt	**UACUAACA**UCGAUUUGUUUA**UAG**
15 nt	**UACUAACA**UCGAUUAUAUGUUUA**UAG**
21 nt	**UACUAACA**UCGUUCUUCUUUCCGAUUAUA**UAG**
27 nt	**UACUAACA**UCGUUCUUCUUUCCGAUUAUAUGUUUA**UAG**
38 nt	**UACUAACA**UCGAUUGCUUCAUUCUUUUUGUUGCUAUAUUAUAUGUU**UAG**
42 nt	**UACUAACA**UCGAAACAUUGCUUCAUUCUUUUUGUUGCUAUAUUAUAUGUU**UAG**
46 nt	**UACUAACA**UCGAAACAACAAUUGCUUCAUUCUUUUUGUUGCUAUAUUAUAUGUU**UAG**
50 nt	**UACUAACA**UCGAAACAACAAACGAUUGCUUCAUUCUUUUUGUUGCUAUAUUAUAUGUU**UAG**

*Branch site and 3' SS sequences in bold

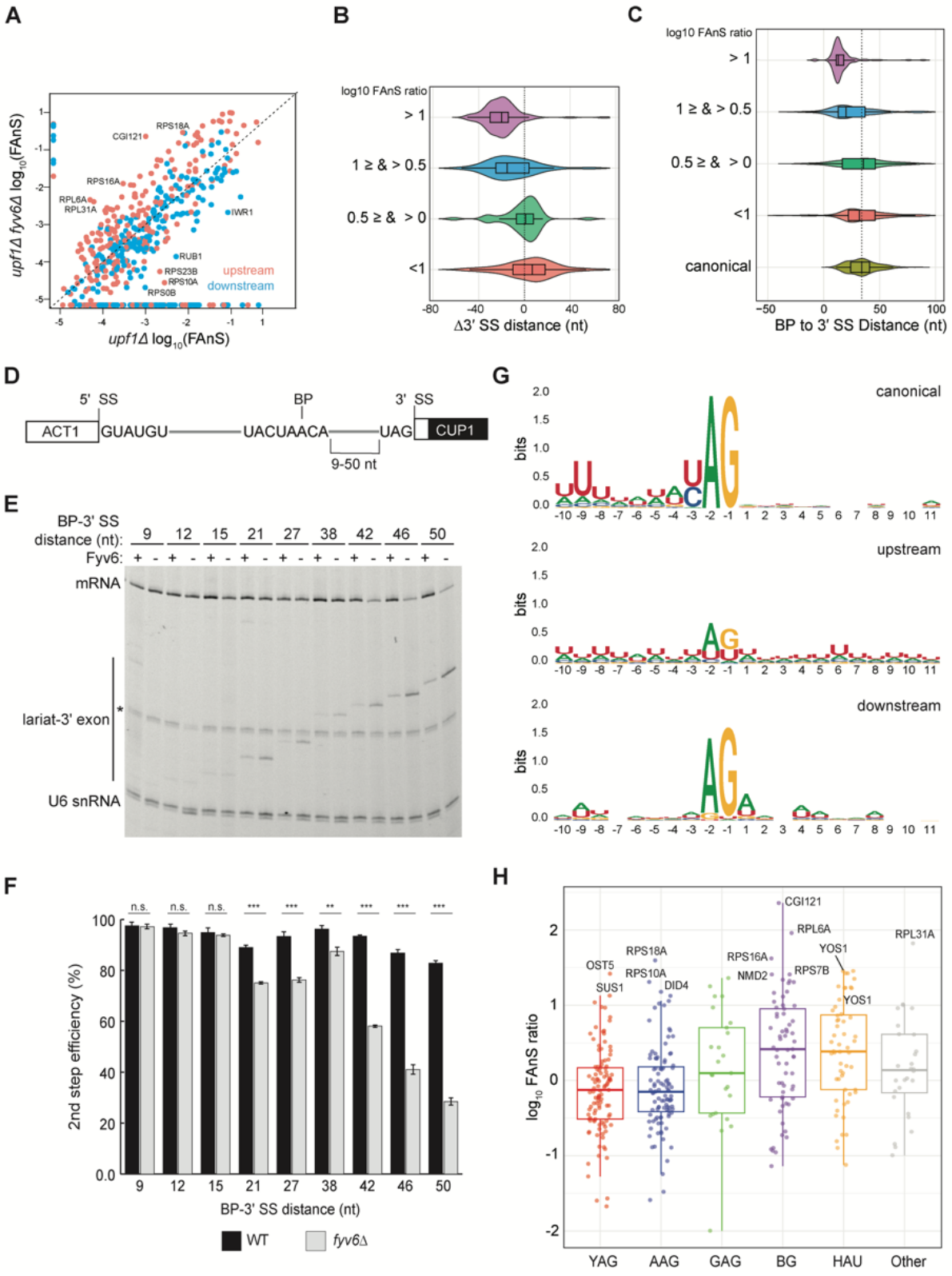
A
upf1Δ fyv6Δ log10(FAnS)
upf1Δ log10(FAnS)
CGI121
RPS18A
RPS16A
RPL6A
RPL31A
IWR1
RUB1
RPS23B
RPS10A
RPS0B
upstream
downstream
B
log10 FAnS ratio
> 1
1 ≥ & > 0.5
0.5 ≥ & > 0
<1
Δ3′ SS distance (nt)
C
log10 FAnS ratio
> 1
1 ≥ & > 0.5
0.5 ≥ & > 0
<1
canonical
BP to 3′ SS Distance (nt)
D
5′ SS
BP
3′ SS
ACT1
GUAUGU
UACUAACA
UAG
CUP1
9-50 nt
E
BP-3′ SS distance (nt):
9 12 15 21 27 38 42 46 50
Fyv6:
+ - + - + - + - + - + - + - + - +
mRNA
lariat-3′ exon
*
U6 snRNA
G
canonical
upstream
downstream
bits
F
n.s.
***
**
2nd step efficiency (%)
BP-3′ SS distance (nt)
WT
fyv6Δ
H
log10 FAnS ratio
CGI121
RPL6A
RPL31A
YOS1
RPS18A
RPS16A
OST5
RPS10A
NMD2
RPS7B
SUS1
DID4
YAG
AAG
GAG
BG
HAU
Other

Figure 5.3 Loss of Fyv6 activates BP proximal, non-consensus 3' SS. A) Plot of the $\log_{10}$ of FAnS values for changes in alternative 3' SS usage in *fyv6Δ upf1Δ* and *upf1Δ* strains. Points are colored based on whether they are upstream (pink) or downstream (blue) to the canonical 3' SS. **B)** Violin plots of the ratios of *fyv6Δupf1Δ* and *upf1Δ* FAnS values based on the distances between the alternative and canonical 3' SS. FAnS ratios >0 are indicative of the site being upregulated in *fyv6Δ*. **C)** Violin plots of the ratios of *fyv6Δupf1Δ* and *upf1Δ* FAnS values based on the distances between the alternative 3' SS and the annotated BP. FAnS ratios >0 are indicative of the site being upregulated in *fyv6Δ*. **D)** Diagram of ACT1-CUP1 reporter showing BP-3' SS distances. **E)** Representative primer extension analysis of RNA products generated from splicing of the ACT1-CUP1 reporter in the presence (WT) or absence of Fyv6 (*fyv6Δ*). Bands for fully spliced mRNA and lariat intermediate are indicated. U6 snRNA was detected as a loading control. The * indicates an unknown product present in every lane. **F)** Quantification of the primer extension results from *N*=3 replicates represented by the ratio of band intensities for mRNA/(mRNA+lariat intermediate). Bars represent the average ratio of the replicates ±SD. Means between WT and *fyv6Δ* groups for each reporter were compared with an unpaired Welch's two-tailed t-test. Significance is indicated: n.s. no significance; ** $p<0.01$; *** $p<0.001$. **G)** Sequence logos of alternative 3' SS with FAnS > 0 sorted by either upstream or downstream of the canonical 3' SS compared against the canonical 3' SS for genes with FanS > 0. **H)** The $\log_{10}$ of the FAnS ratio for alternative 3' SS sorted by the sequences of the 3' SS.

A High-Resolution Spliceosome Structure Reveals Fyv6 Interactions

Previous structures of the yeast P-complex spliceosome were solved with resolutions at the core ranging from 3.3 Å to 3.7 Å (Bai et al., 2017; Liu et al., 2017; Wilkinson et al., 2017), and with peripheral components such as the Prp22 helicase, U2 snRNP, and Prp19 Complex (NTC) at much lower resolutions (5-10 Å) that precluded detailed investigation. To improve the resolution, we purified P complex as previously described by stalling spliceosome disassembly with a dominant negative mutant of Prp22 (S635A) defective in mRNA release (Schwer and Meszaros, 2000; Wilkinson et al., 2017). However, by collecting a much larger cryo-EM dataset on a more modern electron detector, we were able to resolve the structure to 2.3 Å within the catalytic core, as well as visualize peripheral components at resolutions from 3.0 Å – 3.7 Å (**Fig. 5.4A, B; Fig. 5.5-5.8; Table 5.3**). This is the highest resolution spliceosome structure to date. The high-quality density at the active site shows the positions of bases unambiguously, confirming the manner of the non-Watson-Crick base pairing that mediates 3' SS recognition (Liu et al., 2017; Wilkinson et al., 2017). Monovalent ions bound at the active site were previously shown to be important during the 1st step of splicing (Wilkinson et al., 2021), and our data indicate that the positions of these ions are preserved after exon ligation. Additionally, the density also shows a metal ion, likely potassium, that bridges the base pair between +1G of the 5' SS and -1G of the 3' SS (**Fig. 5.4C**).

The 2nd-step splicing factor Prp22 is resolved at 3.0 Å resolution, allowing atomic interpretation for the first time (**Fig. 5.4B** and later figures). Bases 13 – 21 of the 3' exon are visible between the two RecA domains of Prp22 consistent with biochemical footprinting data (Schwer, 2008). The S635A mutation used to stall P complex is also visible; however, the origin of defective mRNA release due to this mutation is not apparent (Schwer and Meszaros, 2000). The loop containing S635A is in a similar conformation in

this structure and in the structure of WT human Prp22 (Felisberto-Rodrigues et al., 2019) (**Fig. 5.8**). Part of the extensive N-terminal domain of Prp22 is resolved and constitutes a long helix that bridges between the Prp22 RecA2 domain and the C-terminal domain of Cwc22; this helix was unassigned in previous P-complex structures (**Fig. 5.9B**).

To improve the resolution of the peripheral, flexible regions, we used a new data-driven regularization strategy (Kimanius et al., 2023) that allows focused refinements of much smaller domains than previously possible (**Fig. 5.6, 5.8**). Combined with AlphaFold2-assisted modelling, we were able to obtain more accurate models for the U2 snRNP, NTC, U5 Sm ring, and Cwc22 N-terminal domains. Our model for the NTC within the P complex is similar to our previous model for the NTC within C and C^{i} complexes (Wilkinson et al., 2021) (**Fig. 5.9C**). Notably however, we observed a star-shaped density coordinated by several conserved lysines from Clf1 and Ntc20, which we were able to model as inositol hexakisphosphate (IP_6) (**Fig. 5.9D**). The same density, at lower resolution, is visible in the C complex (Wilkinson et al., 2021). A distinct, separate IP_6 molecule had previously been observed within all catalytic spliceosomes between the B^{act} and P-complex stages coordinated by Prp8 and (for C* and P-complexes) Slu7 (**Fig. 5.8**). We do not believe that the NTC-bound IP_6 is essential for splicing since deletion of the Clf1 tetratricopeptide repeat that coordinates IP_6 does not result in growth or splicing phenotypes in yeast and Ntc20 is a non-essential splicing factor (Chen et al, 2001; Chung et al., 1999). Nonetheless, it is possible that IP_6 may regulate splicing in some manner via interactions with both Prp8 and the NTC.

We were able to unambiguously assign three, connected long a-helices to Fyv6 (**Fig. 5.9A**). These helices were visible in previous cryo-EM reconstructions of yeast C* and P complexes but were either unassigned or misassigned. In the cases of misassigned densities, Fyv6 was previously attributed to the C-terminal domain of the 1^{st} step factor Yju2 (Wilkinson et al., 2017). Yju2 also has an elongated helical architecture and adopts a similar position on C complex (Wilkinson et al., 2021) and the post-P complex Intron Lariat Spliceosome (ILS) (Wan et al., 2017). We closely inspected the density from previously determined spliceosome structures and concluded that the Yju2 C-terminus is still a better fit than Fyv6 for densities in B* and C complexes, whereas Fyv6 is a better fit than Yju2 for densities in C* and P complexes (**Fig. 5.10**). Therefore, this position of Fyv6 seems to be characteristic of a 2^{nd} step conformation of the spliceosome and consistent with the position of its distant human homolog FAM192A in the human pre-C* complex (Zhan et al., 2022).

Fyv6 makes a multitude of interactions with essential splicing factors (**Fig. 5.4B; 5.9E, F, G**). One conserved region of Fyv6 is the N-terminus with a FVSE motif (**Fig. 5.11**). This motif forms a ‘hook’ that interacts with a hydrophobic patch on the surface of the Prp22 RecA2 domain, where it also forms hydrogen bonds by b-sheet augmentation (**Fig. 5.9E**). The hook is followed by three a-helices. The start of helix 2 sits on top of the intron between the BP and the active site docked 3' SS, where it may act as a steric block to

prevent 3' SS undocking and promote exon ligation (**Fig. 5.9F**) (Liu et al., 2017). Helix 2 also contains a patch of conserved arginines that form salt bridges with Asp317 and Glu318 of Slu7 (**Fig. 5.9F**). The linker between helices 2 and 3 interacts with the NTC component Cef1, and helix 3 interacts with NTC component Syf1 (**Fig. 5.9G**). The binding interface with Syf1 overlaps with the Syf1/Yju2 interaction in the C complex (Wilkinson et al., 2021), suggesting these interactions are mutually exclusive (**Fig. 5.9G**). Fyv6 thus provides a direct link between the Prp22 ATPase, NTC components, and other essential splicing factors over a distance of ~100 Å.

By 3D classification, we were able to resolve a total of three states of the P-complex spliceosome (**Fig. 5.4C-E**). State I, represented by 51% of particles, is the most complete and has all the above-described factors visible, including a stably-docked 3' SS. State II, represented by 20% of particles, is similar to State I but lacks density for the N-terminal half of the 2nd step factor Slu7. The endonuclease-like domain of Prp8, to which this domain of Slu7 binds, is correspondingly shifted. Interestingly, this state retains strong density for the 3' SS, suggesting the N-terminal regions of Slu7 are not involved in maintaining 3' SS docking after exon ligation. Finally, State III, containing 29% of particles, entirely lacks density for the 2nd step factors Slu7, Prp18, and Fyv6. Prp17 and the RNase H domain of Prp8 are still present but are weaker than in States I and II, suggesting they are more flexible in State III. Most of Prp22 can still be observed with weaker density except for the C-terminal domain that interacts with Prp8, which lacks density entirely. Importantly, State III also lacks density for the 3' SS but has density for the 3' exon, suggesting it is indeed post-catalytic and represents a state after loss of 2nd step factors and undocking of the 3' SS from the catalytic core. The coincidence of loss of Fyv6 with loss of Prp18 and Slu7 reinforces the notion that these three factors act together and are important for 3' SS docking and stabilization of Prp17, Prp22, and the RNase H domain of Prp8.

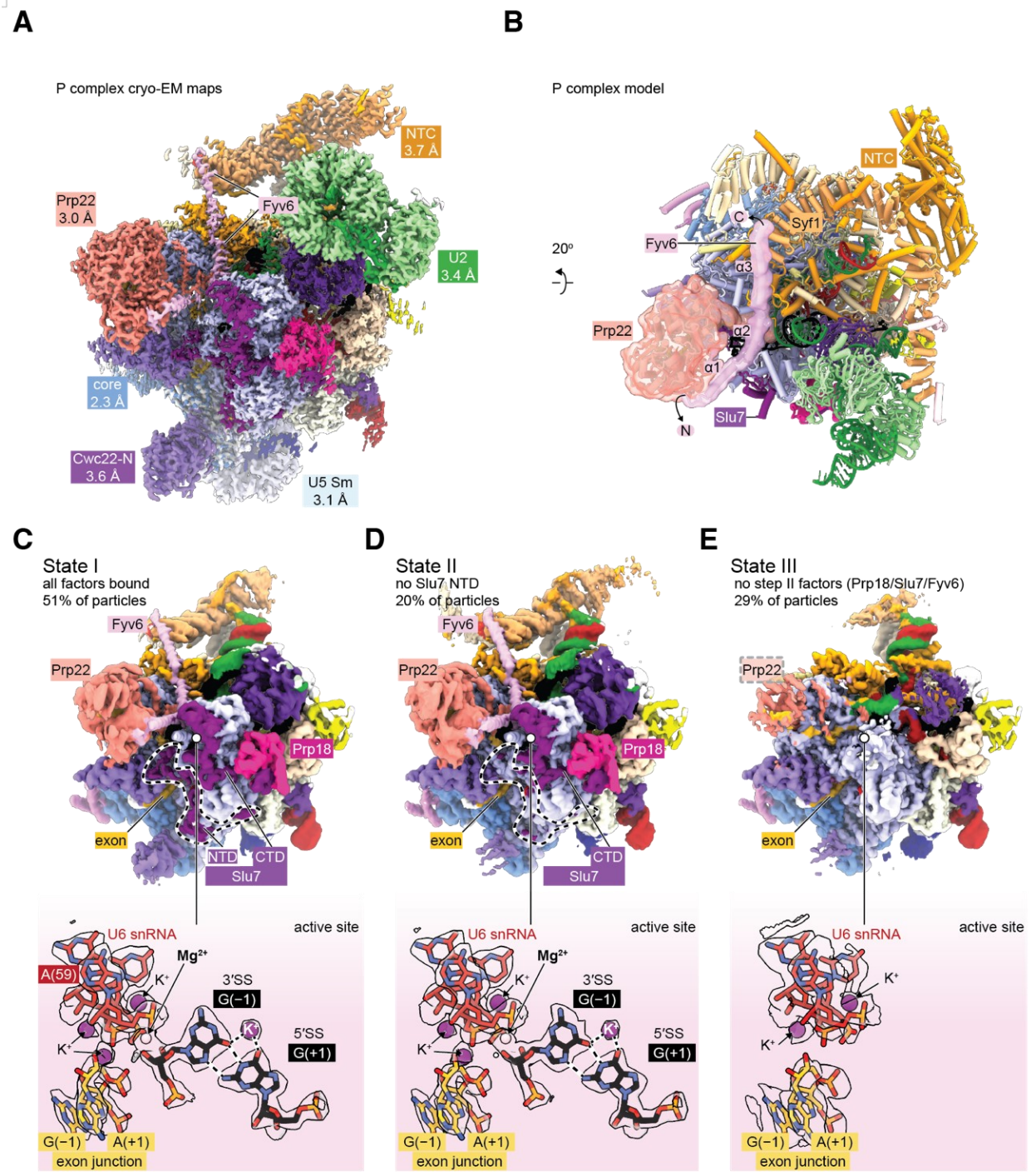

Figure 5.4 Cryo-EM structure of the yeast P complex spliceosome at 2.3 Å resolution. A) A composite density map for P complex showing focused refinements of the Prp22, NTC, U2 snRNP, U5 Sm ring, and Cwc22 N-terminal domain regions. **B)** Overall model for the P complex spliceosome. **C)** Cryo-EM density for state I of P complex (above, low-pass filtered) and for the active site and 3' SS (below, sharpened). **D,E**) As for **C)** but for states II (**D**) and III (**E**).

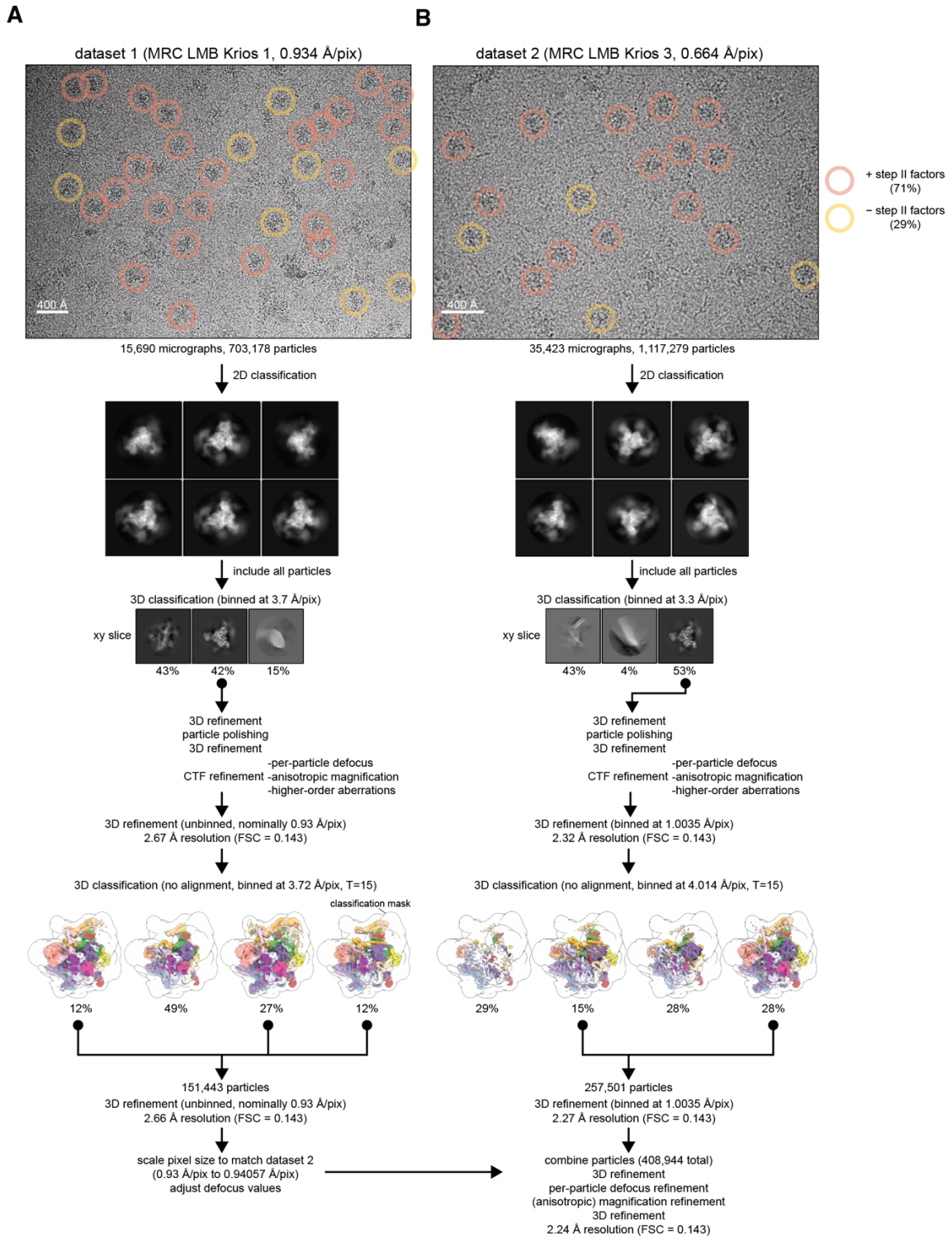

Figure 5.5 Spliceosome cryoEM data collection and general processing. **A)** Example micrograph for dataset 1 with particles in states I or II circled in red and particles in state III circled in yellow. Example 2D classes are shown below, but 2D classification was not used for selection of particles for further processing. **B)** as **A)** but for dataset 2.

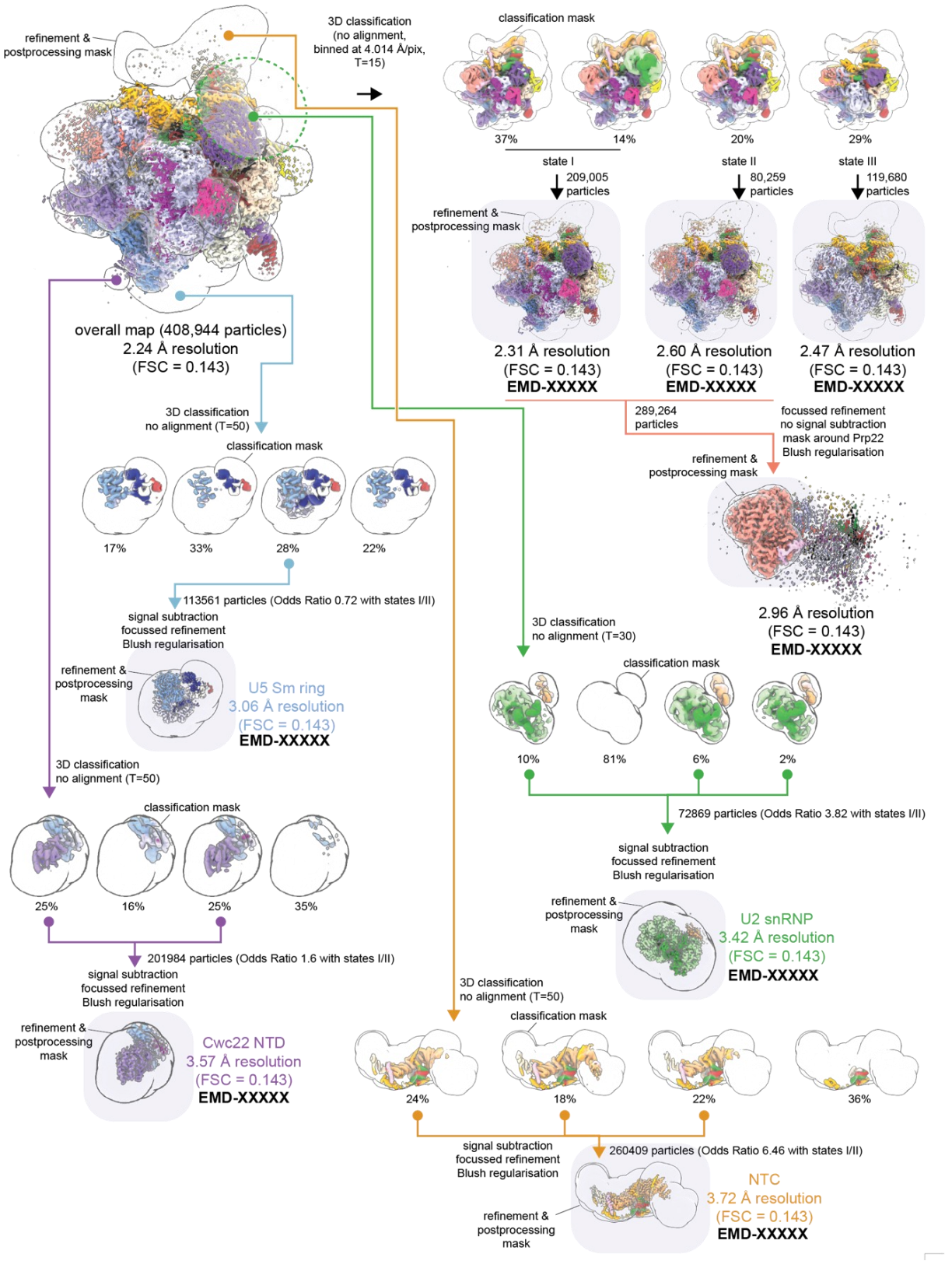

Figure 5.6 Focused classification and refinement scheme for regions of P complex. Final maps deposited to the EMDB are highlighted.

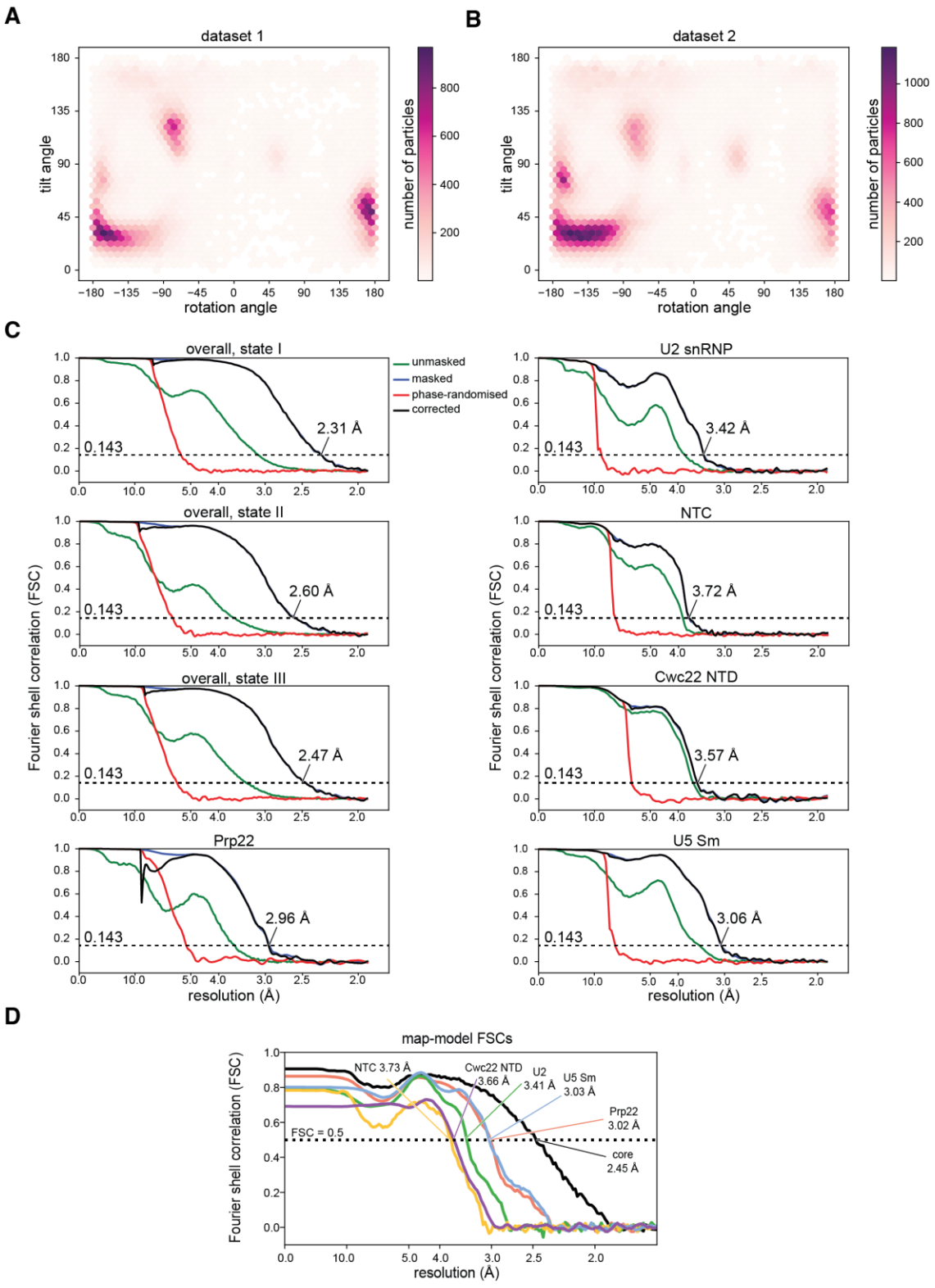

Figure 5.7 Statistics for cryo-EM dataset and map A,B) Orientation distribution plot for state I separated by datasets. **C)** Gold-standard Fourier-Shell Correlation curves for the overall reconstructions and focus-refined maps. **D)** Map-model Fourier-Shell Correlation, calculated using PHENIX.

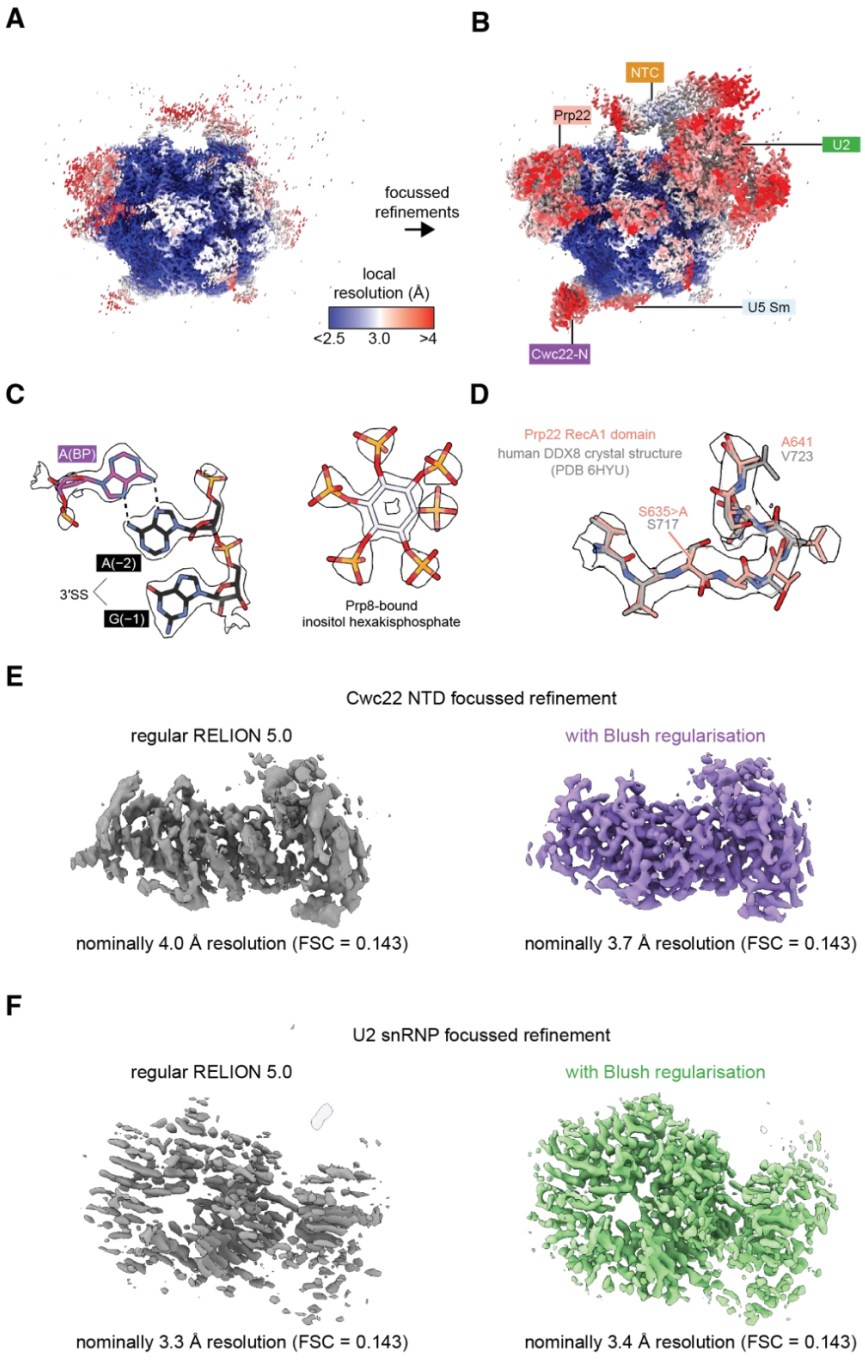

Figure 5.8 Representative cryo-EM densities. **A)** The overall P complex reconstruction (state I) colored by local resolution as calculated within RELION. **B)** Local resolution after focused refinements, colored as in **A**). **C)** Representative density for the P complex core. **D)** Density for the loop containing the S635A dominant-negative mutation in Prp22. The model is superimposed over the crystal structure of DDX8 (human Prp22) in the same region (PDB ID 6HYU; Felisberto-Rodrigues et al., 2019). **E)** Focused refinement of the Cwc22 NTD without and with Blush regularization. All refinement settings and inputs were the same except for the usage of Blush. Only dataset 2 particles were used for this comparison. **F)** As for **E)** but with the U2 snRNP. Note that the resolution estimate is inflated without Blush regularization.

Table 5.3 Cryo-EM data processing, refinement, and validation statistics

	state I	state II	state III	Prp22	U2	NTC	U5	Cwc22N
EM data collection								
Microscope model	Thermo Scientific Titan Krios cryo TEM							
Detector model	Gatan K3							
Number of datasets	2							
Number of micrographs	51113							
Frames per micrograph	40							
Data collection software	EPU							
Voltage (keV)	300							
Electron fluence (e-/Å2)	40							
Magnification	81000 / 130000							
Pixel size (Å)	0.934 / 0.664							
Defocus range (μm)	-1.3 to -3.1							
Energy filter slit width	20 eV							
EMPIAR dataset								
3D reconstructions								
EMDB map entry	NNNNN	NNNNN	NNNNN	NNNNN	NNNNN	NNNNN	NNNNN	NNNNN
PDB coordinate entry	NXYZ							
Data processing								
Initial particle number	1820457							
Final refined particle number	209,005	80,259	119,680	289,264	72,869	260,409	113,561	201,984
Map resolutions (Å):								
Masked, at FSC = 0.143	2.31	2.6	2.47	2.96	3.42	3.72	3.06	3.57
Masked, at FSC = 0.5	270	3.02	2.9	3.32	3.82	3.92	3.38	3.87
Unmasked, at FSC = 0.143	3.14	3.6	3.34	3.6	3.82	3.92	3.5	3.64
Unmasked, at FSC = 0.5	3.94	7.38	4.34	4.3	4.4	4.47	4.15	3.95
Map resolution range[a] (Å)	2.23 – 4.56	2.47 –5.77	2.36 – 6.05	2.95 – 3.35	3.22 – 6.2	3.51 – 10.4	2.96 – 6.23	3.42 – 5.14
Map sharpening *B* factor (Å^2)	-29.1	-28.8	-32.6	-62.6	-66	-121.7	-58	-170.3
Refinement and model statistics								
Model resolution (FSC=0.5) (Å)	2.5			3	3.4	3.7	3	3.7
Map CC (around atoms)	0.81			0.8	0.79	0.71	0.78	0.75
EMRinger score	3.86			3.54	2.65	1.61	4.09	2.95
Model composition								
Non-hydrogen atoms	52145			5972	10642	11883	5035	2125
Protein residues	5607			725	963	1417	602	261

Nucleotides	298	10	136	3	13	0
Ligands	15	0	0	1	0	0
Refinement method		phenix.real_space_refine				
Mean *B* factors (Å^2)						
Protein	33.6	59.88	47	67.4	64.3	23.2
Nucleotides	53.3	94.1	149.5	274	61.1	
Ligand	21.7			85.3		
R.m.s. deviations						
Bond lengths (Å)	0.006	0.005	0.005	0.005	0.005	0.004
Bond angles (°)	1.332	0.854	0.933	0.895	0.929	0.8
Validation						
MolProbity score	2.53	1.86	2.08	1.76	2.11	2.01
All-atom clashscore	5.71	2.94	4.92	3.03	5.03	3.49
Rotamer outliers (%)	10.16	5	10.2	7.16	7.96	6.94
CaBLAM outliers (%)	2.42	1.87	0.55	1.11	1.77	1.17
C-beta deviations (%)	0	0	0	0	0	0
Nucleic acid geometry						
Correct sugar puckers (%)	96.6%	100%	100%	100%	100%	
Good backbone conformations (%)		50%	72.1%	100%	54%	
Ramachandran plot						
Favoured (%)	92.7	96.4	97.76	98.92	97.1	96.5
Allowed (%)	5.6	3.35	2.24	2.01	2.91	3.47
Disallowed (%)	1.54	0.28	0	0.07	0	0

[a] Range from 5th to 95th percentiles of local resolution map within the refinement mask

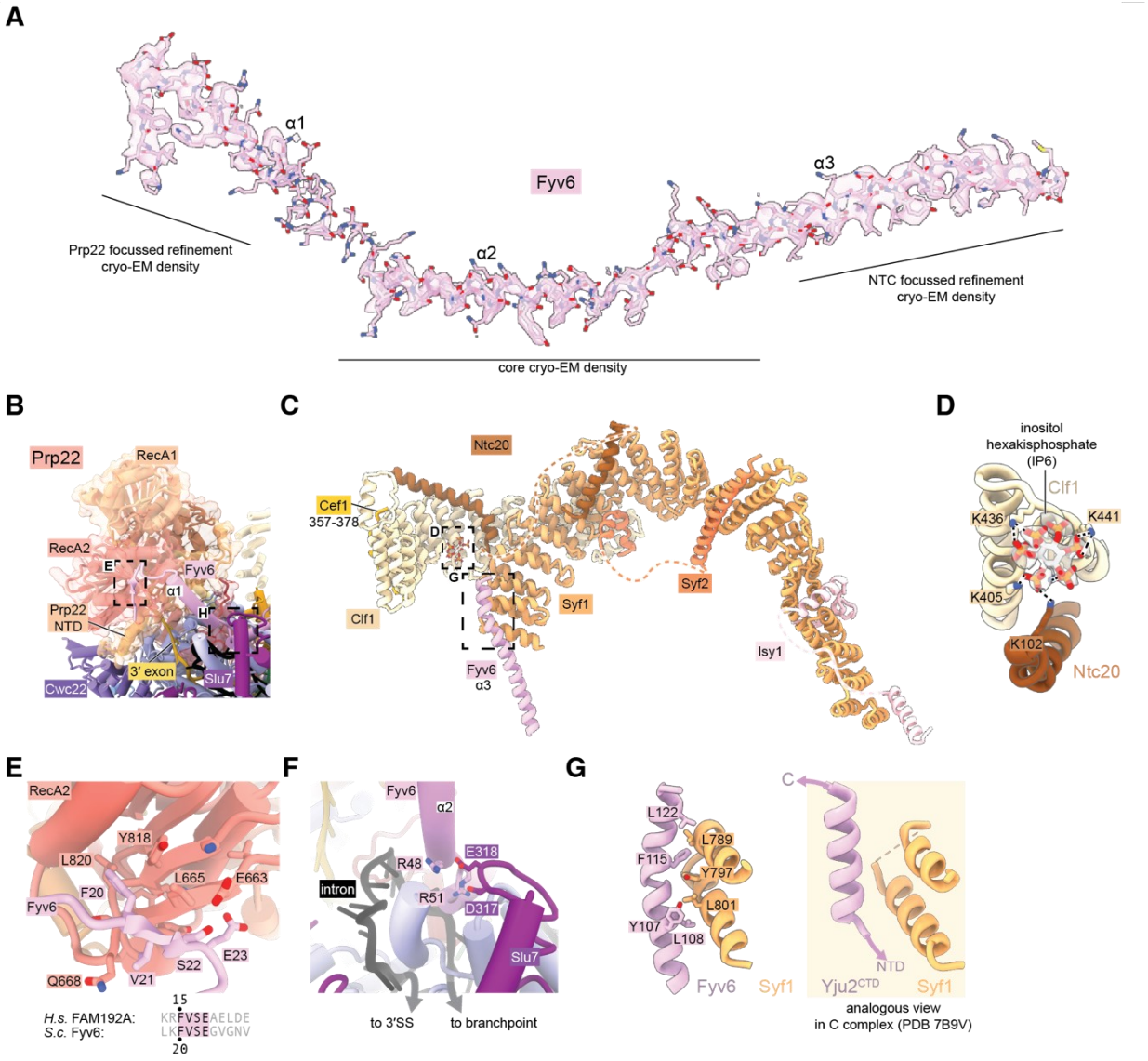

Figure 5.9 Fyv6 and its interactors in the P complex cryo-EM structure. A) Cryo-EM density segmented around Fyv6. **B)** Structure of Prp22. **C)** Structure of the NTC within P complex. **D)** IP_6 site and cryo-EM density in NTC. **E)** Interaction of the hook domain of Fyv6 with the Prp22 RecA2 domain. **F)** Interaction of Fyv6 with Slu7 and the region of the intron between the BP and 3' SS. **G)** Interaction of Fyv6 with Syf1 and an analogous view of the interaction of Syf1 with Yju2 in C complex (Wilkinson et al., 2021).

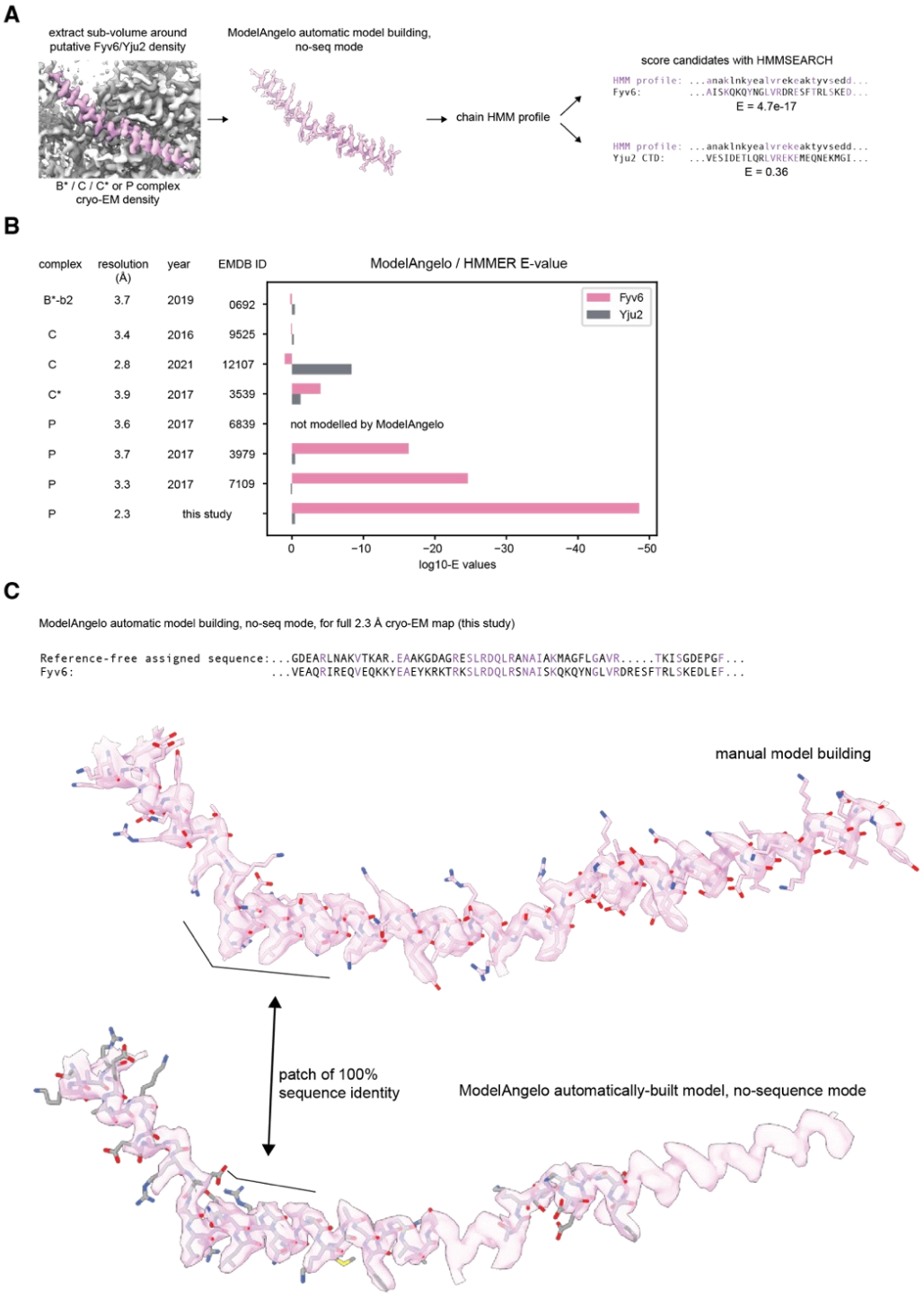

Figure 5.10 Comparison of fits to density of Yju2 versus Fyv6 in each complex. A) Cryo-EM densities were boxed around Fyv6 or Yju2 using ChimeraX. These were used as inputs to ModelAngelo operated in "no-seq" mode. The HMM profiles for the chains built into the putative Fyv6/Yju2 density were used to score all spliceosomal proteins using hmmsearch. **B)** Expectation values (E-values) for hmmsearch against Fyv6 or Yju2 protein sequences. Lower E-values indicate a better match to the profile. **C)** Comparison of the model built by ModelAngelo in no-seq mode to the manually built model for Fyv6 in P complex State I (this study). Despite not providing the sequence for Fyv6, several patches had 100% sequence identity.

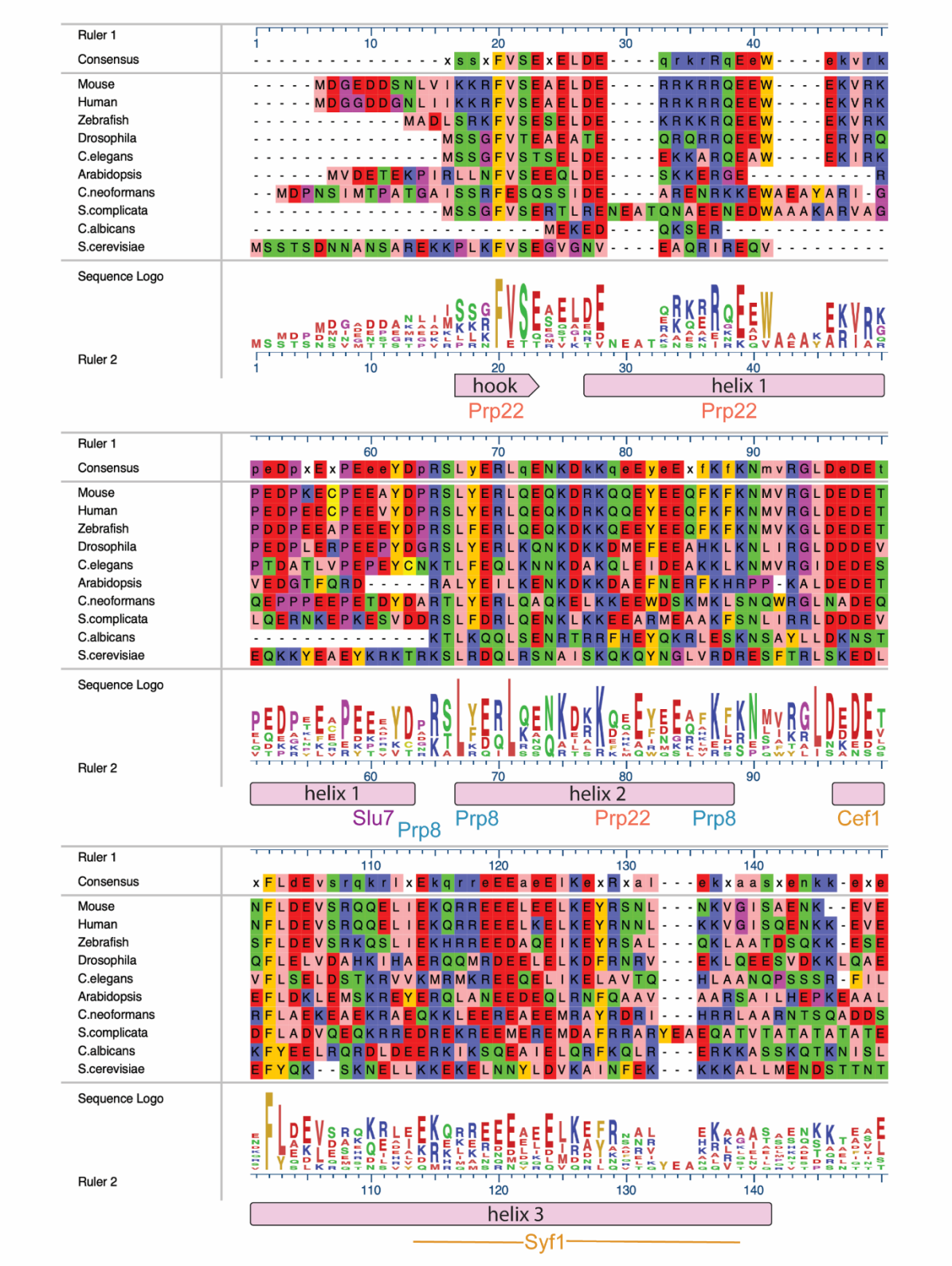

Figure 5.11 (continued next page)

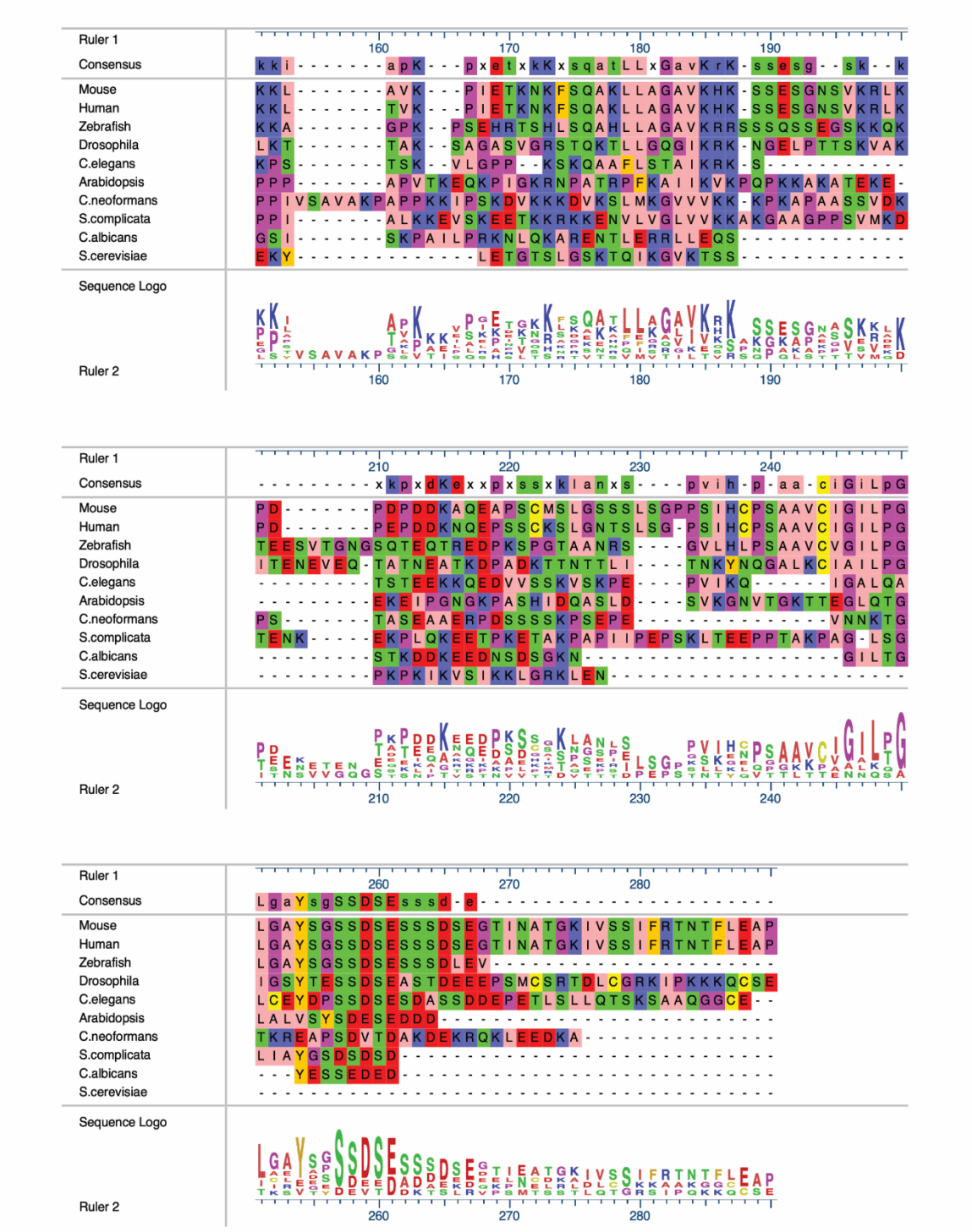

Figure 5.11 Alignments of Fyv6 homolog protein sequences. Sequence-based alignment of Fyv6 and homologs from other model organisms. Alignment, consensus, and sequence logos are shown. Regions of Fyv6 protein structure determined from cryo-EM model are annotated below corresponding sequence. Sequences were aligned and visualized using MegAlign Pro (DNAStar) and Clustal Omega.

The Fyv6/Syf1 Interaction is Critical for Suppressing *fyv6Δ* Phenotypes

To determine the functionally critical regions and interactions of Fyv6, we used the cryo-EM structure to design five truncation mutants (**Fig. 5.12A**). The first three truncations remove amino acids from the N-terminus of Fyv6: deletion of the first 16 amino acids, which are not resolved in the cryo-EM structure (Δ1-16); deletion of the first 23 amino acids, which removes the conserved hook region of Fyv6 that interacts with the Prp22 RecA2 domain (Δ1-23); and deletion of amino acids 1 to 51, which additionally removes the first a-helix and interactions with Prp22, Prp8, and Slu7 (Δ1-51). The remaining two mutants are C-terminal truncations. In the first, amino acids 134 to 173, which do not have clear density in the cryo-EM structure and are not modeled but which do contain a predicted nuclear localization signal (NLS), are deleted (Δ134-173) (Nguyen Ba et al., 2009). In the second, amino acids 103 to 173 are deleted (Δ103-173), which additionally removes the region of the protein that interacts with Syf1.

We tested each truncation mutant for its ability to suppress the *cs* and *ts* phenotypes of the *fyv6Δ* strain when expressed from a plasmid (**Fig. 5.12B**). All of the N-terminal truncation mutations could suppress the *cs* or *ts* phenotypes to some degree. Surprisingly, a strain with the complete deletion of the conserved hook domain that interacts with Prp22 (Δ1-23) grew similarly to the strain expressing WT Fyv6. We wondered if splicing changes could still be present even if the growth phenotype was suppressed; so, we assayed *SUS1* isoform generation by RT-PCR. We previously showed that loss of Fyv6 causes use of an alternative 3' SS in the first intron of this pre-mRNA (Lipinski et al., 2023). The N-terminal truncation mutants also showed no evidence for changes in *SUS1* splicing (**Fig. 5.12C**). The N-terminal region of Fyv6 composed of the hook domain and the first a helix is not required for suppression of either the growth or splicing phenotypes observed when Fyv6 is lost despite its conservation and interactions with critical splicing factors (Prp8, Prp22, and Slu7).

For the C-terminal truncations, the Δ134-173 truncation was also able to suppress the *cs* and *ts* phenotypes, indicating that the predicted NLS is also not essential (**Fig. 5.12B**). However, this mutant did grow more poorly at 16°C than did the Δ1-16 and Δ1-23 truncations and did show some evidence of alternative 3' SS usage in *SUS1* (**Fig. 5.12C**). The largest effects were seen when the C-terminus was truncated further. The Δ103-173 truncation strain phenocopied the *fyv6Δ* strain and grew poorly at 37°C and was dead at 23° or 16°C (**Fig. 5.12B**). Additionally, RT-PCR showed increased use of the alternative 3' SS in *SUS1* with this truncation (**Fig. 5.12C**). To determine if this phenotype could be due to loss of protein expression, we assayed expression of the epitope-tagged proteins by western blot (**Fig. 5.13A**). In all cases, we could detect protein; however, the abundance of the N-terminally epitope-tagged Δ103-173 variant was much lower than the others. Therefore, we also constructed a C-terminally epitope-tagged version of Δ103-173 (**Fig. 5.13B**). The C-terminally-tagged protein expressed at much higher levels but also failed to suppress the *cs* and *ts* phenotypes. These data suggest

that the C-terminal region of Fyv6 that includes a portion of a helix 3 and the Syf1 interaction domain is critical for function.

To obtain additional evidence for the importance of the Fyv6/Syf1 interaction, we created a *syf1Δ* shuffle strain containing WT Fyv6 and expressed C-terminal Syf1 truncations (**Fig. 5.12D**). We hypothesized that loss of the Fyv6-interacting domain on Syf1 could phenocopy *fyv6Δ*. We observed just this effect. Loss of the Fyv6 interacting region in Syf1 Δ778-859 and Δ634-859 mutants resulted in either modest (Δ778-859) or strong (Δ634-859) *cs* and *ts* phenotypes. In addition, these Syf1 truncation mutants also triggered use of the alternative 3' SS in *SUS1* (**Fig. 5.12F**). Interestingly, if the Δ778-859 and Δ817-859 Syf1 mutants are expressed in a *fyv6Δ* background, it results in suppression of the *fyv6Δ ts* phenotype (**Fig. 5.12G; *c.f.* Fig. 5.12B, vector**). However, the *SUS1* alternative 3' SS is still used (**Fig. 5.12H**). This indicates that *ts* phenotype suppression and triggering usage of the alternative 3' SS in *SUS1* are separable processes.

As mentioned above, the Syf1/Fyv6 interaction is mutually exclusive with the Syf1/Yju2 interaction that occurs during 5' SS cleavage and is reformed in the ILS spliceosome. One possible explanation for suppression of the *ts* phenotype in *fyv6Δ*/*syf1*(Δ778-859 or Δ817-859) strains is that Fyv6 and Yju2 are competing with one another and binding of either Fyv6 or Yju2 to the 1st or 2nd step spliceosome complex, respectively, results in growth and splicing defects. When the Yju2/Fyv6 binding site on Syf1 is removed, the observed phenotypes could be due to failure of Yju2 to be properly recruited to the P complex spliceosome for displacement of Fyv6. However, recruitment occurs correctly when Fyv6 is deleted since it no longer needs to be displaced. Similarly, phenotypes due to loss of Fyv6 could arise from retention of Yju2 during exon ligation, which might stabilize first step conformations and result in cold sensitivity. A prediction of this model is that failure to properly remove or recruit Yju2 and Fyv6 can result in similar phenotypes. Further work will be needed to understand the binding dynamics of these factors.

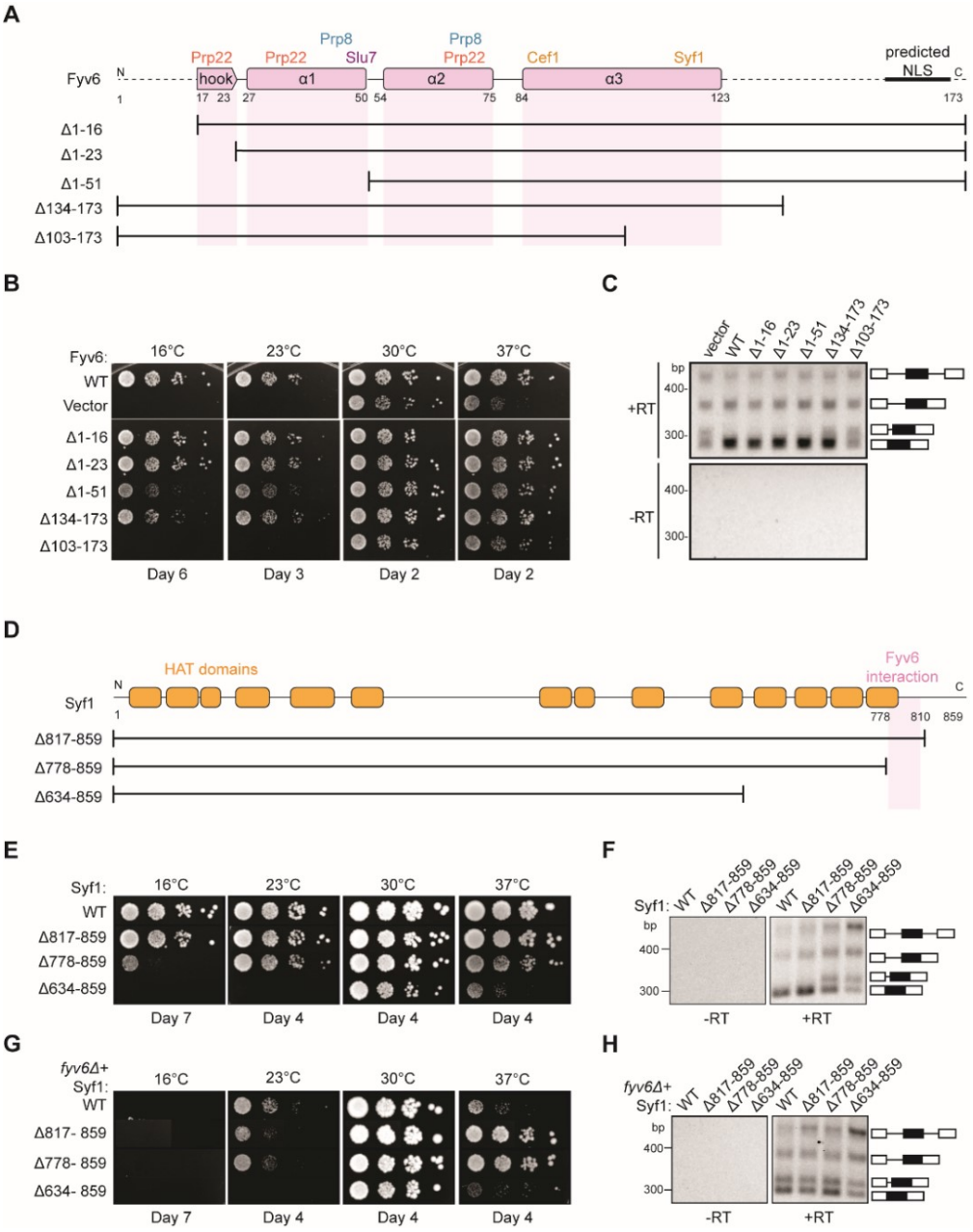

Figure 5.12 Structure-based analysis of Fyv6 domain function. A) Diagram of Fyv6 protein structure, protein interactors, and protein truncations. NLS = nuclear localization signal **B)** Spot dilution assay of strains with Fyv6 truncations on –Trp DO plates at different temperatures. Plates were imaged after the number of days indicated. **C)** Representative gel image of *SUS1* RT-PCR in strains with an empty vector or expressing the indicated Fyv6 variant. (+RT reactions contain reverse transcriptase; -RT control reactions do not contain reverse transcriptase.) **D)** Diagram of Syf1 protein structure and truncations. **E)** Spot dilution assay of strains with Syf1 truncations on –Trp DO plates at different temperatures. Plates were imaged after the number of days indicated. **F)** Representative gel image of *SUS1* RT-PCR in strains with an empty vector or expressing the indicated Syf1 variant. **G)** Temperature growth assay of strains with Syf1 truncations in a *fyv6Δ* background on YPD plates. Plates were imaged after the number of days indicated. **H)** Representative gel image of RT-PCR of *SUS1* in strains with plasmids containing WT Syf1 or a Syf1 truncation in a *fyv6Δ* background.

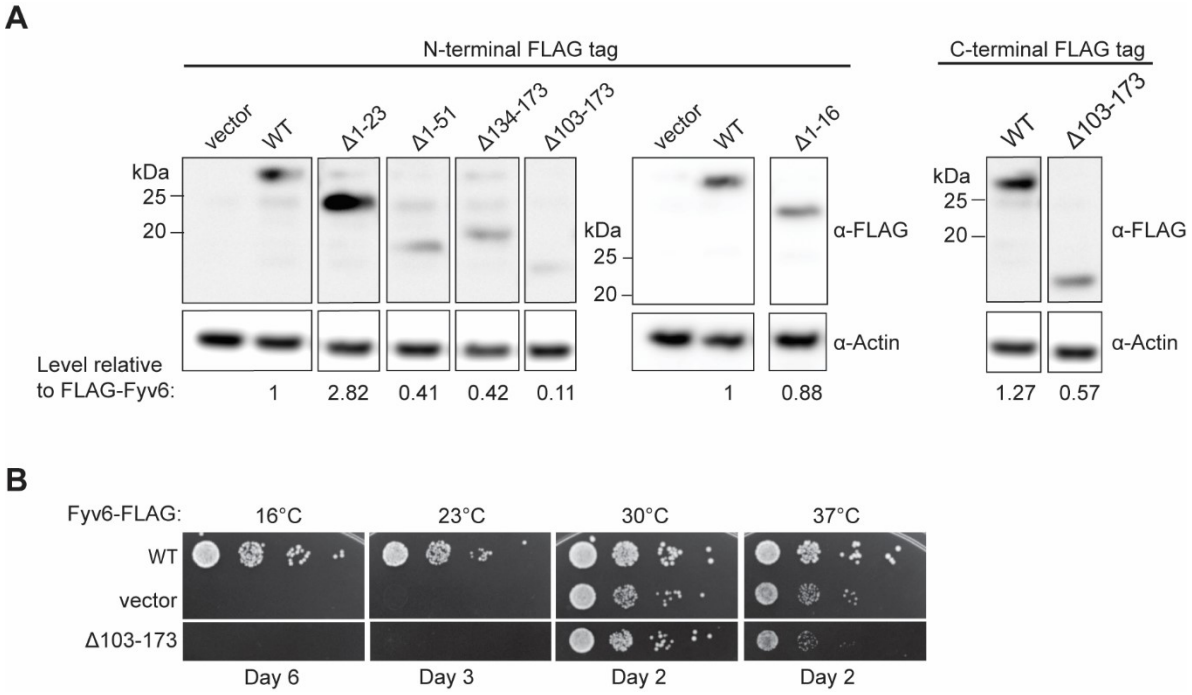

Figure 5.13 Expression of epitope-tagged Fyv6 variants. A) Expression of FLAG-tagged Fyv6 truncations visualized by western blot using anti-FLAG antibody. Actin was used as a loading control. **B)** Temperature growth assay of C-terminally FLAG-tagged WT Fyv6 and the Δ103-173 truncation on –Trp plates. Plates were imaged after days indicated.

Mutations in Many Different Splicing Factors Can Suppress *fyv6Δ* Phenotypes

To gain additional insights into Fyv6 function, we carried out a genetic screen to detect spontaneously-arising suppressors that can correct temperature-dependent growth defects (**Fig. 5.14A**). This method has previously been used to illuminate many aspects of spliceosome biochemistry and non-silent mutations within splicing factors are likely causative for suppression (Montemayor et al. 2014; Kuhn and Brow, 2000; Brow, 2019).

We plated saturated *fyv6Δ* yeast cultures onto 20 different YPD plates and incubated them at 37°C until colonies arose. Within several days, colonies of various sizes appeared and were isolated. A subset of picked colonies were then tested for suppression of the *ts* phenotype. All of these colonies were able to suppress growth defects at 37°C, and many also suppressed the *cs* phenotype at 23°C (**Fig. 5.15**). We then performed whole genome sequencing on the largest of the picked colonies as these were expected to be the strongest suppressors. In 19 of 20 sequenced strains, we were able to identify at least one non-silent mutation in a splicing factor (**Fig. 5.14B**, **Table 5.4**). When we carried out a similar experiment aimed at identifying *cs* suppressors, we were unable to identify any likely causative mutations in splicing factors. Therefore, we focused our analysis on the *ts* suppressor mutations.

Table 5.4 Mutations identified in *fyv6Δ* suppressor strains selected at 37°C

Gene	Chromosome	Position	Substitution	Mutation	Strain(s)	Human Gene	Human Residue
PRP8	VIII	432198	G->T	S1584Y	370201		S1512
		432198	G->A	S1584F*^	372003		S1512
		431365	C->G	V1862L	370801		I1790
		431347	C->T	G1868R	370701		G1796
		431004	G->C	T1982S^	370501		T1910
SLU7	IV	619621	C->T	E9K	370301	SLU7	N/A
		619576	C->T	A24R**	370101		N/A
		619575	G->C				
		619570	C->T	E26K**			N/A
		619562	A->C	N28K**			D58
CDC40 (PRP17)	IV	1203588	G->A	P208L	371102	CDC40	A245
		1203459	G->A	A251V	370601		Y367
		1203266	C->A	K315N	371701		K450
		1203055	C->A	G386W	370901		A511
Cef1	XIII	693490	C->T	A37V	370401	CDC5L	A35
		693905	G->T	M175I	371201, 371302		Q175
		693958	A->C	Q193P	370804		A193
CLF1 (SYF3)	XII	384437	G->A	T33I^^	371001	CRNKL1	P210
RSE1	XIII	175909	A->T	D799E*, ***,****	371601, 371901, 372003	SF3B3	N/A
		175912	G->T	D798E***	371901		N/A
		175899	T->C	K803E***			N/A
		175896	C->T	E804K***			N/A

		175883	A->T	I808K***			N/A
LSR1	II	681838	T->C	A25G	371802	RNU2	A24
PRP22	V	182238	T->G	I1133R**	371601	DHX8	N/A

*Mutations arose together in strain 372003. **Mutations arose together in strain 370101. ***Mutations arose together in strain 371901. ****Mutations arose together in 371601. ^Second-strongest suppressors of *FYV6* deletion based on colony size. ^^Strongest suppressor of *FYV6* deletion based on colony size.

Collectively, the identified mutations tended to fall into two categories: gain of positive charge or substitution with a bulkier side chain. Gain of positive charge mutations were found in Prp8 (G1868R) as well as the second step factors Slu7 (A24R, and N28K), and Prp22 (I1133R). Structurally, amino acids within Prp8 and Prp22 are proximal to RNA within the spliceosome and addition of charged amino acid side chains may stabilize interactions with the phosphodiester RNA backbone. The N-terminus of Slu7 containing A24 and N28 is not resolved in any spliceosome structure to our knowledge and normally contains many negatively charged residues. The majority of identified suppressors involve a mutation to a bulkier side chain including several in Prp8 (S1594F/Y, V1862L), Prp17 (A251V, G386W), Cef1 (A37V), and Clf1 (T33I). The increase in hydrophobic surface area provided by these mutations may help to stabilize interactions with nearby proteins, compensating for structural destabilization due to the absence of Fyv6.

We isolated one suppressor mutation within the U2 snRNA (A25G) near where the C-terminal domain of Prp22 inserts towards the active site and located beneath the Fyv6 binding site. It has previously been observed that mutation of this specific U2 snRNA nucleotide is not deleterious to growth (McPheeters and Abelson, 1992), and is therefore well-tolerated by yeast. The mechanism of suppression by this mutation is not clear; however, it is interesting to note that we also observed suppressor mutations in the U2 snRNP component Rse1 (**Table 5.4**) which is released from the spliceosome prior to catalysis. The Rse1 mutations always occurred coincident with mutations in other splicing factors. It is possible that suppressors in factors present prior to or after exon ligation are able to suppress (or enhance suppression of) *fyv6Δ* by changing how pre-mRNAs compete for a limited set of splicing factors or by reducing blocks to pre-spliceosome assembly caused by accumulation of stalled complexes (Mendoza-Ochoa et al., 2019; Munding et al., 2013). Together these data indicate that mutations in many different splicing factors including second-step factors and the U2 snRNA can be isolated from *fyv6Δ* suppressor strains.

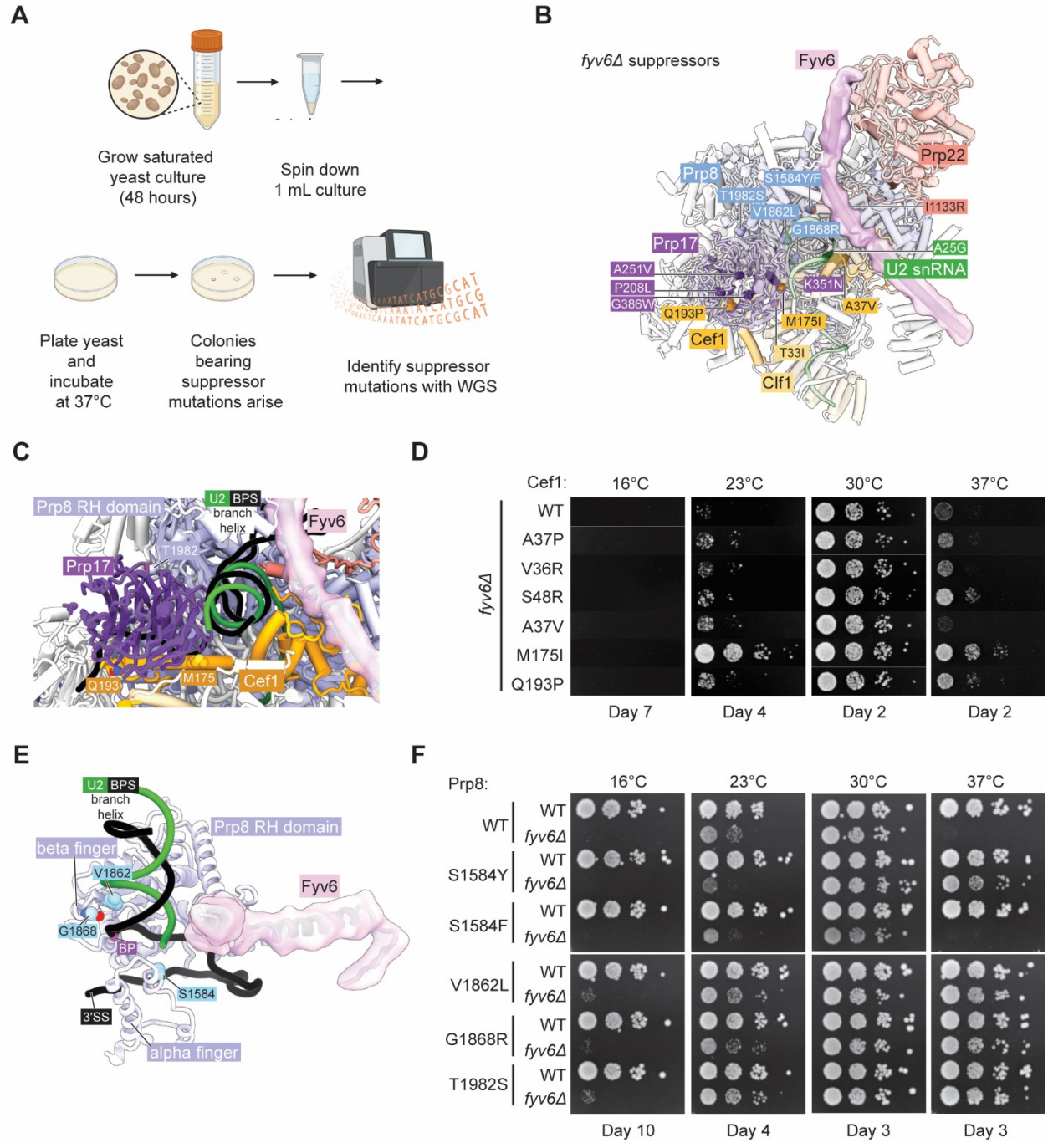

Figure 5.14 Identification of novel suppressors of *fyv6Δ* temperature sensitivity. A) Workflow for the suppressor screen. **B)** P complex spliceosome structure with *fyv6Δ* suppressor mutations labelled. (See **Table 5.4**) **C)** Close-up view of Cef1 and Prp8 suppressor mutations in P complex. **D)** Spot dilution temperature growth assay with Cef1 mutants on YPD plates. Plates were imaged after the number of days indicated. **E)** Close-up view of Prp8 mutations in P complex. **F)** Spot dilution temperature growth assay with Prp8 mutants on YPD plates. Plates were imaged after the number of days indicated.

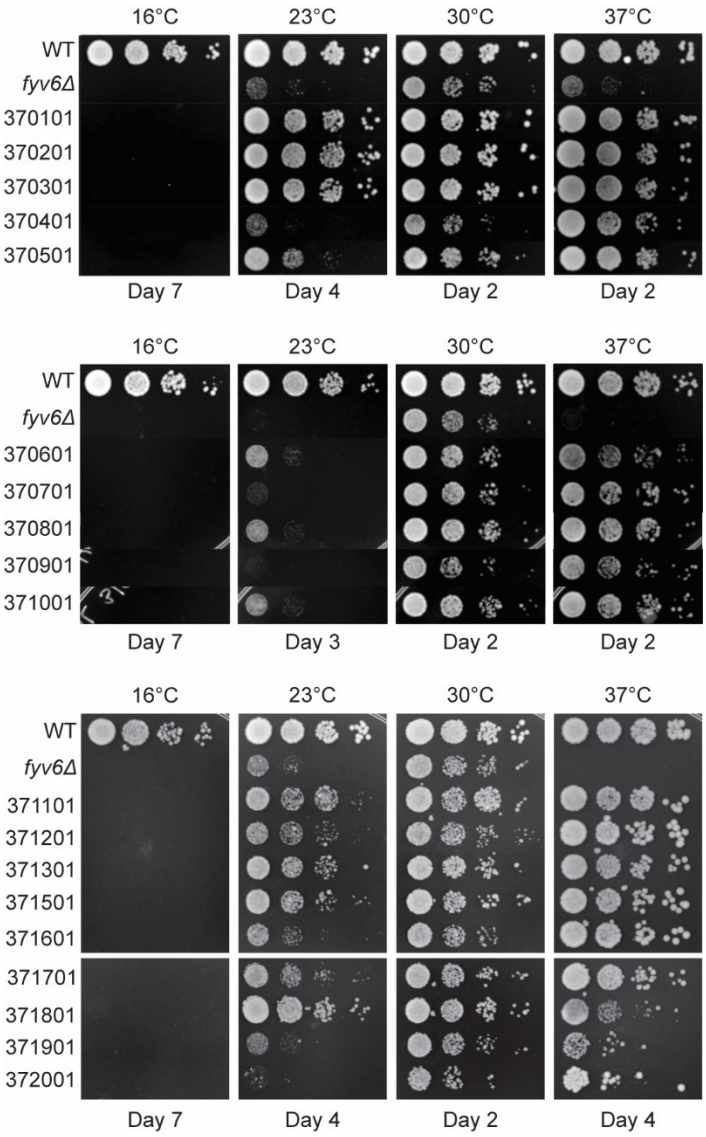

Figure 5.15. Isolated suppressor strains for *fyv6Δ* grown at different temperatures. Strains were grown on YPD plates and imaged after the number of days indicated.

Multiple Second-step Splicing Defects can be Rescued by the Same Suppressor

We decided to study several of these mutations in greater detail and to confirm that they were sufficient for *fyv6Δ ts* phenotype suppression. In the case of Cef1, we obtained several mutants located near the interface between Cef1 and the 2nd step factor Prp17: A37V, M175I, and Q193P (**Fig. 5.14C**). This interface has only been observed in structures of C* and P-complex spliceosomes, suggesting its importance for exon ligation. In addition, it has previously been reported that another mutant of Cef1 A37, A37P, is able to suppress splicing defects due to a BS adenine to guanosine substitution (BS-G) which causes stalling after 5' SS cleavage (Query and Konarska, 2012). Therefore, we wondered if other Cef1 mutants capable of suppressing defects due to BS-G could also

suppress *fyv6Δ* (Cef1 A37P, V36R, and S48R). We created *fyv6Δ* shuffle strains for Cef1 and tested each mutant's ability to suppress *cs* and *ts* phenotypes (**Fig. 5.14D**). All of these Cef1 mutants, both those isolated from our suppressor screen and those identified as BS-G suppressors, could at least weakly suppress the *cs* phenotype at 23°C with our isolated M175I mutant being a particularly strong suppressor. In addition, all of the mutants except Cef1 A37V could suppress the *ts* phenotype at 37°C to varying degrees. These results confirm that single point mutants in Cef1 can suppress *fyv6Δ* growth phenotypes and that at least some BS-G suppressors in Cef1 also suppress *fyv6Δ*. A common mechanism could be that these mutations stabilize the exon ligation conformation of the spliceosome to favor the 2nd step either when a BS-G is present or Fyv6 is absent.

In the case of Prp8, we tested five different suppressor mutations in *prp8* shuffle strains. Identification of mutants in Prp8 that suppress *fyv6Δ* is not surprising since many different alleles of Prp8 are known to modulate the conformational equilibrium between the 1st and 2nd steps. We previously discovered that one Prp8 2nd step allele, P986T, could suppress a *fyv6Δ* growth defect (Fica and Nagai, 2017; Lipinski et al., 2023; Liu et al., 2007; Query and Konarska, 2004).

We identified two suppressors within the Prp8 a-finger domain at position S1584 and another three within the Rnase H-like domain (**Fig. 5.14E, Table 5.4**) (Galej et al. 2013). S1584 is located adjacent to a residue that crosslinks with the BP (aa 1585) (Galej et al., 2013). Proximity to the BP suggests that mutations at S1584 might destabilize RNA bound within the active site after branching, alleviating the stall caused by *fyv6Δ* after the 1st step or facilitating formation of the 2nd step conformation.

Suppressors of other splicing defects have been previously identified within the Prp8 RNase H-like domain, two of which are located at the same position as *fyv6Δ* suppressors. V1862A/Y/D and T1982A were identified as suppressors of, respectively, the *U4-cs*1 allele (Kuhn and Brow, 2000; Kuhn et al., 1999) or the 5' SS U2A mutant (Siatecka et al., 1999). V1862 is located within the RNase H-like domain β-finger (**Fig. 5.14E**), while T1982 is located at the base of the β-finger (**Fig. 5.14C**). We identified *fyv6Δ* suppressors at both V1862 (V1862L) and T1982 (T1982S) as well as a third suppressor (G1868R), also located in the β-finger. Several 2nd step Prp8 alleles are also located proximal to G1868 (N1869D and V1870N) (Query and Konarska, 2004).

Four of the identified Prp8 suppressor mutations could partially suppress either or both of the *cs* and *ts fyv6Δ* phenotypes (**Fig. 5.14F**). Prp8 S1574Y, V1862L, G1868R, and T1982S suppressors were able to support growth at 37°C. Prp8 V1862L and T1982S also partially suppressed the growth defects at 16 and 23°C. Even though we isolated the Prp8 S1584F mutant in our screen and this substitution is very similar to S1584Y, we were not able observe suppression due to this mutant in these strains. Mutations within the Prp8 β-finger are able to compensate for loss of Fyv6 and these mutations are located

near known Prp8 2nd step alleles. As with Cef1 mutants, it is possible that *fyv6Δ* suppression is originating from stabilization of the 2nd step conformation.

Finally, given our observations with the Cef1 and Prp8 mutants, we wondered about the generality of 2nd step splicing factor suppression. Several dominant mutants in the 2nd step factor Prp18 were previously shown to suppress *ts* phenotypes due to various mutations in Slu7 (Aronova et al., 2007). We wondered if these Prp18 mutants would also suppress the *ts* (or *cs*) phenotype due to Fyv6 loss. Indeed, the Prp18 V191A suppressor of *slu7* alleles was able to improve growth at 37°C of *fyv6Δ* yeast (**Fig. 5.16**). Together our results from the suppressor analysis support a model in which a given 2nd step suppressor mutation may be able to restore splicing activity lost by a variety of mechanisms.

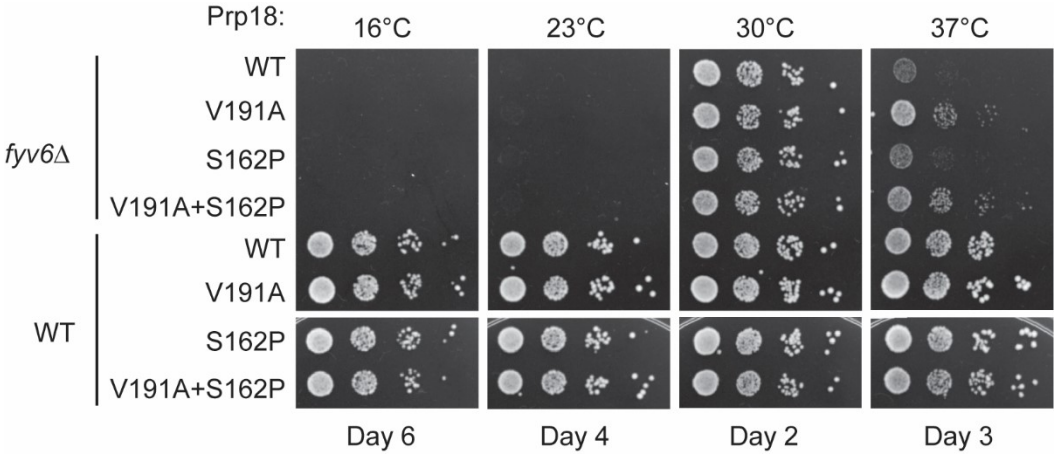

Figure 5.16 Genetic interactions between Prp18 and Fyv6. Spot dilution temperature growth assay for genetic interactions of *FYV6* deletion with Prp18 alleles on –Ura DO plates. Plate images were taken on days indicated.

A Prp22-Dependent Splicing Stall is Relieved by Fyv6 Deletion

The final *fyv6Δ* suppressor mutation that we studied was the Prp22 I1133R mutant. This mutation is located within the C-terminal domain of Prp22 that enters the spliceosome active site and is located proximal to U2 snRNA nucleotide A31 (**Fig. 5.17A**). Fyv6 bridges over this Prp22-U2 snRNA interaction. The I1133R mutation might stabilize Prp22/U2 snRNA contacts that can compensate for Fyv6 loss and promote exon ligation. The Prp22 I1133R mutant is able to weakly suppress both the *cs* and *ts* phenotypes in *fyv6Δ* yeast (**Fig. 5.17B, C**). Given its location in the C-terminal domain, we wondered if it or *fyv6Δ* would show synthetic genetic interactions or epistasis with mutations in the helicase domain. We tested two helicase domain mutations, R805A and G810A, both of which are known to impact Prp22-dependent proofreading (Mayas et al., 2006). The Prp22 R805A mutant is by itself *cs* (Schwer and Meszaros, 2000) and showed a synthetic lethal interaction with *fyv6Δ* (**Fig. 5.17B, C**). This was only slightly rescued by

simultaneously introducing I1133R into Prp22. In contrast, Prp22 G810A mutant is viable in the presence and absence of Fyv6 as well as in combination with the I1133R suppressor mutation (**Fig. 5.17B, C**). Prp22 G810A by itself does not suppress the *ts fyv6Δ* phenotype; however, the Prp22 G810A/I1133R double mutant is a stronger suppressor of the *ts* phenotype than either mutant alone. These additive genetic interactions suggest that the Prp22 *fyv6Δ* suppressor mutation (I1133R) and the helicase domain mutations (R805A and G810A) impact distinct Prp22 functions. These functions could correspond to Prp22's role in promoting exon ligation during the 2nd step (the C-terminal domain mutation) and its role in promoting ATP-dependent mRNA release and proofreading (the helicase domain mutations).

Recently, it was reported that the Prp22 G810A mutant could activate usage of some alternative 3' SS that are also activated by Prp18 deletion (Roy et al., 2023). We therefore wondered if any of the Prp22 mutants could activate alternative 3' SS also activated by Fyv6 deletion. We isolated RNA from these strains and analyzed *SUS1* splicing by RT-PCR (**Fig. 5.17D, E**). None of the Prp22 mutants activated usage of the *SUS1* alternative 3' SS or enhanced usage of this site beyond what is observed upon Fyv6 deletion. However, we noticed that the Prp22 helicase mutants, R805A and G810A, both caused an accumulation of unspliced pre-mRNA. This is consistent with these mutations causing a block during pre-mRNA splicing and having a different effect on splicing than Prp22 I1133R. Surprisingly, the block due to the G810A mutation was relieved in the *fyv6Δ* strains and little pre-mRNA accumulation was observed in the absence of Fyv6. This result suggests that Fyv6 could act as a negative regulator of splicing of the SUS1 transcript when a Prp22 helicase domain mutant is present. Removing Fyv6 restores splicing by activation of the *SUS1* BP proximal 3' SS.

To further explore the connections between Fyv6 and Prp22 fidelity, we used the ACT1-CUP1 assay in which yeast growth in the presence of increasing concentrations of Cu^{2+} is proportional to the extent of splicing of a reporter RNA (**Fig. 5.17F, G**; **Fig. 5.18, 5.19**). Since the Prp22 R805A mutant grew poorly, we only assayed the Prp22 G810A, I1133R, and G810A/I1133R (double) mutants in the presence and absence of Fyv6. Consistent with previous results, we found that *fyv6Δ* is deleterious for splicing of the WT ACT1-CUP1 reporter containing consensus splice sites. Copper tolerance was improved when *fyv6Δ* was combined with Prp22 I1133R. Unexpectedly, Prp22 I1133R was actually deleterious when Fyv6 was present for this reporter. If this Prp22 mutant and Fyv6 both promote the 2nd step, then their combination may have an additive, negative effect on splicing due to over-stabilization of this spliceosome conformation.

Finally, we assayed strains containing a 3' SS UAG to gAG reporter since it was previously shown that Prp22 helicase domain mutants could increase usage of nonconsensus 3' SS due to faulty proofreading activity (**Fig. 5.17F, G**). In agreement with this, we observed a dramatic increase in copper tolerance when Prp22 contained both the C-terminal domain and helicase mutants (double). This suggests that splicing of

nonconsensus 3' SS splice sites is enhanced both by stabilization of the 2nd step conformation (I1133R) and perturbing mRNA release (G810A). This effect on Prp22 is Fyv6 dependent since loss of Fyv6 results in very poor copper tolerance of strains containing the gAG reporter, likely since this 3' SS is located far from the BP (38 nt). These results show that *fyv6Δ* can be suppressed with a Prp22 C-terminal domain mutant and that mutations in the Prp22 C-terminal domain can have distinct functional consequences from those located in the helicase domain. In the case of *SUS1*, stalling of splicing due to the Prp22 G810A mutant can be relieved by removing Fyv6 and activation of a BP proximal 3' SS.

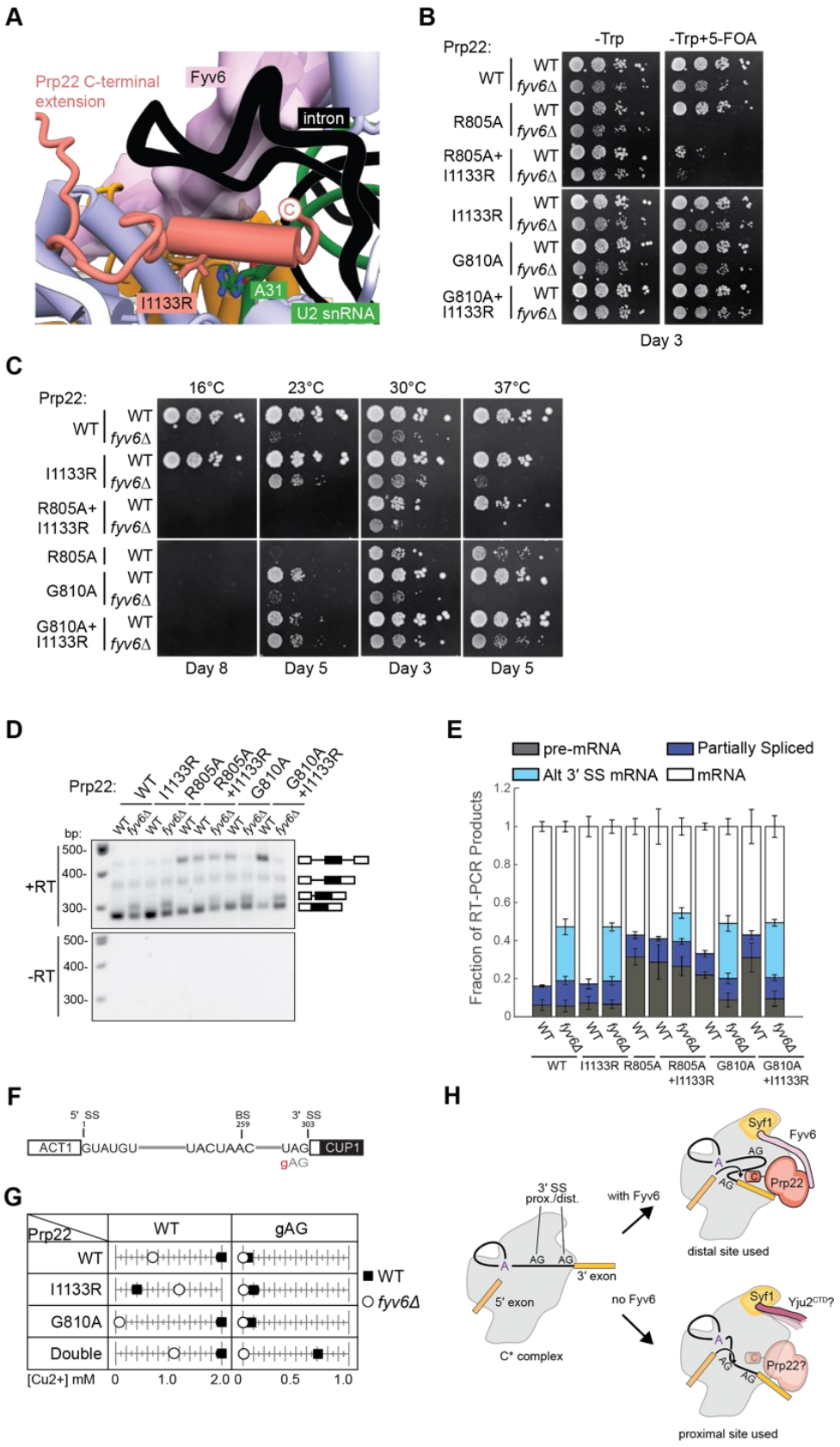
A
Prp22 C-terminal extension
Fyv6
intron
C
I1133R
A31
U2 snRNA
B
Prp22:
-Trp
-Trp+5-FOA
WT
R805A
R805A+ I1133R
I1133R
G810A
G810A+ I1133R
fyv6Δ
Day 3
C
16°C
23°C
30°C
37°C
Day 8
Day 5
Day 3
Day 5
D
Prp22:
bp:
+RT
-RT
E
pre-mRNA
Partially Spliced
Alt 3′ SS mRNA
mRNA
Fraction of RT-PCR Products
F
5′ SS
BS
3′ SS
ACT1
GUAUGU
UACUAAC
UAG
gAG
CUP1
G
Prp22
WT
gAG
I1133R
G810A
Double
[Cu2+] mM
H
3′ SS prox./dist.
with Fyv6
no Fyv6
5′ exon
3′ exon
C* complex
Syf1
Fyv6
Prp22
distal site used
Yju2CTD?
Prp22?
proximal site used

Figure 5.17 Genetic interactions between Prp22 and Fyv6. A) Location of the Prp22 I1133R *fyv6Δ* suppressor in the P complex structure. **B)** Spot dilution growth assay with Prp22 mutants in WT or *fyv6Δ* backgrounds grown on –Trp DO or –Trp+5-FOA plates. Plates were imaged after 3 days at 30°C. **C)** Spot dilution growth assay with strains containing Prp22 mutants on YPD plates at different temperatures. Plates were imaged after the number of days indicated. **D)** Representative gel of *SUS1* RT-PCR products in strains with Prp22 mutants in the presence or absence of Fyv6. **E)** Quantitation of band intensities for each *SUS1* splice isoform as a fraction of total product in the lane. Bars indicate average ± SD of *N*=3 replicates. **F)** ACT1-CUP1 reporter construct with the location of the 3' SS substitution noted. **G)** Copper tolerance for each ACT1-CUP1 reporter with each Prp22 allele in strains with (■) or without (○) Fyv6 present. Shown is a representative, single replicate of *N*=3. **H)** Model for the role of Fyv6 during the 2nd step. Fyv6 interacts with Syf1 and Prp22 and promotes usage of BP distal 3' SS. When Fyv6 is absent, BP proximal 3' SS are favored. It is possible that lack of Fyv6 also results in the presence of Yju2 during exon ligation and relaxing of the requirement for Prp22 at this step.

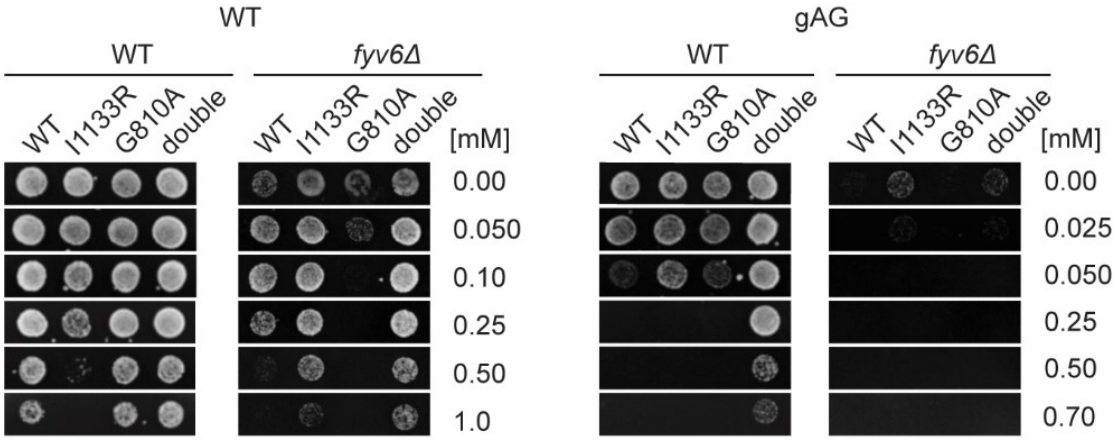

Figure 5.18 ACT1-CUP1 plate images for the data shown in Figure 7H. Images of yeast growth on copper-containing –Leu DO media shown after 48 (WT) or 72 h (*fyv6Δ*) for strains containing the indicated ACT1-CUP1 reporters.

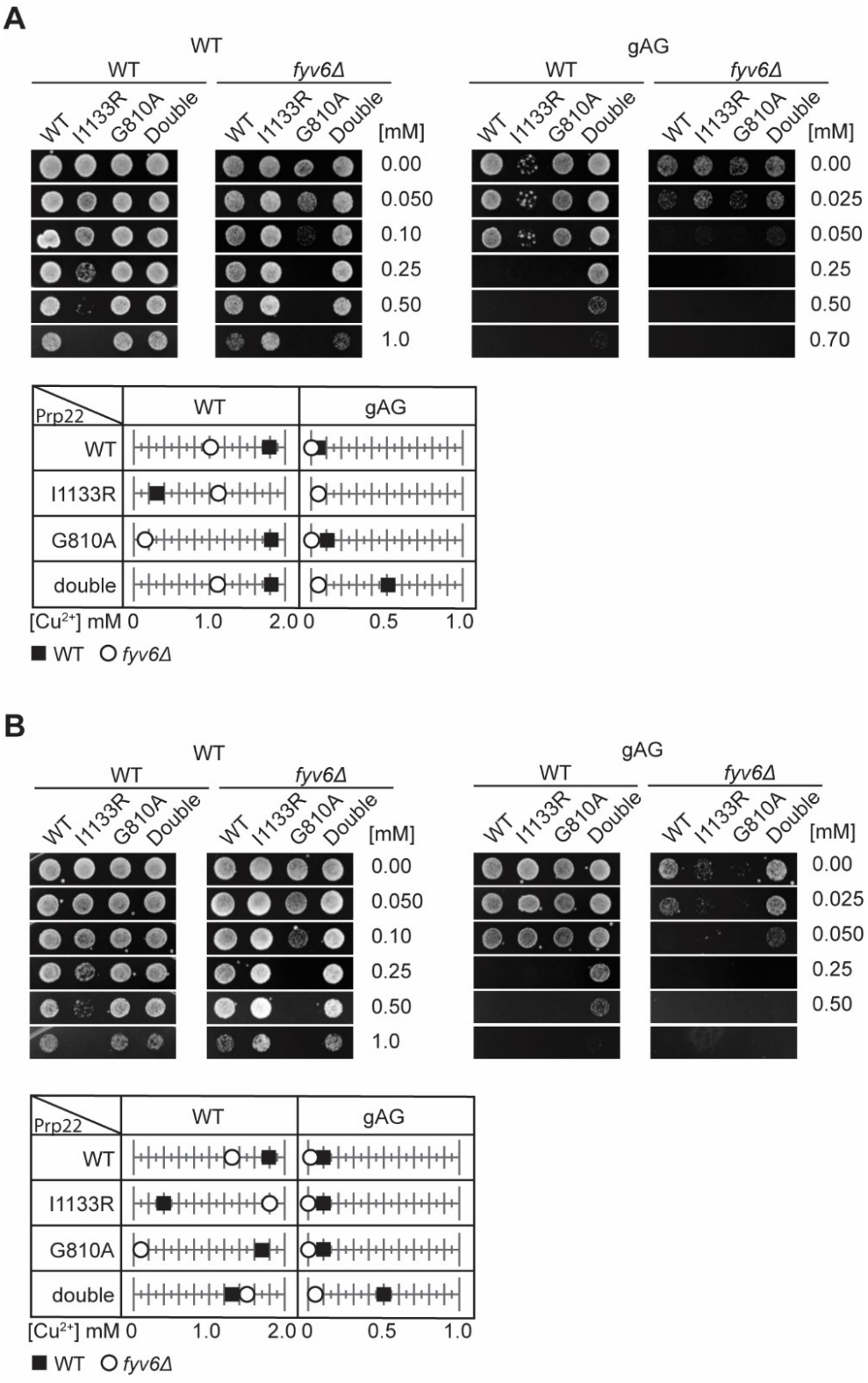

Figure 5.19 Replicate ACT1-CUP1 assays to those shown in Figure 7. A,B) Plate images and quantitation of copper tolerances for two additional replicates of the assay shown in Fig. 7G. Images of representative yeast growth on copper-containing –Leu D.O. media shown after 48 (WT) or 72 h (*fyv6Δ*) for strains containing the indicated ACT1-CUP1 reporters.

5.4 Discussion

Here, we establish Fyv6 as a 2nd step splicing factor in yeast that promotes usage of BP-distal 3' SS. Structural, genetic, and biochemical studies reveal critical contacts and interactions between Fyv6 and the splicing machinery. From these data, we propose that Yju2/Syf1 interactions need to be disrupted for proper Fyv6 recruitment and that Fyv6 can enforce the Prp22-dependence of at least some splicing events. Combined, our data allow us to propose a model for Fyv6's role during the 2nd step (**Fig. 5.17H**). In this model, Fyv6 is recruited to the C* complex after Prp16-dependent release of Yju2, which exposes the binding site for Fyv6 on Syf1. Once bound, Fyv6 (possibly in concert with Slu7 and Prp18) facilitates docking of BP distal 3' SS into the active site. In the absence of Fyv6, the use of BP distal 3' SS becomes more difficult and BP proximal 3' SS are favored. Under these conditions, the requirement for Prp22 during exon ligation is reduced and Yju2 may also potentially rebind Syf1. One or both of these factors could contribute to formation of a lower fidelity spliceosome with relaxed sequence constraints for exon ligation.

Structural Insights into the Recruitment and Release of 2nd step Factors

A key principle of spliceosome structural biochemistry is the presence of mutually exclusive interactions that coordinate the stepwise and directional progression of splicing (Townsend et al., 2020). Our cryo-EM structure of the P complex spliceosome suggests another such interaction is involved in the 1st to 2nd-step transition during which Fyv6 displaces Yju2. Yju2 and Fyv6 occupy overlapping binding sites on the splicing factor Syf1 and, in fact, unassigned density in spliceosome P complexes due to Fyv6 was previously mis-interpreted as Yju2 due to their structural similarity (Wilkinson et al., 2017; Wilkinson et al., 2021). We propose that Prp16-promoted spliceosome remodeling results in release of Yju2 from the spliceosome C complex and permits mutually exclusive binding of Fyv6. Fyv6 remains bound in the P complex but then dissociates to permit re-binding of Yju2 in the ILS complex.

Previously, Chiang and Cheng noted that while the N-terminal half of Yju2 was able to restore 1st step activity to Yju2 depleted extracts, it is much more stably associated if a second polypeptide containing the C-terminal, Syf1-interacting half was included (Chiang and Cheng, 2013). It is possible that competition between Yju2 and Fyv6 for Syf1 binding can explain these results. If only the N-terminal half of Yju2 is present, Fyv6 could associate prematurely with Syf1 and result in destabilization of the 1st step conformation in favor of 2nd step interactions. This may then result in weaker binding for the N-terminal half of Yju2. Failure to properly recruit and exchange the Yju2/Fyv6 interaction on Syf1 may also be the origin of some of the yeast growth and splicing phenotypes we observe (**Fig. 5.12**).

The need for Yju2/Fyv6 exchange hints at two distinct functions for Fyv6: a role in Yju2 displacement to prevent its re-binding during exon ligation and a role in stabilization

of the 2nd step spliceosome conformation. These two functions can potentially result in the distinct *cs* and *ts* phenotypes in yeast. While more work is needed to dissect these phenotypes, it is possible that the *cs* phenotype results from increased stability of the Yju2/Syf1 interaction at lower temperatures and the need for Fyv6 to act as a competitor of this interaction. The *ts* phenotype on the other hand may result from destabilization of the 2nd step conformation of the spliceosome active site due to loss of Fyv6. It is possible that loss of Fyv6 results in destabilization due to removal of a buttress that holds the RNase H-like domain of Prp8 in place (along with Prp18), changes in how the C-terminal domain of Prp22 interacts with the spliceosome active site, removal of an impediment to reversal of the U2 snRNA/BS duplex conformational change that occurs during the C to C* complex transition, or a combination of these factors.

The high quality of our cryo-EM data combined with 3D classification also provide new insights into the requirements for 3' SS docking into the active site and potentially for how the P complex spliceosome is disassembled into the ILS complex. It was previously speculated that the N-terminal domain of Slu7 may be critical for 3' SS docking by reducing the entropic cost of this process when the BP to 3' SS distance is long (Wilkinson et al., 2017). In our state II P complex structure, the 3' SS remains docked but this domain of Slu7 is not observed. This suggests that either Slu7 N-terminal domain interactions with Cwc22 and Prp8 are not essential for 3' SS docking into the active site for exon ligation and/or that these interactions are not essential for maintaining 3' SS docking within P complex once they form in C*.

In the state III P complex structure, densities for the 2nd step factors Fyv6, Prp18, and Slu7 are entirely missing and the 3' SS is no longer docked into the active site. This result suggests that 3' SS undocking is coupled to 2nd step factor release. Prp22 is still present in this structure, and it is possible that state III represents a conformation occurring after exon ligation but prior to Prp22-dependent mRNA release. In this case, the state III structure implies that Prp22 ATPase activity is not needed for dissociation of Fyv6, Prp18, and Slu7. It is also interesting to note that the C-terminal domain of Prp22 can no longer be resolved in the state III structure. What exactly activates the Prp22 ATPase to promote proofreading or mRNA release is not known. It is conceivable that undocking of the 3' SS from the active site can be sensed by the Prp22 C-terminal domain and this contributes to ATPase activation. This model differs from that previously proposed for Prp22-enabled 3' SS sampling in which that ATPase activity of Prp22 is needed to undock the 3' SS (Semlow et al., 2016). However, sampling could involve multiple steps: ATP-independent 3' SS undocking and ATP-dependent entry into the sampling state. Whether or not the 2nd step factors must release or alter conformation to enable sampling is unknown. Regardless of the exact mechanism, it is important to recognize that the P complex structures determined here were all obtained using a Prp22 mutant defective in mRNA release. The chemical and kinetic competencies of these states will require future investigations.

Multiple, Non-Redundant 2nd Step Factors Promote Use of BP Distal 3' SS

The challenge of 3' SS selection involves not just selection of the correct nucleotide sequence (YAG) but also identifying which YAG sequence to use. Here, we define a function for Fyv6 in facilitating use of BP distal 3' SS, especially those located ≥ 21 nt away from the branch. This function is similar to those of other 2nd step factors important for efficient splicing when the BP-3' SS distance is ≥12 nt (Slu7, Prp18) or ≥21 nt (Prp22) (Brys and Schwer, 1996; Schwer and Gross, 1998; Zhang and Schwer, 1997). Human Slu7 has additionally been reported to be important for preventing suppression of the productive 3' SS in favor of up- or downstream alternative 3' SS (Chua and Reed, 1999). Our previous study using competing 3' SS in the ACT1-CUP1 intron did not provide evidence for suppression of the productive 3' SS in the absence of Fyv6, suggesting that this function is not shared between human Slu7 and yeast Fyv6. Nonetheless, it is important to recognize that the splicing machinery employs multiple factors that enforce usage of BP distal 3' SS. These functions are not redundant as changes in 3' SS are still observed in the absence of Fyv6 but presence of Prp18 and vice versa (Roy et al., 2023).

Intron architecture can also contribute to 3' SS selection since canonical “AG” 3' SS are depleted in humans in the region between the BP and annotated 3' SS (the AG exclusion zone, or AGEZ) (Corvelo et al., 2010; Gooding et al., 2006). AG sites are also typically depleted in the corresponding region in yeast introns (data not shown). This would suggest that annotated 3' SS are also the most optimal sites located downstream of the BP. However, an important assumption of this model is that non-AG 3' SS can also be effectively proofread and not used. Deletion of either Prp18 or Fyv6 leads to activation of not just BP proximal 3' SS but also those with highly divergent sequences from the YAG consensus (Roy et al., 2023). This suggests that an AGEZ alone is insufficient for productive 3' SS usage and that mechanisms to ensure preferential use of YAG sites are critical.

It is likely that Fyv6 and Prp18 enforce usage of YAG sites by common as well as factor-specific mechanisms. Both proteins stabilize the conformation of the Prp8 RNase H-like domain in P complex. Numerous splice site suppressor mutations map to this domain including those that enhance splicing of non-YAG 3' SS (Galej et al., 2013). Fyv6 and Prp18 may be essential for establishing a “high fidelity” conformation of the RNase H-like domain that enforces use of canonical YAG 3' SS and preventing activation of cryptic sites within the AGEZ. In support of this, we note that in our state III P complex structure that lacks both Fyv6 and Prp18 and a stably docked 3' SS, the RNase H-like domain is also more flexible.

In addition to this possibility, Fyv6 and Prp18 may also contribute uniquely to 3' SS fidelity and selection. The conserved loop region of Prp18 inserts into the spliceosome active site and contacts the 3' SS. Partial deletion of this region of Prp18 phenocopies at least some of the changes in 3' SS selection when Prp18 is deleted entirely (Crotti et al., 2007; Roy et al., 2023). This indicates that even when the Prp8 RNaseH-like domain is

still stabilized by Prp18 (and Fyv6) disruptions to 3' SS selection can still occur. Unlike Prp18, Fyv6 does not contact the spliceosome active site. However, it is the only 2nd step factor that contacts exon ligation fidelity factor Prp22. While these contacts are not essential for suppressing use of an alternative 3' in *SUS1*, a mutation in Prp22 can nonetheless suppress a *fyv6Δ* phenotype and loss of Fyv6 relieves a Prp22 helicase mutant-dependent stall in *SUS1* splicing. These observations suggest that Fyv6 and Prp22 are functionally connected to one another.

Some 2nd Step Factors May Enforce Usage of Prp22 during Exon Ligation

We were particularly struck by the similarities in the BP-3' SS distances for the requirement of Prp22 during exon ligation and for promotion of BP distal 3' SS usage by Fyv6, both ≥21 nt (Schwer and Gross, 1998). In addition, the alternative, BP proximal 3' SS that are activated when Fyv6 is deleted show much less sequence conservation relative to the annotated 3' SS and include sequences (GAG, BG) known to be proofread and rejected by Prp22. Together, this suggests that one function of Fyv6 is to enforce Prp22-dependence of exon ligation by stabilizing a spliceosome conformation that more easily enables usage of BP distal 3' SS (**Fig. 5.17H**). As these sites require Prp22 for exon ligation, it is also likely that they are subject to Prp22-dependent proofreading, and this may enforce accurate exon ligation at the YAG 3' SS consensus. When Fyv6 is absent, many alternative, BP proximal 3' SS are activated. These sites could have a relaxed requirement for Prp22 and proofreading. While the Fyv6-dependence of Prp22 occupancy is beyond the scope of the current manuscript, loss of Prp18 (which also activates usage of non-canonical, BP-proximal 3' SS) does result in lower occupancy of Prp22 on the *ACT1* intron (Strittmatter et al., 2021). This suggests that Prp22 occupancy, and possible involvement in the exon ligation step, can be influenced by 2nd step factors.

There is precedent for nonessential splicing factors enforcing use of a proofreading ATPase. Cus2 is not essential for pre-spliceosome formation (Yan et al., 1998). However, when Cus2 is present then ATPase activity of Prp5 is essential for disruption of the Hsh155/Cus2 interaction (Perriman et al., 2003; Talkish et al., 2019). In this case, Cus2 is not essential for Prp5 proofreading of BP selection, and the connection between the ATP hydrolysis activity of Prp5, removal of Cus2, and proofreading is unclear (Liang and Cheng, 2015; Talkish et al., 2019; Xu and Query, 2007). It is possible that enforcement of Prp5 ATPase activity by Cus2 is important for other aspects of pre-spliceosome formation *in vivo* and the impact of Prp5 and Cus2 on BP selection has not yet been studied transcriptome-wide using high-resolution, sequencing-based approaches. The C-terminal domain of the 1st step splicing factor Yju2 (the same domain that competes with Fyv6 for Syf1 binding) is also not essential for 5' SS cleavage. In the absence of this domain, the N-terminal domain of Yju2 is sufficient for promoting the 1st step reaction, and the spliceosome can transition to the 2nd step and carry out exon ligation without Prp16, albeit inefficiently (Chiang and Cheng, 2013). Importantly, addition of the C-terminal domain on

a separate polypeptide reduces Prp16-independent activity. This suggests that the Yju2 C-terminal domain enforces the use of the Prp16 ATPase. As we propose with Fyv6 and Prp22, the C-terminal domain of Yju2 may not just help to promote the 1st to 2nd step transition by enforcing Prp16-dependence but could also facilitate Prp16-dependent proofreading during 5' SS cleavage (Koodathingal et al., 2010; Semlow et al., 2016). General operating principles of the yeast splicing machinery may not just include proofreading but also use of splicing factors that enforce usage of the proofreading ATPases.

5.5 Materials and Methods

The yeast strains and plasmids used in this study are described in **Tables 5.5 and 5.6**. Yeast transformation, plasmid shuffling/5-FOA selection, and growth were carried out using standard techniques and media (Treco and Lundblad 1993; Sikorski and Boeke 1991).

Cryo-electron microscopy

Purification of dominant-negative Prp22

Yeast Prp22 with a dominant-negative S635A mutation was cloned into pRS424 and pRS426 vectors after an N-terminal CBP-His-TEV tag (Galej et al., 2013). The protein was expressed in 24 L of BCY123 yeast cells essentially as described (Galej et al., 2013). Recombinant Prp22 was then purified first by Calmodulin-Sepharose in a buffer containing 50 mM Tris-HCl pH 8.5, 350 mM NaCl, 2 mM $CaCl_2$, 1 mM Mg acetate, 1 mM imidazole, 0.1% NP-40 substitute, 5 mM b-mercaptoethanol and eluted with 3 mM EGTA. The protein was dialyzed into 20 mM Tris-HCl pH 7.4, 350 mM NaCl, 10 mM imidazole, 0.1% NP-40 substitute and further purified over Ni-NTA resin at 500 mM NaCl and eluted with 200 mM imidazole. The eluted protein was treated with His-tagged TEV protease during dialysis against 20 mM HEPES-KOH pH 7.9, 300 mM KCl, 0.2 mM EDTA, 20% glycerol, 0.5 mM DTT, before removal of the protease with Ni-NTA resin. The final protein was frozen in liquid nitrogen and stored at -80 °C.

Preparation and purification of P complex for cryo-EM

P complex was purified as described in (Wilkinson et al., 2017). Brr2-TAPS yeast were grown in a 120 L fermenter and splicing extract was prepared by the liquid nitrogen method essentially as previously described (Lin et al., 1985). A pre-mRNA substrate consisting of 20 nt 5'-exon, 95 nt intron, and 32 nt 3'-exon from the UBC4 pre-mRNA followed by 3xMS2 stem loops was generated by run-off in vitro transcription from an EcoRI-linearized plasmid template. The RNA product was labelled at the 3' end with fluorescein-5-thiosemicarbazide (Wu et al., 1996). A 180 mL in vitro splicing reaction was performed containing 60 mM potassium phosphate pH 7, 2 mM ATP, 2.5 mM $MgCl_2$, 3% w/v PEG8000, 3 nM pre-mRNA, 37.5 nM MS2-MBP protein, and 40% (v/v) splicing extract treated with dominant negative Prp22 S635A. Reactions were incubated for 30 min at 23°C, then incubated for 20 min with 5 mM of a DNA oligonucleotide complementary to the 3'-exon (oligo sequence 5'-ATGAAGTAGGTGGAT-3') to induce cleavage of the 3'-MS2 tag by the endogenous RNase H activity of the splicing extract. The reaction mixture was centrifuged through a 40% glycerol cushion in buffer A (20 mM HEPES, pH 7.9, 75 mM KCl, 0.25 mM EDTA). The cushion was collected and applied to amylose resin in the presence of 0.025% NP-40 substitute. After overnight incubation at 4^{o}C the resin was washed and eluted with buffer A containing 5% glycerol, 0.01% NP-40 substitute and 12 mM maltose. Fractions containing spliceosomes were concentrated to 0.1 mg/mL then dialyzed against buffer A for 3 h.

For cryo-EM grid preparation, a freshly glow-discharged Cu300 R2/2 holey carbon grid with a 2-nm layer of amorphous carbon (Quantifoil) was mounted in the chamber of a Vitrobot Mark IV (Thermo Fisher Scientific) maintained at 4°C and 100% humidity. Spliceosomes (3 µL) were applied and after 60 s was manually blotted using Ø55 grade 595 filter paper (Ted Pella), followed by application of another 3 µL of spliceosomes, waiting 60 s, before blotting again and plunging into liquid ethane.

Cryo-EM data collection

Cryo-EM data were collected as two datasets from two grids made from the same sample. Dataset 1 was acquired using a Thermo Scientific Titan Krios cryo-TEM (LMB Krios 1) and dataset 2 was acquired using a Thermo Scientific Titan Krios G3i cryo-TEM (LMB Krios 3). Both used a K3 direct detector (Gatan) operated in super-resolution mode with 2-fold binning, and an energy filter with slit width of 20 eV. Micrographs were collected automatically using EPU in AFIS mode, yielding a total of 51,113 movies at 81,000x or 130,000x magnification with a real nominal pixel size of 0.93 Å or 0.669 Å, with defocus ranging from -1.3 µm to -3.1 µm and a total fluence per micrograph of 40 $e^-/Å^2$, fractionated into 40 frames.

Cryo-EM data processing

All cryo-EM data were processed using RELION-5.0 (Kimanius et al., 2023). Both datasets were initially processed separately. Movies were corrected for motion using the RELION implementation of MotionCor2, with 4x4 patches and dose-weighting. CTF parameters were estimated using CTFFIND-4.1. Particle picking was done using Topaz with the general model (Bepler et al., 2019), yielding 1,820,457 particles. 3D classification was performed on 4-fold (dataset 1) or 5-fold (dataset 2) binned particles using a P-complex reference map (EMD-10140), and 299,741 (dataset 1) or 599,123 (dataset 2) were selected and refined to 2.99 Å or 2.78 Å resolution respectively. After Bayesian Polishing and CTF refinement (per-particle defocus, anisotropic magnification, beam tilt, trefoil, and 4^{th} order aberrations) the resolution improved to 2.67 Å or 2.32 Å respectively. To improve the general density quality, 3D classification without alignment was performed with a soft mask around the whole complex with T=15. Particles with good density, without selecting for Fyv6 occupancy, were selected (151,443 and 257,501 for datasets 1 and 2) and refined to 2.66 Å or 2.27 Å respectively. The particles were then merged, refined, and then refined for anisotropic magnification. The overall scaling factors from the “anisotropic” magnification matrix was then used to calculate the corrected pixel size for dataset 1 as 0.94057 Å, keeping dataset 2 constant at 0.669 Å per pixel. Defocus values for dataset 1 were then scaled by the square of 0.94057 / 0.93, with an empirical constant correction of -10.06 Å (Wilkinson et al., 2019). Finally, another round of per-particle defocus refinement and anisotropic magnification refinement was performed to reduce errors in defocus correction and dataset scaling. The combined particles refined to 2.24 Å resolution. Resolution is reported using the gold-standard Fourier Shell Correlation with 0.143 cutoff.

The refined particles after merging were subjected to 3D classification without alignment, with a soft mask around the entire spliceosome, with T=15. The resulting four classes were used to define States I, II and III. To improve the densities for the periphery of the spliceosome that displayed local flexibility, focused refinements were used with Blush regularization in RELION-5 to reduce overfitting (Kimanius et al., 2023). The Prp22 density was improved using a soft-mask just around the Prp22 helicase domain for refinement of particles from States I and II, yielding a 2.96 Å map. Densities for the U2 snRNP, NTC, U5 Sm ring, and Cwc22 N-terminal domain were each improved first by 3D classification without alignment of particles from all three states, selecting particles with strong density (72,869 for the U2 snRNP, 260,409 for the NTC, 113,561 for the U5 Sm ring, 201,984 for Cwc22-NTD), performing signal subtraction to remove density for the rest of the spliceosome (recentering on the mask center of mass and re-boxing to 300 pixels, or 192 pixels for Cwc22-NTD), and finally 3D refinement starting with 1.8 degree local angular searches, giving 3.42 Å for the U2 snRNP, 3.72 Å for the NTC, 3.06 Å for the U5 Sm, and 3.57 Å for Cwc22-NTD. In all cases, control refinements without Blush regularization did not produce interpretable maps (**Fig. 5.8).**

Model building

The model was built using Coot (Casañal et al., 2020) and ISOLDE (Croll, 2018) starting with PDB 6EXN (Wilkinson et al., 2017) as an initial model for the core and PDB 7B9V (Wilkinson et al., 2021) as an initial model for the periphery. The core was improved by manual fixing in Coot, while the peripheral U2 snRNP and NTC were improved by applying torsion and distance restraints to reference models produced by AlphaFold2. The model was refined using PHENIX real_space_refine (Liebschner et al., 2019), just performing one macro-cycle. Figures were generated using UCSF ChimeraX.

Plasmid cloning

Fyv6 plasmids were made by PCR amplification from gDNA of the *FYV6* coding sequence plus ~250 nt upstream and downstream, restriction digest with NotI and SalI, and ligation into pRS414 or pRS416. An N-terminal 1x FLAG tag was added onto Fyv6 in the plasmids by PCR mutagenesis. Fyv6 truncations were cloned by Genewiz (Azenta Life Sciences) from pAAH1572. Restriction digest and ligation were used to place Fyv6 constructs into the pRS413 backbone.

Syf1 plasmids were made by PCR amplification from gDNA of the SYF1 coding sequence plus ~275 nt upstream and downstream, restriction digest, and ligation into pRS414 or pRS416.

Cef1 mutant plasmids, Prp22 I1133R mutant plasmid, Prp8 mutant plasmids, ACT1-CUP1 BP-3' SS distance reporters, and Syf1 truncation plasmids were prepared by PCR-based site directed mutagenesis of pAAH1611, pAAH1042, pAAH1440, pAAH0470, and pAAH1625 respectively. All plasmids were fully sequenced to confirm their nucleotide sequence.

Yeast strain creation

The *DBR1* and *FYV6* genes were deleted by replacement with a nourseothricin resistance cassette (NatMX4) or hygromycin resistance cassette (hphMX4), respectively, through homologous recombination (see **Table 5.6**; Goldstein and McCusker, 1999). The *SYF1* and *CEF1* genes were deleted by replacement with a kanamycin resistance cassette (KanMX) through homologous recombination in strains transformed with pRS416-Syf1 or pRS316-Cef1 respectively. Gene deletion was confirmed by colony PCR and/or genomic DNA extraction from the strains and PCR amplification of the corresponding genomic locus.

RNA isolation

Yeast were grown in YPD media until OD_{600} 0.5-0.8. For RNA-sequencing of temperature-shifted samples, cultures were shifted to either 37°C or 16°C or remained at 30°C and grown for 1 hr. A 15 mL volume of culture was harvested via vacuum filtration with mixed cellulose ester filter paper (Whatman, 47 mm diameter) in a Buchner funnel. Filter papers were separated into 5 mL tubes and flash frozen. Filters were stored at –80°C until RNA isolation. Filters were washed with 1 mL TRIzol to collect cells. For all other experiments involving RNA isolation, 10 OD_{600} units of OD_{600} 0.5-0.8 yeast were harvested via centrifugation.

Cell walls were disrupted by incubation with Zymolyase (50 U/mL, Zymo Research) in Y1 buffer (1 M sorbitol, 100 mM EDTA pH 8.0, 13 mM β-mercaptoethanol) (primer extension, RNA-seq of non-temperature-shifted samples) or vortexing with silica disruption beads in TRIzol (RT-PCR, RNA-seq of temperature-shifted samples). Total RNA was then isolated using TRIzol reagent (Thermo Fisher Scientific) according to manufacturer's instructions. Samples were cleaned using a 50 µg Monarch RNA Cleanup Kit (New England Biolabs) and treated with TURBO DNase (Thermo Fisher Scientific) to remove residual contaminating DNA. For RNA-sequencing, quality of RNA samples was assessed prior to library preparation by NanoDrop (concentration, A_{260}/A_{280}, and A_{260}/A_{230}), Qubit (RNA High Sensitivity and RNA IQ), and TapeStation (RNA ScreenTape Analysis). High quality samples showing minimal rRNA degradation with RIN >7 and A260/A280 ratios of ~2.0 were selected for library preparation.

Library preparation and RNA sequencing

Library preparation was conducted by the University of Wisconsin-Madison Biotechnology Center Gene Expression Center. mRNA was selected from total RNA samples through polyA enrichment with the TruSeq Stranded mRNA kit (Illumina). Sequencing was performed with an NovaSeq6000 sequencer (Illumina).

Bioinformatic analysis of RNA-seq datasets

FASTQC was used for quality control of reads pre- and post- trimming (Andrews, 2010). Trimming of reads was accomplished using fastp (Chen et al., 2018) prior to mapping to the SacCer3 genome (Ensembl, R64-1-1) with STAR (Dobin et al., 2013). All samples were aligned with STAR in a first pass. Novel junctions from all samples (WT and *fyv6Δ*) were combined and filtered for likely false-positive junctions. STAR was run in a second pass with the additional input of the filtered set of novel junctions. Indexing of bam files was accomplished with SAMtools (Li et al., 2009). Quantitation of read counts aligned to the SacCer3 transcriptome was accomplished with Salmon (Patro et al., 2017). Differential gene expression analysis was conducted with DeSeq2 (Love et al., 2014). Junction reads within genes were counted with featureCounts (Liao et al., 2014) and used to calculate fraction of Annotation Splicing (FAnS) based on methods in Roy et al., 2023. Canonical branch points were used as defined in the Ares Intron Database (Grate and Ares, 2002). Differential splicing analysis was conducted with SpliceWiz (Wong et al., 2024). The following Docker Images were used: biocontainers/fastp; alexdobin/star:2.7.10a_alpha_220506; pegi3s/feature-counts:2.0.0; combinelab/salmon:1.10.3; jysgro/fastqc:ub2306_12.1. Additional details and specific flags used for RNA-seq data analysis are as follows:

fastp: fastp --detect_adapter_for_pe -Q -q 20 -u 40 -l 36 --poly_g_min_len 10 -g -5 -3 -W 4 -M 20

STAR: All samples were aligned with STAR in a first pass. Novel junctions from all samples (WT and *fyv6Δ*) were combined and filtered for likely false-positive junctions (non-canonical junctions, column5 > 0; junctions supported by multi mappers only, column 7 > 0; junctions supported by two few reads, column 7 > 2). STAR was run in a second pass with the additional input of the filtered set of novel junctions as follows: --runMode alignReads --runThreadN 20 --genomeDir ./genome --sjdbGTFfile annotation.gtf --alignIntronMin 10 --alignIntronMax 2000 --readFilesCommand zcat --readFilesIn "R1_trimmed.fastq.gz" "R2_trimmed.fastq.gz" --outSAMtype BAM SortedByCoordinate --limitBAMsortRAM 6000000000

SpliceWiz: novelSplicing = TRUE, novelSplicing_requireOneAnnotatedSJ = TRUE, novelSplicing_minSamples = 1, novelSplicing_minSamplesAboveThreshold = 1, novelSplicing_countThreshold = 10, novelSplicing_useTJ = TRUE

featureCounts: featureCounts -a annotation.gtf -o Sample --splitOnly -J -f -p -T 5 -R BAM -B -C Aligned.sortedByCoord.out.bam

Salmon: Salmon was run on the free resource usegalaxy.org with the following flags: --libType A --incompatPrior '0.0' --biasSpeedSamp '5' --fldMax '1000' --fldMean '250' --fldSD '25' --forgettingFactor '0.65' --maxReadOcc '100' --numBiasSamples '2000000' --numAuxModelSamples '5000000' --numPreAuxModelSamples '5000'

--numGibbsSamples '0' --numBootstraps '0' --thinningFactor '16' --sigDigits '3' --vbPrior '1e-05'

DeSeq2: DeSeq2 was run on the free resource usegalaxy.org with the following flags: deseq2.R -o 'output.dat' -p -A 0.1 -H -f '[["Fyv6", [{"WT": ["WT1.tabular", "WT2.tabular"]}, {"Fyv6": ["Fyv61.tabular", "Fyv62.tabular"]}]]]' -l '{"Fyv61.tabular": "Fyv61.tabular", "Fyv62.tabular": "Fyv62.tabular", "WT1.tabular": "WT1.tabular", "WT2.tabular": "WT2.tabular"}' -t 1 -P -V 10 -i -y salmon -x

FAnS Calculation: Junction reads within genes were counted with featureCounts. Genes were filtered based on those listed in the Ares Intron database. Canonical 5′ and 3′ SS were annotated based on the highest number of junction counts. All junctions were filtered based on presence in all four RNA-seq samples. Read counts for filtered junctions were combined for *fyv6Δ* and WT replicates. Alternative junctions were separated into those that shared a canonical 5′ SS and an alternative 3′ SS or those that shared a canonical 3′ SS and an alternative 5′ SS. FAnS was calculated based on the number of junction reads for a unique alternative 3′ SS sharing a canonical 5′ SS divided by the number of junction reads for canonical 5′ SS and 3′ SS within a sample to adjust for changes in expression. Ratios of FAnS were calculated by dividing the FAnS value for *fyv6Δ* by the FAnS value for WT.

Primer extension

RNA was isolated using the protocol above. IR700 dye conjugated probes (Integrated DNA Technologies) were used for primer extension with SuperScript III (Thermo Fisher) of the ACT1-CUP1 reporters (2 pmol yAC6: /5IRD700/GGCACTCATGACCTTC) and U6 snRNA (0.4 pmol yU6: /5IRD700/GAACTGCTGATCATGTCTG) (Carrocci et al., 2017; van der Feltz et al., 2021). Primer extension products were visualized on a 7% (w/v) denaturing polyacrylamide gel run at 35 W for 80 min at RT. Gels were imaged with an Amersham Typhoon NIR laser scanner (Cytiva), and band intensities were quantified with Image J (version 1.53v, 2022).

RT-PCR

Yeast cultures were inoculated from cultures grown to stationary phase overnight and grown until OD_{600} = 0.7-0.9. RNA was isolated and depleted of contaminating DNA using the MasterPure Yeast RNA Purification Kit (LGC Biosearch Technologies) protocol as previously described (Carrocci et al., 2017) (RT-PCR shown in Fig. 4C) or using the RNA isolation protocol above (all others).

RT-PCR reactions were set up using the Access RT-PCR system (Promega Corporation) following the kit procedure with 100 ng of input RNA per 25 µl reaction. Primers to amplify *SUS1* were SUS1-exon1 5'-TGGATACTGCGCAATTAAAGAGTC-3' and SUS1-exon3 5'-TCATTGTGTATCTACAATCTCTTCAAG-3'. Reaction products were

separated on 2% (w/v) agarose-TBE gels and imaged using ethidium bromide fluorescence.

Selection of *fyv6Δ* suppressors

Strain yAAH3353 was previously described (Lipinski et al. 2023). Twenty single colonies of yAAH3353 were selected from a YPD plate and used to inoculate 20 x 5mL YPD liquid culture. Cultures were grown for 48 h at 30°C. A 1 mL volume of culture was spun down, resuspended in 200 µL YPD liquid medium, and spread on a YPD plate containing 40 mg/L adenine hemisulfate (to prevent ade2 reversion). Duplicate plates for each overnight were placed in incubators at either 18°C or 37°C. Many colonies of various sizes appeared on each plate by 9 days for 37°C and 12 days for 18°C. Colonies were picked after that time and annotated first by growth temperature, then by overnight number, and finally roughly by colony size with 01 generally being the largest colony on a plate (ex: 370101 from a plate grown at 37°C inoculated from overnight 01 and the largest colony selected from the plate). Picked colonies were grown overnight in 5mL YPD at 30°C and frozen stocks were made. Anywhere from 0 to ~15 colonies were picked per plate, covering the variety of colony sizes that arose.

Sequencing of yeast genomic DNA

Yeast were grown to saturated stationary phase in 5mL YPD and ~10 OD units were sent as a frozen yeast pellet to Azenta or SeqCoast for genomic DNA extraction, library, preparation, and sequencing with 30x/400 mbp coverage.

Bioinformatic analysis of WGS data

Azenta and SeqCoast utilized the SacCer3 genome (Ensembl, R64-1-1) for variant calling and provided a list of >6,000 SNPs. Since the strain is not a perfect match for this reference, non-unique SNPs were filtered from each dataset by taking the union of two sets with each set obtained from colonies isolated on different plates (to reduce the chance of computationally eliminating the same variant SNP). Filtering reduced the dataset to <100 unique variants per colony. Few unique SNPs were within protein- or snRNA-coding genes. Of SNPs within protein-coding regions, single amino acid changes were frequently found within splicing factors (see **Table 5.3**). Datasets were reexamined upon identification of suppressor mutations for suppressor mutation duplication that were potentially filtered out during comparison of SNP datasets.

Yeast temperature growth assays

Yeast were grown to stationary phase overnight in the appropriate liquid medium indicated in the figure legends. The cultures were first diluted to OD_{600} 0.5, serially diluted 1:10, 1:100, and 1:1000, and stamped onto agar plates containing the appropriate growth

medium. The plates were incubated at 16°, 23°, 30°, and 37°C, and plates were imaged after the number of days indicated in each figure.

ACT1-CUP1 copper tolerance assays

Yeast strains expressing ACT1-CUP1 reporters were grown to stationary phase in -Leu DO media, diluted to OD_{600} = 0.5 in 10% (v/v) glycerol, and spotted onto -Leu DO plates containing 0 to 2.5 mM $CuSO_4$ (Lesser and Guthrie 1993; Carrocci et al., 2018). Plates were scored and imaged after 48 h of growth at 30°C for WT strains and after 72 h of growth at 30°C for *fyv6Δ* strains due to differential growth between strains.

Table 5.5 Yeast strains

Strain name	Genotype	Description	Source
yAAH0434	MATα cup1Δ ura3 his3 trp1 lys2 ade2 leu2	Cu^{2+} sensitive strain	David Brow
yAAH3353	yAAH0434 + fyv6Δ::hphMX	Cu^{2+} sensitive *fyv6Δ* strain	Lipinski et al., 2023
yAAH3399	yAAH0434+upf1Δ::KanMX	Cu^{2+} sensitive *upf1Δ* strain	Lipinski et al., 2023
yAAH3400	yAAH3353+upf1Δ::KanMX	Cu^{2+} sensitive *fyv6Δ* *upf1Δ* strain	Lipinski et al., 2023
yAAH3403	yAAH0434 + dbr1Δ::NatR	Cu^{2+} sensitive *dbr1Δ* strain	This study
yAAH3404	yAAH3353 + dbr1Δ::NatR	Cu^{2+} sensitive *fyv6Δ dbr1Δ* strain	This study
yAAH3388	yAAH3353 + pAAH1555	Cu^{2+} sensitive Fyv6 shuffle strain	This study
yAAH3393	yAAH3353 + pAAH0135	vector control, contains pRS414	This study
yAAH3419	yAAH3353 + pAAH1572	FLAG-Fyv6 (N-terminal tag)	This study
yAAH3434	yAAH3353 + pAAH1577	FLAG-Fyv6-Δ1-16	This study
yAAH3435	yAAH3353 + pAAH1578	FLAG-Fyv6-Δ1-23	This study
yAAH3436	yAAH3353 + pAAH1579	FLAG-Fyv6-Δ1-51	This study
yAAH3437	yAAH3353 + pAAH1580	FLAG-Fyv6-Δ134-173	This study
yAAH3438	yAAH3353 + pAAH1581	FLAG-Fyv6-Δ103-173	This study
yAAH3418	yAAH3353 + pAAH1573	Fyv6-FLAG (C-terminal tag)	This study
yAAH3442	yAAH3353 + pAAH1586	Fyv6-Δ103-173 -FLAG	This study
yAAH3455	yAAH0434 + pAAH1602	$Prp18^{WT}$ merodiploid	This study
yAAH3456	yAAH0434 + pAAH1603	$Prp18^{WT}$/$Prp18^{V191A}$ merodiploid	This study
yAAH3457	yAAH0434 + pAAH1604	$Prp18^{WT}$ /$Prp18^{S162P}$ merodiploid	This study
yAAH3458	yAAH0434 + pAAH1605	$Prp18^{WT}$ /$Prp18^{S162P+V191A}$ merodiploid	This study
yAAH3459	yAAH3353 + pAAH1602	$Prp18^{WT}$ merodiploid in *fyv6Δ* strain	This study

yAAH3460	yAAH3353 + pAAH1603	$Prp18^{WT}$/$Prp18^{V191A}$ merodiploid in *fyv6Δ* strain	This study
yAAH3461	yAAH3353 + pAAH1604	$Prp18^{WT}$ /$Prp18^{S162P}$ merodiploid in *fyv6Δ* strain	This study
yAAH3462	yAAH3353 + pAAH1605	$Prp18^{WT}$ /$Prp18^{S162P+V191A}$ merodiploid in *fyv6Δ* strain	This study
yAAH3405	yAAH3403 + pAAH0470	*dbr1Δ* + WT ACT1-CUP1 (38 nt)	This study
yAAH3410	yAAH3404 + pAAH0470	*dbr1Δ fyv6Δ* + WT ACT1-CUP1 (38 nt)	This study
yAAH3498	yAAH3403 + pAAH1632	*dbr1Δ* + ACT1-CUP1 (9 nt)	This study
yAAH3499	yAAH3403 + pAAH1633	*dbr1Δ* + ACT1-CUP1 (12 nt)	This study
yAAH3500	yAAH3403 + pAAH1634	*dbr1Δ* + ACT1-CUP1 (21 nt)	This study
yAAH3501	yAAH3403 + pAAH1635	*dbr1Δ* + ACT1-CUP1 (27 nt)	This study
yAAH3503	yAAH3404 + pAAH1633	*dbr1Δ fyv6Δ* + ACT1-CUP1 (9 nt)	This study
yAAH3504	yAAH3404 + pAAH1634	*dbr1Δ fyv6Δ* + ACT1-CUP1 (12 nt)	This study
yAAH3505	yAAH3404 + pAAH1635	*dbr1Δ fyv6Δ* + ACT1-CUP1 (21 nt)	This study
yAAH3506	yAAH3404 + pAAH1632	*dbr1Δ fyv6Δ* + ACT1-CUP1 (27 nt)	This study
yAAH3509	yAAH3403 + pAAH1636	*dbr1Δ* + ACT1-CUP1 (15 nt)	This study
yAAH3510	yAAH3403 + pAAH1637	*dbr1Δ* + ACT1-CUP1 (42 nt)	This study
yAAH3511	yAAH3403 + pAAH1638	*dbr1Δ* + ACT1-CUP1 (46 nt)	This study
yAAH3512	yAAH3403 + pAAH1639	*dbr1Δ* + ACT1-CUP1 (50 nt)	This study
yAAH3513	yAAH3404 + pAAH1636	*dbr1Δ fyv6Δ* + ACT1-CUP1 (15 nt)	This study
yAAH3514	yAAH3404 + pAAH1637	*dbr1Δ fyv6Δ*+ ACT1-CUP1 (42 nt)	This study
yAAH3515	yAAH3404 + pAAH1638	*dbr1Δ fyv6Δ*+ ACT1-CUP1 (46 nt)	This study
yAAH3516	yAAH3404 + pAAH1639	*dbr1Δ fyv6Δ*+ ACT1-CUP1 (50 nt)	This study
yAAH3517	yAAH0434+syf1Δ::KanMX + pAAH1624	Cu^{2+} sensitive Syf1 shuffle strain	This study
yAAH3518	yAAH3517+ fyv6Δ::HygR	Cu^{2+} sensitive Syf1 shuffle strain with *fyv6Δ*	This study

yAAH3519	yAAH0434+syf1Δ::KanMX + pAAH1625	WT Syf1	This study
yAAH3577	yAAH0434+syf1Δ::KanMX + +pAAH1666	Syf1 Δ817-859	This study
yAAH3578	yAAH0434+syf1Δ::KanMX + fyv6Δ::HygR + pAAH1666	*fyv6Δ* + Syf1 Δ817-859	This study
yAAH3579	yAAH0434+syf1Δ::KanMX + +pAAH1667	Syf1 Δ778-859	This study
yAAH3580	yAAH0434+syf1Δ::KanMX + fyv6Δ::HygR + pAAH1667	*fyv6Δ* + Syf1 Δ778-859	This study
yAAH3581	yAAH0434+syf1Δ::KanMX + +pAAH1668	Syf1 Δ634-859	This study
yAAH3582	yAAH0434+syf1Δ::KanMX + fyv6Δ::HygR + pAAH1668	*fyv6Δ* + Syf1 Δ634-859	This study
yAAH3583	yAAH0434+syf1Δ::KanMX + fyv6Δ::HygR + pAAH1625 (WT Syf1)	*fyv6Δ* + WT Syf1	This study
yAAH3593	yAAH0434 + cef1Δ::KanMX + pAAH1658	Cu^{2+} sensitive Cef1 shuffle strain	This study
yAAH3594	yAAH3353 + cef1Δ::KanMX + pAAH1658	Cu^{2+} sensitive Cef1 shuffle strain with fyv6Δ	This study
yAAH3634	yAAH0434 + cef1Δ::KanMX + pAAH1611	$Cef1^{WT}$	This study
yAAH3636	yAAH0434 + cef1Δ::KanMX + pAAH1642	$Cef1^{M175I}$	This study
yAAH3637	yAAH0434 + cef1Δ::KanMX + pAAH1612	$Cef1^{A37P}$	This study
yAAH3638	yAAH0434 + cef1Δ::KanMX + pAAH1641	$Cef1^{A37V}$	This study
yAAH3639	yAAH0434 + cef1Δ::KanMX + pAAH1614	$Cef1^{V36R}$	This study
yAAH3640	yAAH0434 + cef1Δ::KanMX + pAAH1613	$Cef1^{S48R}$	This study
yAAH3641	yAAH0434 + cef1Δ::KanMX + pAAH1643	$Cef1^{Q193P}$	This study

yAAH3655	yAAH3353 + cef1Δ::KanMX + pAAH1612	*fyv6Δ* Cef1^{A37P}	This study
yAAH3656	yAAH3353 + cef1Δ::KanMX + pAAH1642	*fyv6Δ* Cef1^{M175I}	This study
yAAH3657	yAAH3353 + cef1Δ::KanMX + pAAH1641	*fyv6Δ* Cef1^{A37V}	This study
yAAH3658	yAAH3353 + cef1Δ::KanMX + pAAH1614	*fyv6Δ* Cef1^{V36R}	This study
yAAH3659	yAAH3353 + cef1Δ::KanMX + pAAH1613	*fyv6Δ* Cef1^{S48R}	This study
yAAH3660	yAAH3353 + cef1Δ::KanMX + pAAH1643	*fyv6Δ* Cef1^{Q193P}	This study
yAAH3661	yAAH3353 + cef1Δ::KanMX + pAAH1658	*fyv6Δ* Cef1^{WT}	This study
yAAH0117	ade2, cup1Δ:ura3 his3 leu2 lys2 prp8Δ:lys2 trp1 pJU169:PRP8(URA)	Cu^{2+} sensitive Prp8 shuffle strain	Christine Guthrie
yAAH3352	yAAH117 + fyv6Δ::HygR	Cu^{2+} sensitive *fyv6Δ* Prp8 shuffle strain	Lipinski et al., 2023
yAAH3093	ade2, cup1Δ:ura3 his3 leu2 lys2 prp8Δ:lys2 trp1 +pAAH1440	Prp8^{WT}	This study
yAAH3683	ade2, cup1Δ:ura3 his3 leu2 lys2 prp8Δ:lys2 trp1 +pAAH1659	Prp8^{S1584Y}	This study
yAAH3684	ade2, cup1Δ:ura3 his3 leu2 lys2 prp8Δ:lys2 trp1 +pAAH1660	Prp8^{S1584F}	This study
yAAH3685	ade2, cup1Δ:ura3 his3 leu2 lys2 prp8Δ:lys2 trp1 +pAAH1661	Prp8^{V1862L}	This study
yAAH3686	ade2, cup1Δ:ura3 his3 leu2 lys2 prp8Δ:lys2 trp1 +pAAH1662	Prp8^{G1868R}	This study
yAAH3687	ade2, cup1Δ:ura3 his3 leu2 lys2 prp8Δ:lys2 trp1 +pAAH1663	Prp8^{T1982S}	This study
yAAH3688	ade2, cup1Δ:ura3 his3 leu2 lys2 prp8Δ:lys2 trp1 fyv6Δ::HygR +pAAH1659	*fyv6Δ* Prp8^{S1584Y}	This study

yAAH3689	ade2, cup1Δ:ura3 his3 leu2 lys2 prp8Δ:lys2 trp1 fyv6Δ::HygR +pAAH1660	*fyv6Δ* Prp8^{S1584F}	This study
yAAH3690	ade2, cup1Δ:ura3 his3 leu2 lys2 prp8Δ:lys2 trp1 fyv6Δ::HygR +pAAH1661	*fyv6Δ* Prp8^{V1862L}	This study
yAAH3691	ade2, cup1Δ:ura3 his3 leu2 lys2 prp8Δ:lys2 trp1 fyv6Δ::HygR +pAAH1662	*fyv6Δ* Prp8^{G1868R}	This study
yAAH3692	ade2, cup1Δ:ura3 his3 leu2 lys2 prp8Δ:lys2 trp1 fyv6Δ::HygR +pAAH1663	*fyv6Δ* Prp8^{T1982S}	This study
yAAH3693	ade2, cup1Δ:ura3 his3 leu2 lys2 prp8Δ:lys2 trp1 fyv6Δ::HygR +pAAH1440	*fyv6Δ* Prp8^{WT}	This study
yAAH1930	MATa ade2 cup1Δ::ura3 his3 leu2 lys2 trp1 ura3 GAL+ prp22Δ::loxP pPrp22 (URA3)	Prp22 Shuffle Strain	Charles Query
yAAH1931	MATa ade2 cup1Δ::ura3 his3 leu2 lys2 trp1 ura3 GAL+ prp22Δ::loxP +pAAH1042	Prp22^{WT}	Charles Query
yAAH3377	MATa ade2 cup1Δ::ura3 his3 leu2 lys2 trp1 ura3 GAL+ prp22Δ::loxP fyv6Δ::hygR pPrp22 (URA3)	f*yv6Δ* Prp22 Shuffle Strain	Lipinski et al., 2023
yAAH3379	MATa ade2 cup1Δ::ura3 his3 leu2 lys2 trp1 ura3 GAL+ prp22Δ::loxP fyv6Δ::hygR + pAAH1042	*fyv6Δ* Prp22^{WT}	Lipinski et al., 2023
yAAH3558	MATa ade2 cup1Δ::ura3 his3 leu2 lys2 trp1 ura3 GAL+ prp22Δ::loxP +pAAH1648	Prp22^{I1133R}	This study
yAAH3359	MATa ade2 cup1Δ::ura3 his3 leu2 lys2 trp1 ura3 GAL+ prp22Δ::loxP fyv6Δ::hygR +pAAH1648	*fyv6Δ* Prp22^{I1133R}	This study
yAAH3606	MATa ade2 cup1Δ::ura3 his3 leu2 lys2 trp1 ura3 GAL+ prp22Δ::loxP +pAAH1665	Prp22^{R805A}	This study

yAAH3607	MATa ade2 cup1Δ::ura3 his3 leu2 lys2 trp1 ura3 GAL+ prp22Δ::loxP +pAAH1664	Prp22^{G810A}	This study
yAAH3608	MATa ade2 cup1Δ::ura3 his3 leu2 lys2 trp1 ura3 GAL+ prp22Δ::loxP fyv6Δ::hygR +pAAH1664	*fyv6Δ* Prp22^{G810A}	This study
yAAH3612	yAAH1930 + pAAH1675	pre-5-FOA selection Prp22$^{R805A+I1133R}$	This study
yAAH3613	yAAH1930 + pAAH1674	pre-5-FOA selection Prp22$^{G810A+I1133R}$	This study
yAAH3614	yAAH3377 + pAAH1665	pre-5-FOA selection *fyv6Δ* Prp22^{R805A}	This study
yAAH3615	yAAH3377 + pAAH1675	pre-5-FOA selection *fyv6Δ* Prp22$^{R805A+I1133R}$	This study
yAAH3616	yAAH3377 + pAAH1674	pre-5-FOA selection *fyv6Δ* Prp22$^{G810A+I1133R}$	This study
yAAH3632	MATa ade2 cup1Δ::ura3 his3 leu2 lys2 trp1 ura3 GAL+ prp22Δ::loxP fyv6Δ::hygR +pAAH1675	*fyv6Δ* Prp22$^{R805A+I1133R}$	This study
yAAH3633	MATa ade2 cup1Δ::ura3 his3 leu2 lys2 trp1 ura3 GAL+ prp22Δ::loxP fyv6Δ::hygR +pAAH1674	*fyv6Δ* Prp22$^{G810A+I1133R}$	This study
yAAH3635	MATa ade2 cup1Δ::ura3 his3 leu2 lys2 trp1 ura3 GAL+ prp22Δ::loxP +pAAH1674	Prp22$^{G810A+I1133R}$	This study
yAAH3662	MATa ade2 cup1Δ::ura3 his3 leu2 lys2 trp1 ura3 GAL+ prp22Δ::loxP +pAAH1675	Prp22$^{R805A+I1133R}$	This study
yAAH3663	yAAH1931 + pAAH0470	Prp22^{WT} + WT ACT1-CUP1	This study
yAAH3664	yAAH1931 + pAAH0526	Prp22^{WT} + ACT1-CUP1 U301G	This study
yAAH3665	yAAH1931 + pAAH0527	Prp22^{WT} + ACT1-CUP1 A302U	This study
yAAH3666	yAAH3607 + pAAH0470	Prp22^{G810A} + WT ACT1-CUP1	This study
yAAH3667	yAAH3607 + pAAH0527	Prp22^{G810A} + ACT1-CUP1 A302U	This study
yAAH3694	yAAH3607 + pAAH0526	Prp22^{G810A} + ACT1-CUP1 U301G	This study

yAAH3695	yAAH3379 + pAAH0470	*fyv6Δ* Prp22^{WT} + WT ACT1-CUP1	This study
yAAH3696	yAAH3379 + pAAH0526	*fyv6Δ* Prp22^{WT} + ACT1-CUP1 U301G (gAG)	This study
yAAH3697	yAAH3379 + pAAH0527	*fyv6Δ* Prp22^{WT} + ACT1-CUP1 A302U (UuG)	This study
yAAH3698	yAAH3608 + pAAH0470	*fyv6Δ* Prp22^{WT} + WT ACT1-CUP1	This study
yAAH3699	yAAH3608 + pAAH0526	*fyv6Δ* Prp22^{WT} + ACT1-CUP1 U301G (gAG)	This study
yAAH3700	yAAH3608 + pAAH0527	*fyv6Δ* Prp22^{WT} + ACT1-CUP1 A302U (UuG)	This study
yAAH3701	yAAH1931 + pAAH0880	Prp22^{WT} + ACT1-CUP1 BSG	This study
yAAH3702	yAAH3607 + pAAH0880	Prp22^{G810A} + ACT1-CUP1 BSG	This study
yAAH3703	yAAH3379 + pAAH0880	*fyv6Δ* Prp22^{WT} + ACT1-CUP1 BSG	This study
yAAH3704	yAAH3608 + pAAH0880	*fyv6Δ* Prp22^{G810A} + ACT1-CUP1 BSG	This study
yAAH3748	yAAH3559 + pAAH0470	*fyv6Δ* Prp22^{I1133R} + WT ACT1-CUP1	This study
yAAH3749	yAAH3559 + pAAH0526	*fyv6Δ* Prp22^{I1133R} + ACT1-CUP1 U301G (gAG)	This study
yAAH3750	yAAH3559 + pAAH0527	*fyv6Δ* Prp22^{I1133R} + ACT1-CUP1 A302U (UuG)	This study
yAAH3751	yAAH3559 + pAAH0880	*fyv6Δ* Prp22^{I1133R} + ACT1-CUP1 BSG	This study
yAAH3771	yAAH3635 + pAAH0470	Prp22$^{G810A+I1133R}$ + WT ACT1-CUP1	This study
yAAH3772	yAAH3635 + pAAH0526	Prp22$^{G810A+I1133R}$ + ACT1-CUP1 U301G	This study
yAAH3773	yAAH3635 + pAAH0527	Prp22$^{G810A+I1133R}$ + ACT1-CUP1 A302U	This study
yAAH3774	yAAH3635 + pAAH0880	Prp22$^{G810A+I1133R}$ + ACT1-CUP1 BSG	This study
yAAH3779	yAAH3558 + pAAH0470	Prp22^{I1133R} + WT ACT1-CUP1	This study
yAAH3780	yAAH3558 + pAAH0526	Prp22^{I1133R} + ACT1-CUP1 U301G	This study

yAAH3781	yAAH3558 + pAAH0527	Prp22^{I1133R} + ACT1-CUP1 A302U	This study
yAAH3782	yAAH3558 + pAAH0880	Prp22^{I1133R} + ACT1-CUP1 BSG	This study
yAAH3783	yAAH3633 + pAAH0470	*fyv6Δ* Prp22$^{G810A+I1133R}$ + WT ACT1-CUP1	This study
yAAH3784	yAAH3633 + pAAH0526	*fyv6Δ* Prp22$^{G810A+I1133R}$ + ACT1-CUP1 U301G	This study
yAAH3785	yAAH3633 + pAAH0527	*fyv6Δ* Prp22$^{G810A+I1133R}$ + ACT1-CUP1 A302U	This study
yAAH3786	yAAH3633 + pAAH0880	*fyv6Δ* Prp22$^{G810A+I1133R}$ + ACT1-CUP1 BSG	This study
BCY123	MATa pep4::HIS3 prb1::LEU2 bar1::HIS6 lys2::GAL1/10-GAL4 can1 ade2 trp1 ura3 his3 leu2-3,112	Protease-deficient yeast protein expression strain	Galej et al., 2013

Table 5.6 Plasmids

Plasmid ID	Plasmid name	Description	Source
pAAH0135	pRS414	CEN6/ARSH4 TRP1 Vector for Fyv6 plasmids	Mumberg et al., 1995 (ATCC# 87519)
pAAH1555	pRS416-Fyv6	FYV6 +/- ~250 bp (URA3 CEN6/ARSH4), used for Fyv6 shuffle strain	This study
pAAH1556	pRS414-Fyv6	FYV6 +/- ~250 bp (TRP1 CEN6/ARSH4)	This study
pAAH1572	pRS414-FLAG-Fyv6	N-terminally FLAG tagged FYV6 +/- ~250 bp (TRP1 CEN6/ARSH4)	This study
pAAH1577	pRS414-FLAG-Fyv6-Δ1-16	FLAG-Fyv6 Δ1-16 (TRP1 CEN ARS); in Fyv6 Δ1-16 truncation strain	This study
pAAH1578	pRS414-FLAG-Fyv6-Δ1-23	FLAG-Fyv6 Δ1-23 (TRP1 CEN ARS); in Fyv6 Δ1-23 truncation strain	This study
pAAH1579	pRS414-FLAG-Fyv6-Δ1-51	FLAG-Fyv6 Δ1-51 (TRP1 CEN ARS); in Fyv6 Δ1-51 truncation strain	This study
pAAH1580	pRS414-FLAG-Fyv6-Δ134-173	FLAG-Fyv6 Δ134-173 (TRP1 CEN ARS); in Fyv6 Δ134-173 truncation strain	This study
pAAH1581	pRS414-FLAG-Fyv6-Δ103-173	FLAG-Fyv6 Δ103-173 (TRP1 CEN ARS); in Fyv6 Δ103-173 truncation strain	This study
pAAH1573	pRS414-Fyv6-FLAG	C-terminally FLAG tagged FYV6 +/- ~250 bp (TRP1 CEN6/ARSH4)	This study
pAAH1586	pRS414-Fyv6-Δ103-173 -FLAG	Fyv6 Δ103-173-FLAG (TRP1 CEN ARS); in Fyv6 Δ103-173 truncation C-terminally FLAG tagged strain	This study
pAAH1602	p360-Prp18 WT	$Prp18^{WT}$ (URA3 CEN)	Aronova et al., 2007; gift from Beate Schwer
pAAH1603	p360-Prp18-11	$Prp18^{V191A}$ (URA3 CEN)	Aronova et al., 2007; gift from Beate Schwer
pAAH1604	p360-Prp18-18	$Prp18^{S162P}$ (URA3 CEN)	Aronova et al., 2007; gift from Beate Schwer
pAAH1605	p360-Prp18-11/18	$Prp18^{S162P+V191A}$ (URA3 CEN)	Aronova et al., 2007; gift from Beate Schwer

pAAH0470	ACT1-CUP1 WT	WT ACT1-CUP1 reporter; 38 nt BP-3' SS spacing. (GAP promoter, LEU2)	Gift from Charles Query.
pAAH1632	ACT1-CUP1-9nt	ACT1-CUP1 reporter with 9 nt BP-3' SS spacing. (GAP promoter, LEU2)	This study
pAAH1633	ACT1-CUP1-12nt	ACT1-CUP1 reporter with 12 nt BP-3' SS spacing. (GAP promoter, LEU2)	This study
pAAH1634	ACT1-CUP1-21nt	ACT1-CUP1 reporter with 21 nt BP-3' SS spacing. (GAP promoter, LEU2)	This study
pAAH1635	ACT1-CUP1-27nt	ACT1-CUP1 reporter with 27 nt BP-3' SS spacing. (GAP promoter, LEU2)	This study
pAAH1636	ACT1-CUP1-15nt	ACT1-CUP1 reporter with 27 nt BP-3' SS spacing. (GAP promoter, LEU2)	This study
pAAH1637	ACT1-CUP1-42nt	ACT1-CUP1 reporter with 42 nt BP-3' SS spacing. (GAP promoter, LEU2)	This study
pAAH1638	ACT1-CUP1-46nt	ACT1-CUP1 reporter with 46 nt BP-3' SS spacing. (GAP promoter, LEU2)	This study
pAAH1639	ACT1-CUP1-50nt	ACT1-CUP1 reporter with 50 nt BP-3' SS spacing. (GAP promoter, LEU2)	This study
pAAH1624	pRS416-Syf1	SYF1 +/- ~275 bp (URA3 CEN6/ARSH4), used for SYF1 shuffle strain	This study
pAAH1625	pRS414-Syf1	SYF1 +/- ~275 bp (TRP1 CEN6/ARSH4)	This study
pAAH1666	pRS414-Syf1Δ817-859	SYF1 Δ817-859 (TRP1 CEN6/ARSH4)	This study
pAAH1667	pRS414-Syf1Δ778-859	SYF1 Δ778-859 (TRP1 CEN6/ARSH4)	This study
pAAH1668	pRS414-Syf1Δ634-859	SYF1 Δ634-859 (TRP1 CEN6/ARSH4)	This study
pAAH1611	pRS314-Cef1-WT	$Cef1^{WT}$ (TRP1 CEN)	Query and Konarska, 2012 ; Gift from Charles Query.
pAAH1612	pRS314-Cef1-A37P	$Cef1^{A37P}$ (TRP1 CEN)	Query and Konarska, 2012; Gift from Charles Query.

pAAH1613	pRS314-Cef1-S48R	Cef1^{S48R} (TRP1 CEN)	Query and Konarska, 2012; Gift from Charles Query.
pAAH1614	pRS314-Cef1-9-8	Cef1^{V36R} (TRP1 CEN)	Query and Konarska, 2012; Gift from Charles Query.
pAAH1641	pRS314-Cef1-A37V	Cef1^{A37V} (TRP1 CEN)	This study
pAAH1642	pRS314-Cef1-M175I	Cef1^{M175I} (TRP1 CEN)	This study
pAAH1643	pRS314-Cef1-Q193P	Cef1^{Q193P} (TRP1 CEN)	This study
pAAH1658	PRS316-Cef1	WT Cef1 (URA3 CEN)	Query and Konarska, 2012; Gift from Charles Query.
pAAH1440	pRS424-Prp8	WT Prp8 (full length) (TRP1 2μ)	This study
pAAH1659	pRS424-Prp8-S1584Y	Prp8 S1584Y (TRP1 2μ)	This study
pAAH1660	pRS424-Prp8-S1584F	Prp8 S1584F (TRP1 2μ)	This study
pAAH1661	pRS424-Prp8-V1862L	Prp8 V1862L (TRP1 2μ)	This study
pAAH1662	pRS424-Prp8-G1868R	Prp8 G1868R (TRP1 2μ)	This study
pAAH1663	pRS424-Prp8-T1982S	Prp8 T1982S (TRP1 2μ)	This study
pAAH1042	pPrp22-WT	WT Prp22 (TRP1 CEN)	Gift from Charles Query.
pAAH1648	pPrp22-I1133R	Prp22 I1133R (TRP1 CEN)	This study
pAAH1664	pPrp22-G810A	Prp22 G810A (TRP1 CEN)	Schwer and Meszaros, 2000; Gift from Beate Schwer
pAAH1665	pPrp22-R805A	Prp22 R805A (TRP1 CEN)	Schwer and Meszaros, 2000; Gift from Beate Schwer
pAAH1674	pPrp22-G810A+I1133R	Prp22 G810A+I1133R (TRP1 CEN)	This study

pAAH1675	pPrp22-R805A+I1133R	Prp22 R805A+I1133R (TRP1 CEN)	This study
pRS424-CBP-His-TEV-Prp22-S635A	pRS424-CBP-His-TEV-Prp22-S635A	GAL/GAPDH promoter, Prp22 expression (TRP1 2μ)	This study
pRS426-CBP-His-TEV-Prp22-S635A	pRS426-CBP-His-TEV-Prp22-S635A	GAL/GAPDH promoter, Prp22 expression (URA3 2μ)	This study
pAAH0880	ACT1-CUP1 BSG (A259G)	BS A259G reporter used for ACT1-CUP1 assays. (GAP promoter, LEU2)	Gift from Charles Query.
pAAH0526	ACT1-CUP1 U301G	3' SS gAG reporter used for ACT1-CUP1 assays. (GAP promoter, LEU2)	Gift from Charles Query.
pAAH0527	ACT1-CUP1 A302U	3' SS UuG reporter used for ACT1-CUP1 assays. (GAP promoter, LEU2)	Gift from Charles Query.

5.6 References

Andrews, S. (2010). *FastQC: a quality control tool for high throughput sequence data* [Computer software]. http://www.bioinformatics.babraham.ac.uk/projects/fastqc

Aronova, A., Bacíková, D., Crotti, L. B., Horowitz, D. S., & Schwer, B. (2007). Functional interactions between Prp8, Prp18, Slu7, and U5 snRNA during the second step of pre-mRNA splicing. *RNA*, *13*(9), 1437–1444. https://doi.org/10.1261/rna.572807

Bai, R., Yan, C., Wan, R., Lei, J., & Shi, Y. (2017). Structure of the Post-catalytic Spliceosome from Saccharomyces cerevisiae. *Cell*, *171*(7), 1589-1598.e8. https://doi.org/10.1016/j.cell.2017.10.038

Bepler, T., Morin, A., Rapp, M., Brasch, J., Shapiro, L., Noble, A. J., & Berger, B. (2019). Positive-unlabeled convolutional neural networks for particle picking in cryo-electron micrographs. *Nature Methods*, *16*(11), Article 11. https://doi.org/10.1038/s41592-019-0575-8

Brow, D. A. (2019). An Allosteric Network for Spliceosome Activation Revealed by High-Throughput Suppressor Analysis in Saccharomyces cerevisiae. *Genetics*, *212*(1), 111–124. https://doi.org/10.1534/genetics.119.301922

Brys, A., & Schwer, B. (1996). Requirement for SLU7 in yeast pre-mRNA splicing is dictated by the distance between the branchpoint and the 3' splice site. *RNA*, *2*(7), 707–717.

Carrocci, T. J., Paulson, J. C., & Hoskins, A. A. (2018). Functional analysis of Hsh155/SF3b1 interactions with the U2 snRNA/branch site duplex. *RNA*, *24*(8), 1028–1040. https://doi.org/10.1261/rna.065664.118

Carrocci, T. J., Zoerner, D. M., Paulson, J. C., & Hoskins, A. A. (2017). SF3b1 mutations associated with myelodysplastic syndromes alter the fidelity of branchsite selection in yeast. *Nucleic Acids Research*, *45*(8), 4837–4852. https://doi.org/10.1093/nar/gkw1349

Casañal, A., Lohkamp, B., & Emsley, P. (2020). Current developments in Coot for macromolecular model building of Electron Cryo-microscopy and Crystallographic Data. *Protein Science*, *29*(4), 1055–1064. https://doi.org/10.1002/pro.3791

Chen, C.-H., Tsai, W.-Y., Chen, H.-R., Wang, C.-H., & Cheng, S.-C. (2001). Identification and Characterization of Two Novel Components of The Prp19p-associated Complex, Ntc30p and Ntc20p *. *Journal of Biological Chemistry*, *276*(1), 488–494. https://doi.org/10.1074/jbc.M006958200

Chen, S., Zhou, Y., Chen, Y., & Gu, J. (2018). fastp: An ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics*, *34*(17), i884–i890. https://doi.org/10.1093/bioinformatics/bty560

Chiang, T.-W., & Cheng, S.-C. (2013). A Weak Spliceosome-Binding Domain of Yju2 Functions in the First Step and Bypasses Prp16 in the Second Step of Splicing.

Molecular and Cellular Biology, *33*(9), 1746–1755. https://doi.org/10.1128/MCB.00035-13

Chua, K., & Reed, R. (1999). The RNA splicing factor hSlu7 is required for correct 3′ splice-site choice. *Nature*, *402*(6758), 207–210. https://doi.org/10.1038/46086

Chung, C.-S., Wai, H. L., Kao, C.-Y., & Cheng, S.-C. (2023). An ATP-independent role for Prp16 in promoting aberrant splicing. *Nucleic Acids Research*, *51*(20), 10815–10828. https://doi.org/10.1093/nar/gkad861

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

Company, M., Arenas, J., & Abelson, J. (1991). Requirement of the RNA helicase-like protein PRP22 for release of messenger RNA from spliceosomes. *Nature*, *349*(6309), 487–493. https://doi.org/10.1038/349487a0

Corvelo, A., Hallegger, M., Smith, C. W. J., & Eyras, E. (2010). Genome-Wide Association between Branch Point Properties and Alternative Splicing. *PLOS Computational Biology*, *6*(11), e1001016. https://doi.org/10.1371/journal.pcbi.1001016

Croll, T. I. (2018). ISOLDE: A physically realistic environment for model building into low-resolution electron-density maps. *Acta Crystallographica Section D: Structural Biology*, *74*(6), 519–530. https://doi.org/10.1107/S2059798318002425

Crotti, L. B., Bačíková, D., & Horowitz, D. S. (2007). The Prp18 protein stabilizes the interaction of both exons with the U5 snRNA during the second step of pre-mRNA splicing. *Genes & Development*, *21*(10), 1204–1216. https://doi.org/10.1101/gad.1538207

Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., & Gingeras, T. R. (2013). STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics*, *29*(1), 15–21. https://doi.org/10.1093/bioinformatics/bts635

Felisberto-Rodrigues, C., Thomas, J. C., McAndrew, C., Le Bihan, Y.-V., Burke, R., Workman, P., & van Montfort, R. L. M. (2019). Structural and functional characterisation of human RNA helicase DHX8 provides insights into the mechanism of RNA-stimulated ADP release. *Biochemical Journal*, *476*(18), 2521–2543. https://doi.org/10.1042/BCJ20190383

Feltz, C. van der, Nikolai, B., Schneider, C., Paulson, J. C., Fu, X., & Hoskins, A. A. (2021). Saccharomyces cerevisiae Ecm2 modulates the catalytic steps of pre-mRNA splicing. *RNA*, *27*(5), 591–603. https://doi.org/10.1261/rna.077727.120

Fica, S. M., & Nagai, K. (2017). Cryo-electron microscopy snapshots of the spliceosome: Structural insights into a dynamic ribonucleoprotein machine. *Nature Structural & Molecular Biology*, *24*(10), Article 10. https://doi.org/10.1038/nsmb.3463

Fica, S. M., Oubridge, C., Wilkinson, M. E., Newman, A. J., & Nagai, K. (2019). A human postcatalytic spliceosome structure reveals essential roles of metazoan factors for exon ligation. *Science*, *363*(6428), 710–714. https://doi.org/10.1126/science.aaw5569

Frank, D., & Guthrie, C. (1992). An essential splicing factor, SLU7, mediates 3' splice site choice in yeast. *Genes & Development*, *6*(11), 2112–2124. https://doi.org/10.1101/gad.6.11.2112

Galej, W. P., Oubridge, C., Newman, A. J., & Nagai, K. (2013). Crystal structure of Prp8 reveals active site cavity of the spliceosome. *Nature*, *493*(7434), 638–643. https://doi.org/10.1038/nature11843

Gautam, A., Grainger, R. J., Vilardell, J., Barrass, J. D., & Beggs, J. D. (2015). Cwc21p promotes the second step conformation of the spliceosome and modulates 3′ splice site selection. *Nucleic Acids Research*, *43*(6), 3309–3317. https://doi.org/10.1093/nar/gkv159

Goldstein, A. L., & McCusker, J. H. (1999). Three new dominant drug resistance cassettes for gene disruption in Saccharomyces cerevisiae. *Yeast*, *15*(14), 1541–1553. https://doi.org/10.1002/(SICI)1097-0061(199910)15:14<1541::AID-YEA476>3.0.CO;2-K

Gooding, C., Clark, F., Wollerton, M. C., Grellscheid, S.-N., Groom, H., & Smith, C. W. (2006). A class of human exons with predicted distant branch points revealed by analysis of AG dinucleotide exclusion zones. *Genome Biology*, *7*(1), R1. https://doi.org/10.1186/gb-2006-7-1-r1

Grate, L., & Ares, M. (2002). Searching yeast intron data at ares lab web site. In C. Guthrie & G. R. Fink (Eds.), *Methods in Enzymology* (Vol. 350, pp. 380–392). Academic Press. https://doi.org/10.1016/S0076-6879(02)50975-7

Kawashima, T., Douglass, S., Gabunilas, J., Pellegrini, M., & Chanfreau, G. F. (2014). Widespread Use of Non-productive Alternative Splice Sites in Saccharomyces cerevisiae. *PLOS Genetics*, *10*(4), e1004249. https://doi.org/10.1371/journal.pgen.1004249

Kimanius, D., Jamali, K., Wilkinson, M. E., Lövestam, S., Velazhahan, V., Nakane, T., & Scheres, S. H. W. (2023). *Data-driven regularisation lowers the size barrier of cryo-EM structure determination* (p. 2023.10.23.563586). bioRxiv. https://doi.org/10.1101/2023.10.23.563586

Koodathingal, P., Novak, T., Piccirilli, J. A., & Staley, J. P. (2010). The DEAH Box ATPases Prp16 and Prp43 Cooperate to Proofread 5′ Splice Site Cleavage during Pre-mRNA Splicing. *Molecular Cell*, *39*(3), 385–395. https://doi.org/10.1016/j.molcel.2010.07.014

Kuhn, A. N., & Brow, D. A. (2000). Suppressors of a Cold-Sensitive Mutation in Yeast U4 RNA Define Five Domains in the Splicing Factor Prp8 That Influence

Spliceosome Activation. *Genetics*, *155*(4), 1667–1682. https://doi.org/10.1093/genetics/155.4.1667

Kuhn, A. N., Li, Z., & Brow, D. A. (1999). Splicing Factor Prp8 Governs U4/U6 RNA Unwinding during Activation of the Spliceosome. *Molecular Cell*, *3*(1), 65–75. https://doi.org/10.1016/S1097-2765(00)80175-6

Lesser, C. F., & Guthrie, C. (1993). Mutational analysis of pre-mRNA splicing in Saccharomyces cerevisiae using a sensitive new reporter gene, CUP1. *Genetics*, *133*(4), 851–863. https://doi.org/10.1093/genetics/133.4.851

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., & 1000 Genome Project Data Processing Subgroup. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics*, *25*(16), 2078–2079. https://doi.org/10.1093/bioinformatics/btp352

Liang, W.-W., & Cheng, S.-C. (2015). A novel mechanism for Prp5 function in prespliceosome formation and proofreading the branch site sequence. *Genes & Development*, *29*(1), 81–93. https://doi.org/10.1101/gad.253708.114

Liao, Y., Smyth, G. K., & Shi, W. (2014). featureCounts: An efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics*, *30*(7), 923–930. https://doi.org/10.1093/bioinformatics/btt656

Liebschner, D., Afonine, P. V., Baker, M. L., Bunkóczi, G., Chen, V. B., Croll, T. I., Hintze, B., Hung, L. W., Jain, S., McCoy, A. J., Moriarty, N. W., Oeffner, R. D., Poon, B. K., Prisant, M. G., Read, R. J., Richardson, J. S., Richardson, D. C., Sammito, M. D., Sobolev, O. V., … Adams, P. D. (2019). Macromolecular structure determination using X-rays, neutrons and electrons: Recent developments in Phenix. *Acta Crystallographica. Section D, Structural Biology*, *75*(Pt 10), 861–877. https://doi.org/10.1107/S2059798319011471

Lin, R. J., Newman, A. J., Cheng, S. C., & Abelson, J. (1985). Yeast mRNA splicing in vitro. *Journal of Biological Chemistry*, *260*(27), 14780–14792. https://doi.org/10.1016/S0021-9258(17)38641-6

Lipinski, K. A., Senn, K. A., Zeps, N. J., & Hoskins, A. A. (2023). Biochemical and Genetic Evidence Supports Fyv6 as a Second-Step Splicing Factor in Saccharomyces cerevisiae. *RNA*, rna.079607.123. https://doi.org/10.1261/rna.079607.123

Liu, L., Query, C. C., & Konarska, M. M. (2007). Opposing classes of prp8 alleles modulate the transition between the catalytic steps of pre-mRNA splicing. *Nature Structural & Molecular Biology*, *14*(6), Article 6. https://doi.org/10.1038/nsmb1240

Liu, S., Li, X., Zhang, L., Jiang, J., Hill, R. C., Cui, Y., Hansen, K. C., Zhou, Z. H., & Zhao, R. (2017). Structure of the yeast spliceosomal postcatalytic P complex. *Science*, *358*(6368), Article 6368. https://doi.org/10.1126/science.aar3462

Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*, *15*(12), 550. https://doi.org/10.1186/s13059-014-0550-8

Mayas, R. M., Maita, H., & Staley, J. P. (2006). Exon ligation is proofread by the DExD/H-box ATPase Prp22p. *Nature Structural & Molecular Biology*, *13*(6), Article 6. https://doi.org/10.1038/nsmb1093

McPheeters, D. S., & Abelson, J. (1992). Mutational analysis of the yeast U2 snRNA suggests a structural similarity to the catalytic core of group I introns. *Cell*, *71*(5), 819–831. https://doi.org/10.1016/0092-8674(92)90557-s

Mendoza-Ochoa, G. I., Barrass, J. D., Maudlin, I. E., & Beggs, J. D. (2019). Blocking late stages of splicing quickly limits pre-spliceosome assembly in vivo. *RNA Biology*, *16*(12), 1775–1784. https://doi.org/10.1080/15476286.2019.1657788

Montemayor, E. J., Curran, E. C., Liao, H. H., Andrews, K. L., Treba, C. N., Butcher, S. E., & Brow, D. A. (2014). Core structure of the U6 small nuclear ribonucleoprotein at 1.7-Å resolution. *Nature Structural & Molecular Biology*, *21*(6), Article 6. https://doi.org/10.1038/nsmb.2832

Mumberg, D., Müller, R., & Funk, M. (1995). Yeast vectors for the controlled expression of heterologous proteins in different genetic backgrounds. *Gene*, *156*(1), 119–122. https://doi.org/10.1016/0378-1119(95)00037-7

Munding, E. M., Shiue, L., Katzman, S., Donohue, J. P., & Ares, M. (2013). Competition between Pre-mRNAs for the Splicing Machinery Drives Global Regulation of Splicing. *Molecular Cell*, *51*(3), 338–348. https://doi.org/10.1016/j.molcel.2013.06.012

Nguyen Ba, A. N., Pogoutse, A., Provart, N., & Moses, A. M. (2009). NLStradamus: A simple Hidden Markov Model for nuclear localization signal prediction. *BMC Bioinformatics*, *10*(1), 202. https://doi.org/10.1186/1471-2105-10-202

Patro, R., Duggal, G., Love, M. I., Irizarry, R. A., & Kingsford, C. (2017). Salmon provides fast and bias-aware quantification of transcript expression. *Nature Methods*, *14*(4), 417–419. https://doi.org/10.1038/nmeth.4197

Perriman, R., Barta, I., Voeltz, G. K., Abelson, J., & Ares, M. (2003). ATP requirement for Prp5p function is determined by Cus2p and the structure of U2 small nuclear RNA. *Proceedings of the National Academy of Sciences*, *100*(24), 13857–13862. https://doi.org/10.1073/pnas.2036312100

Plaschka, C., Newman, A. J., & Nagai, K. (2019). Structural Basis of Nuclear pre-mRNA Splicing: Lessons from Yeast. *Cold Spring Harbor Perspectives in Biology*, *11*(5), a032391. https://doi.org/10.1101/cshperspect.a032391

Query, C. C., & Konarska, M. M. (2004). Suppression of Multiple Substrate Mutations by Spliceosomal prp8 Alleles Suggests Functional Correlations with Ribosomal Ambiguity Mutants. *Molecular Cell*, *14*(3), 343–354. https://doi.org/10.1016/S1097-2765(04)00217-5

Query, C. C., & Konarska, M. M. (2012). CEF1/CDC5 alleles modulate transitions between catalytic conformations of the spliceosome. *RNA*, *18*(5), 1001–1013. https://doi.org/10.1261/rna.029421.111

Roy, K. R., Gabunilas, J., Neutel, D., Ai, M., Yeh, Z., Samson, J., Lyu, G., & Chanfreau, G. F. (2023). Splicing factor Prp18p promotes genome-wide fidelity of consensus 3′-splice sites. *Nucleic Acids Research*, gkad968. https://doi.org/10.1093/nar/gkad968

Sayani, S., Janis, M., Lee, C. Y., Toesca, I., & Chanfreau, G. F. (2008). Widespread Impact of Nonsense-Mediated mRNA Decay on the Yeast Intronome. *Molecular Cell*, *31*(3), 360–370. https://doi.org/10.1016/j.molcel.2008.07.005

Schwer, B. (2008). A Conformational Rearrangement in the Spliceosome Sets the Stage for Prp22-Dependent mRNA Release. *Molecular Cell*, *30*(6), Article 6. https://doi.org/10.1016/j.molcel.2008.05.003

Schwer, B., & Gross, C. H. (1998). Prp22, a DExH-box RNA helicase, plays two distinct roles in yeast pre-mRNA splicing. *The EMBO Journal*, *17*(7), 2086–2094. https://doi.org/10.1093/emboj/17.7.2086

Schwer, B., & Guthrie, C. (1992). A conformational rearrangement in the spliceosome is dependent on PRP16 and ATP hydrolysis. *The EMBO Journal*, *11*(13), 5033–5039. https://doi.org/10.1002/j.1460-2075.1992.tb05610.x

Schwer, B., & Meszaros, T. (2000). RNA helicase dynamics in pre-mRNA splicing. *The EMBO Journal*, *19*(23), 6582–6591. https://doi.org/10.1093/emboj/19.23.6582

Semlow, D. R., Blanco, M. R., Walter, N. G., & Staley, J. P. (2016). Spliceosomal DEAH-Box ATPases Remodel Pre-mRNA to Activate Alternative Splice Sites. *Cell*, *164*(5), Article 5. https://doi.org/10.1016/j.cell.2016.01.025

Semlow, D. R., & Staley, J. P. (2012). Staying on message: Ensuring fidelity in pre-mRNA splicing. *Trends in Biochemical Sciences*, *37*(7), 263–273. https://doi.org/10.1016/j.tibs.2012.04.001

Siatecka, M., Reyes, J. L., & Konarska, M. M. (1999). Functional interactions of Prp8 with both splice sites at the spliceosomal catalytic center. *Genes & Development*, *13*(15), 1983–1993.

Sikorski, R. S., & Boeke, J. D. (1991). In Vitro Mutagenesis and Plasmid Shuffling: From Cloned Gene to Mutant Yeast. *Methods in Enzymology*, *194*(C), 302–318. https://doi.org/10.1016/0076-6879(91)94023-6

Strittmatter, L. M., Capitanchik, C., Newman, A. J., Hallegger, M., Norman, C. M., Fica, S. M., Oubridge, C., Luscombe, N. M., Ule, J., & Nagai, K. (2021). psiCLIP reveals dynamic RNA binding by DEAH-box helicases before and after exon ligation. *Nature Communications*, *12*(1), 1488. https://doi.org/10.1038/s41467-021-21745-9

Talkish, J., Igel, H., Hunter, O., Horner, S. W., Jeffery, N. N., Leach, J. R., Jenkins, J. L., Kielkopf, C. L., & Ares, M. (2019). Cus2 enforces the first ATP-dependent step of

splicing by binding to yeast SF3b1 through a UHM–ULM interaction. *RNA*, *25*(8), 1020–1037. https://doi.org/10.1261/rna.070649.119

Townsend, C., Leelaram, M. N., Agafonov, D. E., Dybkov, O., Will, C. L., Bertram, K., Urlaub, H., Kastner, B., Stark, H., & Lührmann, R. (2020). Mechanism of protein-guided folding of the active site U2/U6 RNA during spliceosome activation. *Science*, *370*(6523), eabc3753. https://doi.org/10.1126/science.abc3753

Treco, D. A., & Lundblad, V. (1993). Preparation of Yeast Media. *Current Protocols in Molecular Biology*, *23*(1), 13.1.1-13.1.7. https://doi.org/10.1002/0471142727.mb1301s23

Wan, R., Yan, C., Bai, R., Lei, J., & Shi, Y. (2017). Structure of an Intron Lariat Spliceosome from Saccharomyces cerevisiae. *Cell*, *171*(1), 120-132.e12. https://doi.org/10.1016/j.cell.2017.08.029

Wilkinson, M. E., Fica, S. M., Galej, W. P., & Nagai, K. (2021). Structural basis for conformational equilibrium of the catalytic spliceosome. *Molecular Cell*. https://doi.org/10.1016/j.molcel.2021.02.021

Wilkinson, M. E., Fica, S. M., Galej, W. P., Norman, C. M., Newman, A. J., & Nagai, K. (2017). Postcatalytic spliceosome structure reveals mechanism of 3′–splice site selection. *Science*, *358*(6368), 1283–1288. https://doi.org/10.1126/science.aar3729

Wilkinson, M. E., Kumar, A., & Casañal, A. (2019). Methods for merging data sets in electron cryo-microscopy. *Acta Crystallographica Section D: Structural Biology*, *75*(9), 782–791. https://doi.org/10.1107/S2059798319010519

Wong, A. C. H., Wong, J. J.-L., Rasko, J. E. J., & Schmitz, U. (2024). SpliceWiz: Interactive analysis and visualization of alternative splicing in R. *Briefings in Bioinformatics*, *25*(1), bbad468. https://doi.org/10.1093/bib/bbad468

Wu, T., Ruan, K., & Liu, W. (1996). A Fluorescence-Labeling Method for Sequencing Small RNA on Polyacrylamide Gel. *Nucleic Acids Research*, *24*(17), 3472–3473. https://doi.org/10.1093/nar/24.17.3472

Xu, Y.-Z., & Query, C. C. (2007). Competition between the ATPase Prp5 and Branch Region-U2 snRNA Pairing Modulates the Fidelity of Spliceosome Assembly. *Molecular Cell*, *28*(5), 838–849. https://doi.org/10.1016/j.molcel.2007.09.022

Yan, D., Perriman, R., Igel, H., Howe, K. J., Neville, M., & Ares Jr., M. (1998). CUS2, a Yeast Homolog of Human Tat-SF1, Rescues Function of Misfolded U2 through an Unusual RNA Recognition Motif. *Molecular and Cellular Biology*, *18*(9), 5000–5009. https://doi.org/10.1128/MCB.18.9.5000

Zhan, X., Lu, Y., Zhang, X., Yan, C., & Shi, Y. (2022). Mechanism of exon ligation by human spliceosome. *Molecular Cell*, *82*(15), Article 15. https://doi.org/10.1016/j.molcel.2022.05.021

Zhang, X., & Schwer, B. (1997). Functional and physical interaction between the yeast splicing factors Slu7 and Prp18. *Nucleic Acids Research*, *25*(11), 2146–2152. https://doi.org/10.1093/nar/25.11.2146

Zhang, X., Zhan, X., Yan, C., Zhang, W., Liu, D., Lei, J., & Shi, Y. (2019). Structures of the human spliceosomes before and after release of the ligated exon. *Cell Research*, *29*(4), Article 4. https://doi.org/10.1038/s41422-019-0143-x

CHAPTER 6

Discussion

CHAPTER 6: Discussion

6.1 Summary

The studies presented in this thesis highlight the significance of RNA-protein interactions in pre-mRNA splicing. First, the synthesis of U6 snRNA by RNA polymerase III (RNAP III) is investigated, revealing that U6 snRNA function can be maintained when synthesized by RNAP II. This flexibility allows for relaxation of sequence restraints imposed by RNAP III promoters and terminators. Moreover, the addition of U4 snRNA Sm protein binding site enhances U6-II stability, likely by binding the Sm complex in vivo. Disruption of post-transcriptional processing events demonstrates the coupling of transcription to snRNA processing, emphasizing the importance of processing fidelity in maintaining snRNP levels in cells and incorporating correctly processed, functional U6 snRNAs into larger subcomplexes and spliceosomes.

In another study, the mechanism of spliceosome activation, particularly the release of U4 snRNA by Brr2 helicase activity and structural rearrangement of U6, is explored. Endogenous fluorescent labeling strategies, such as MS2 or Mango sequences, are compatible with spliceosome function as genetic insertion into the U4 sequence minimally affects cell growth and splicing activity. These findings pave the way for single molecule studies of spliceosome activation where U4 snRNA release can be monitored, providing insights into active site formation.

Chapters 4 and 5 focused on the second transesterification step, exon ligation. Fyv6 (FAM192A in humans) is identified as a second step splicing factor that, when absent, impacts 3' splice site (SS) selection transcriptome wide. Structural analysis reveals Fyv6's unique interaction with Prp22 ATPase and its mutually exclusive binding with the first step factor Yju2. Prp22 has a known role in maintaining fidelity in 3′ SS selection. As 3′ SS fidelity is globally lost when spliceosome lack Fyv6, a functional connection between Fyv6 and Prp22 is suggested that directly impacts interactions of RNAs within the spliceosome active site.

6.2 RNAP transcribed U6 relaxes sequence constraints to allow for incorporation of RNA tags within the U6 5′ stem loop

We were able to show that transcription of the U6 snRNA in yeast is not dependent on RNAP III and can be switched to RNAP II. An advantageous outcome of U6 RNAP II transcription is that the 5′ stem loop (SL) can be replaced with RNA tags for affinity purification or fluorescent labeling. Indeed, we have made variants of U6-II that replace the 5′ SL with the Mango-II RNA aptamer. Yeast expressing U6-II-Mango are viable; however, further characterization on U6-II-Mango variants has not been performed. Mango tags have been previously used as a handle for pulling down RNP complexes

from yWCE and could facilitate the purification of U6-containing complexes. Particularly with U6-II-Sm it would be interesting to examine the constituent components of the U6 snRNP or U4/U6 di-snRNP.

We were able to show by immunoprecipitation (**Fig 2.11**) that U6-II-Sm can be pulled down with both tagged Lsm and Sm proteins, indicating both proteins can bind to U6-II-Sm. We were unable to clarify whether Sm and Lsm rings can simultaneously bind a single U6-II-Sm molecule, although there is theoretically enough space between the binding sequences for both rings to fit. Purification of U6-containing molecules would allow for quantitation of molecular weight by mass photometry, a technique that measures protein mass based on refracted light (Young et al. 2018). U6 complexes containing both an Sm and Lsm ring will be shifted by roughly 77 kDa from a population of U6 complexes containing either the Lsm or Sm ring, allowing for observation of a population of U6•Lsm•Sm should it exist.

6.3 Mango affinity purification of free U4 snRNP in U6-II expressing yeast strains

Structural analysis of free U4 and U5 snRNPs has long been desired as it will show arrangements of protein and RNA components prior to assembly into di- and tri-snRNPs, yielding insights into how these complexes are formed. The low abundance of free U4 and U5 snRNPs in extracts (**Fig. 2.6,2.7**) has made isolation of a homogenous population of particles for cryo-EM difficult. A structure of the human 20S U5 snRNP was very recently published (Schneider et al. 2024; Riabov et al. 2024). However, the structure of the free U4 snRNP in any species has yet to be solved.

In Chapter 2, we showed that transcription of U6 by RNAP II (U6-II) altered the distributions of snRNPs. Fractions of free U4 and U5 snRNPs increased in yWCE where U6-II is expressed, as lower levels of free U6 snRNP were likely limiting to di-snRNP and/or tri-snRNP assembly **(Fig. 2.6,2.7**). Insertion of a tag into the U4 snRNA would facilitate affinity purification of the free snRNP particles in strains expressing U6-II. I have shown in Chapter 3, that insertion of Mango or MS2 tags into the U4 snRNA has been shown to result in few growth defects and are otherwise tolerated well by yeast (**Fig. 3.4**). Given there is an established protocol for RNP purification from Mango tagged snRNAs (Panchapakesan et al. 2017) isolation of free U4 snRNPs for cryo-EM could be accomplished from yeast strains also expressing U6-II.

Yeast expressing U6-II and U4-Mango-III are viable. However, insertion of Mango into the U4 snRNA may further alter snRNA levels in cells. A decrease in the amount of free U4 snRNP would not be advantageous to isolating a large population of homogenous molecules for cryo-EM. An initial primer extension experiment with total RNA isolated from U6-II U4-Mango-III expressing yeast showed that levels of U4 and U6 snRNAs are increased upon expression of U4-Mango-III when compared to levels of the U2 snRNA (**Fig. 6.1;** see Chapter 3 methods). However, ratios of snRNAs in strains expressing U4-

Mango-III still show higher amounts of U4 relative to U6 (**Fig. 6.1;** quantification not shown).

Experiments looking at snRNP distributions are needed to clarify if the free U4 snRNP abundance is maintained in strains expressing U4-Mango-III before cryo-EM is attempted. Pulldowns with U4-Mango-III would isolate all U4-containing complexes including di-snRNP, tri-snRNP, and spliceosomes. Larger complexes may be able to be removed with size exclusion chromatography or computationally eliminated from sets of particles during cryo-EM data processing and refinement. Particle distributions may be able to be easily visualized using mass photometry, instead of more labor-intensive methods such as glycerol gradient sedimentation, to determine if affinity purification with Mango tagged U4 is a viable method for obtaining cryo-EM samples.

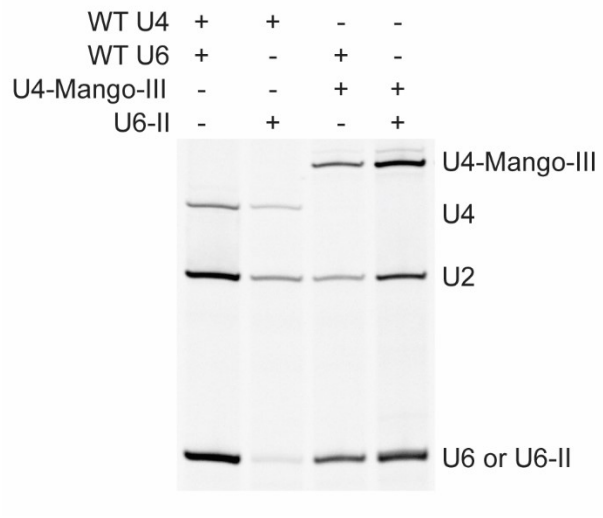

Figure 6.1 Primer extension of U2, U4, and U6 snRNAs from total RNA.

6.4 CoSMoS experiments for examining activation dynamics

Methods for tagging described in Chapter 3 may be applicable for a wide variety of microscope experiments examining spliceosome activation. Initial experiments would involve understanding the dynamics of spliceosomes during activation. Labeling of Prp3 by fusion to a SNAP tag and the U4 snRNA by methods outlined in Chapter 3 would allow for observation of the order of release of these molecules in relation to each other. Measurements of the time from tri-snRNP arrival prior to departure of the U4 snRNA and Prp3 could additionally give information about activation processes, such as delayed or stalled activation. Conditions in WCE could be modified to understand how Brr2 activity during activation is regulated by hydrolysis of GTP by Snu114 (Small et al, 2006; Bartels et al. 2003). Addition of various GTP:GDP ratios to yWCE extracts during CoSMoS

experiments would allow for comparison between conditions when GTP is enriched or depleted, thus allowing for investigation of how lack of Snu114 activity influences activation dynamics. Additionally, the impact of mutants that either stabilize or destabilize U4/U6 di-snRNA Stem I or Stem II regions or competing structures in U6 (telestem, ISL, or interactions with U2 snRNA) on U4 release from the spliceosome could be studied to refine current models of snRNA interactions and remodeling that drive activation.

6.5 Fyv6 and Yju2 exchange between the first and second steps of splicing

Fyv6 spans the spliceosome over 100Å and forms numerous contacts with splicing proteins. Determining a precise mechanism of action for Fyv6 in promoting the second step of splicing has therefore been difficult. The region of Fyv6 that we identified as essential for function and that reproduces the *fyv6*Δ phenotype is the C-terminal helix contacting Syf1 (**Fig 5.12**). Yju2 interacts with Syf1 in an similar location, such that presence of Yju2 is mutually exclusive with Fyv6 (**Fig. 5.9**). Cryo-EM analysis of Yju2 and Fyv6 densities suggests that Yju2 is present in splicing complexes before (B*) and immediately after the first step of splicing (C), whereas Fyv6 is likely present before (C*) and immediately after (P) the second step of splicing. An exchange between Yju2 and Fyv6 during the first to second step is therefore required, but an exact mechanism is unknown.

To examine Yju2 and Fyv6 interactions with Syf1 further, *in vitro* splicing assays can be conducted with *fyv6*Δ splicing extracts supplemented with truncated portions of Fyv6. We have shown that addition of recombinant full-length Fyv6 in splicing extracts lacking Fyv6 can rescue first step splicing defects (**Fig. 6.2A,B**; See Chapter 3 Methods; Fyv6 purification protocol not detailed). An interesting experiment would be to add Fyv6 peptides corresponding to the Syf1 interacting region in splicing extracts. If displacement of Yju2 from Syf1 is sufficient to restore second-step splicing efficiency, then we can more confidently conclude that the block to the second step of splicing is due to a failure to displace Yju2 from Syf1. ATP-dependent helicase activity of Prp16 is responsible for remodeling the spliceosome after the second step and displacing Yju2. However, Yju2 is still likely loosely associated with the spliceosome or present in high enough concentrations that it can readily rebind once displaced. Rebinding of Yju2 might stabilize first step conformations, however binding of Fyv6 after Yju2 displacement would prevent Yju2 from rebinding. A caveat to adding a Syf1-interacting Fyv6 peptide to splicing extracts may be that Yju2 is unable to bind Syf1 initially, likely resulting in a reduction in first and second-step splicing efficiencies, or that larger peptides are required to see an effect.

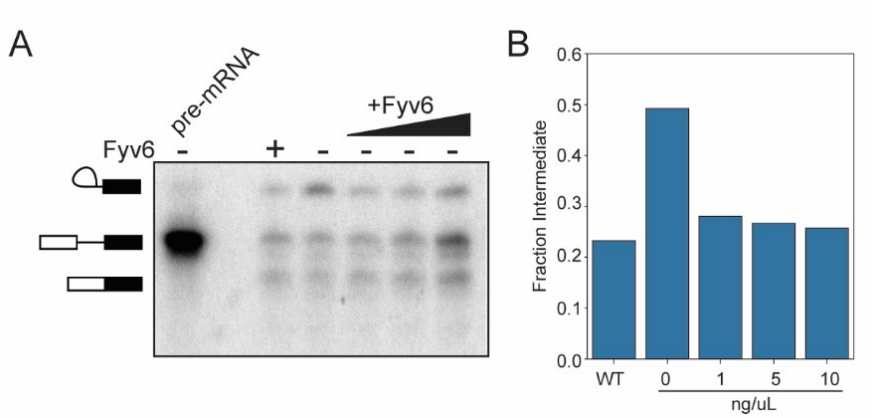

Figure 6.2 Addition of recombinant Fyv6 protein to extracts rescues intermediate accumulation in *fyv6Δ* strains. A) Gel of splicing products with addition of increasing amounts of recombinant Fyv6. B) Quantitation of splicing intermediates.

The Yju2-Fyv6 exchange could also be examined genetically. Yeast expressing C-terminal truncations of Yju2 are viable, but growth is slowed and completion of the second step becomes very inefficient (Chiang and Cheng, 2013). Growth phenotypes at different temperatures were not characterized with the Yju2 C-terminal truncations. Since yeast are viable with C-terminal Yju2 truncations, it would be interesting to express Yju2 truncations in a *fyv6*Δ background. We have some evidence already that absence of the Syf1-Yju2 interaction allows for some bypass of growth phenotypes seen in *fyv6*Δ. Specifically, truncation of the Syf1 regions before and after the Fyv6-interaction region (*Δ817-859* and *Δ778-859*) rescue the heat sensitive *fyv6*Δ growth phenotype (**Fig. 5.12**). We might expect to see a similar rescue in the temperature-sensitive growth phenotype when the C-terminal portion of Yju2 that interacts with Syf1 is truncated, allowing Syf1 to reposition between the two catalytic steps. In strains where Fyv6 is present, truncation of Yju2 may also lead to the premature recruitment of Fyv6 potentially destabilizing the first step conformation of the spliceosome. Both Yju2 and Fyv6 may play similar stabilizing roles for the first and second steps of splicing, respectively, where they mediate the positioning of NTC components.

However, in addition to the mutual exclusivity of the Fyv6 C-terminal and Yju2 C-terminal interactions with Syf1, the center regions of these two elongated, helical proteins are located perpendicular to each other when comparing C to P complexes where Yju2 or Fyv6 are respectively present (Wan et al. 2019; Chapter 5). Further truncations of Yju2 beyond the Syf1-interaction C-terminal region may therefore be needed to functionally bypass any stabilization of first step spliceosome conformations by Yju2, particularly over the active site which is remodeled between the first and second steps of splicing.

6.6 Evidence suggests regulation of Prp22 by Fyv6

Fyv6 extensively contacts the Rec A2 domain of Prp22 in our P complex structure (**Fig. 5.9B**). Despite removal of these regions only having a modest effect on cold sensitivity of yeast, we have shown additional connections between these two proteins including the rescue of accumulated *SUS1* pre-mRNAs in a release defective Prp22 mutant (G810A; **Fig. 5.17D,E**) by deletion of FYV6 and identification of a mutation in Prp22 that suppresses *fyv6*Δ growth phenotypes (**Fig. 5.14**). There are multiple parallels between Prp22 and Fyv6 including arrival at the spliceosome during approximately the same step (C* complex) and the dependency of both Prp22 and Fyv6 when BP-3′ SS distances are greater than 21 nt (**Fig. 5.3;** Schwer and Gross, 1998). Despite genetic and physical interactions between Prp22 and Fyv6, the functional relationship remains largely ambiguous.

One region of Fyv6 that interacts with Prp22 that we were not able to probe was the second α-helix that bridges over the spliceosome active site. Truncation of the second α-helix was not feasible as it would result short, likely poorly expressed protein, complicating phenotypic analysis. We had made several single point mutations within this region but found no effect on the temperature or cold sensitive phenotype (data not shown). Interestingly, one of the mutations we chose was directly adjacent to the Prp22 mutation identified in the suppressor screen. The Prp22 R1134A mutation we tested had no effect on temperature or cold sensitivity in yeast lacking Fyv6 (data not shown) in contrast to the I1133R mutation that rescued both growth phenotypes (**Fig. 5.17C**). In yeast with Fyv6 present, we found that I1133R reduced splicing of a WT ACT1-CUP1 reporter (**Fig 5.17F,G**). The presence of two adjacent arginine residues may hyper-stabilize interactions of the Prp22 C-terminal insertion within the active site. Expression of the combined mutations I1133R and R1134A in yeast may rescue splicing defects of the WT ACT1-CUP1 reporter seen in I1133R alone since only a single arginine is present instead of two adjacent arginines.

It is worth noting that the Prp22 C-terminus is enriched in positively charged residues, with many positions highly conserved across species (**Fig. 6.3**). Charged residues in this region could mediate interactions of the Prp22 C-terminus with the U2-intron helix near the active site. Mutation of all charged residues to alanine in the C-terminus of Prp22 may result in lethality, demonstrating that charged residues in this region of Prp22 are likely required for function.

The impact of Fyv6 on Prp22 proofreading activity is an area of ongoing research in the Hoskins Lab. While experimental evidence with nonconsensus ACT1-CUP1 reporters (**Fig. 5.17F,G**) and upregulation of nonconsensus SS in RNA-seq datasets (**Fig. 5.3H**) suggests a link between Fyv6 and Prp22 proofreading, the exact mechanism is unclear. Additionally, the contribution of the Prp22 C-terminus to the proofreading has not been examined. Finally, a mechanism explaining the decrease in splicing efficiency seen for pre-mRNA substrates with longer BP - 3′ SS distances for spliceosomes lacking Prp22,

Fyv6, or other second-step splicing factors is lacking. Future research is needed to characterize interactions of second-step factors with the active site and the impact on different splicing substrates.

Figure 6.3 Conserved charged regions of Prp22 may be functionally important for interactions with Fyv6 and proofreading activities. Alignment of Prp22 sequences across species highlighting the quantity and conservation of positively charged residues within the C-terminus.

6.7 Conclusions

The exploration of RNA-protein interactions in various facets of splicing elucidates key mechanisms underlying the intricate process of pre-mRNA maturation. The ability to transcribe U6 snRNA with RNAP II not only broadens the understanding of snRNP synthesis but also presents a novel avenue for incorporating RNA tags for affinity purification or fluorescent labeling, facilitating the study of splicing complexes. Furthermore, the investigation into the exchange dynamics between Fyv6 and Yju2 during splicing steps sheds light on their roles in spliceosome assembly and function, offering insights into the regulatory mechanisms governing the transition between splicing stages. Additionally, the proposed rip-cord model of Prp22 regulation by Fyv6 unveils a potential mechanism for fine-tuning spliceosome activity, highlighting the intricate interplay between RNA-protein interactions and splicing fidelity. These findings underscore the complexity of splicing regulation and provide a foundation for further research into the molecular mechanisms driving accurate pre-mRNA processing.

6.8 References

Bartels C., Urlaub H., Lührmann R., Fabrizio P. 2003. Mutagenesis suggests several roles of Snu114p in pre-mRNA splicing. *J Biol Chem.* **278**:28324–28334. doi: 10.1074/jbc.M303043200

Chiang TW, Cheng SC. 2013. A Weak Spliceosome-Binding Domain of Yju2 Functions in the First Step and Bypasses Prp16 in the Second Step of Splicing. Molecular and *Cellular Biology* **33**(9), 1746–1755. doi: 10.1128/MCB.00035-13

Gonzalo MOI, Barrass DJ, Maudlin IE, Beggs JD. 2019. Blocking Late Stages of Splicing Quickly Limits Pre-Spliceosome Assembly in Vivo. *RNA Biology* **16(**12): 1775–84. doi: 10.1080/15476286.2019.1657788.

Panchapakesan SSS, Ferguson ML, Hayden EJ, Chen X, Hoskins AA, Unrau PJ. 2017. Ribonucleoprotein purification and characterization using RNA Mango. *RNA* **23**(10):1592-1599. doi: 10.1261/rna.062166.117

Riabov Bassat, D., Visanpattanasin, S., Vorländer, M.K. et al. 2024. Structural basis of human U5 snRNP late biogenesis and recycling. *Nat Struct Mol Biol*. doi: 10.1038/s41594-024-01243-4

Schneider S, Brandina I, Peter D. et al. 2024. Structure of the human 20S U5 snRNP. *Nat Struct Mol Biol* doi:10.1038/s41594-024-01250-5

Schwer B, Gross CH. 1998. Prp22, a DExH-box RNA helicase, plays two distinct roles in yeast pre-mRNA splicing. *EMBO J* **17**: 2086–2094. doi: 10.1093/emboj/17.7.2086

Small EC, Leggett SR, Winans AA, Staley JP. 2006. The EF-G-like GTPase Snu114p regulates spliceosome dynamics mediated by Brr2p, a DExD/H box ATPase. *Mol Cell.* **23**:389–399. doi: 10.1016/j.molcel.2006.05.043

Wan R, Bai R, Yan C, Lei J, Shi Y. 2019. Structures of the Catalytically Activated Yeast Spliceosome Reveal the Mechanism of Branching. *Cell* **177**(2): 339-351.e13. doi: 10.1016/j.cell.2019.02.006.

Young G, Hundt N, Cole D, et al. 2018. Quantitative mass imaging of single biological macromolecules *Science* **360**:423–427 doi: 10.1126/science.aar5839