Characterizing the Role of p14 in Pre-mRNA Splicing

by

Sayed Madi

A thesis submitted in partial fulfillment of the requirements for the degree of

Master of Science

Department of Biochemistry

University of Alberta

ABSTRACT:

Splicing is an essential step in the post-transcriptional modification of pre-mRNAs in eukaryotes. This process is catalyzed by the spliceosome, a multi-megadalton nuclear protein-RNA enzyme which excises non-coding sequence from RNA transcripts prior to translation. Deficient or improper splicing of introns in humans has been linked to many morbidities including myelodysplastic syndromes (MDS) and retinitis pigmentosa.

SF3b14a/p14 is an evolutionarily conserved 14 kDa protein within the spliceosome that has been shown to crosslink to a highly conserved nucleotide at the 3′ end of introns that performs the first step of splicing. To date, the exact function that p14 plays in splicing remains elusive. Interestingly, siRNA knockdown of p14 in HeLa cells results in cell death within a few days, therefore making it an essential gene in humans.

I have used the model organism *Schizosaccharomyces pombe* to study the role of p14 in pre-mRNA splicing. This was possible because the *S. pombe* p14 knockout is viable. Therefore, by investigating differences between wild-type and p14 knockout cells, I hoped to gather new insights into protein function. To do this, I utilized a combination of stress tests, morphological studies, and mass spectrometry.

My research has shown that p14 may play an important role in the recognition and subsequent splicing of a certain subset of introns. This was supported by mass spectrometry data showing a downregulation in proteins required for organo-nitrogen compound metabolism and mitosis in p14 knockout cells.

Dedication

This thesis is dedicated to my mother, father, and brother who have supported me to no ends during my academic endeavors. Without their kindness, patience, and wisdom this would not have been made possible.

Table of Contents

List of Tables:

List of Figures:

List of Abbreviations:

Δ	Deletion
α	Alpha
3D	Three dimensional
Å	Angstrom
A	Adenosine
aa	Amino acid
Amp^R	Ampicillin resistance
ATP	Adenosine triphosphate
β	Beta
bp	Base-pair
BPD	Beta-propeller domain(s)
BS	Branch Sequence
BPA	Branch point adenosine
C	Cytosine
cDNA	Complementary DNA
CLL	Chronic lymphocytic leukemia
CNBr	Cyanogen bromide
CRL3A	Cullin4A-RING E3 ubiquitin ligase
CTD	C-terminal domain
Cys	Cysteine
ddw	Double distilled water
DEPC	Diethyl pyrocarbonate
DNA	Deoxyribonucleic Acid
ds	Double Stranded
DTT	Dithiolthreitol
E complex	Early complex
EDTA	Ethylene diamine tetra-acetic acid
EM	Electron microscopy
G	Guanine
G418	Geneticin
HA	Hemagglutinin
HeLa	Henrietta Lacks
HRP	Horseradish Peroxidase
HU	Hydroxyurea
Kan	Kanamycin
kb	Kilobases
kDa	Kilo-dalton
MDS	Myelodysplastic syndromes
mRNA	Messenger RNA
nt	Nucleotide
MMS	Methyl methanesulfonate
MPB	Maltose binding protein
NTD	N-terminal Domain
OD	Optical Density
ORF	Open reading frame

PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate-buffered saline
PCR	Polymerase chain reaction
PMSF	Phenylmethylsulfonyl fluoride
PPT	Polypyrimidine tract
RRM	RNA recognition motif
RNA	Ribonucleic Acid
RNP	Ribonucleoprotein
RPM	Revolutions per minute
RT PCR	Reverse transcription polymerase chain reaction
sap	Spliceosome-associated protein
SCI	Single colony isolates
SDS	Sodium dodecyl sulfate
SF	Splicing factor
SnRNP	Small nuclear ribonucleoproteins
SnRNA	Small nuclear RNA
ss	Single Stranded
SS	Splice Site
SUGP1	SURP and G-patch domain containing 1
TE	Tris-EDTA
U	Uracil
U2AF	U2 Auxiliary Factor
UV	Ultraviolet
WT	Wild-type
y	Any pyrimidine base
ZnF	Zinc finger domain

Chapter 1
Introduction to pre-mRNA splicing

I-1.1: Introduction.

Gene regulation in eukaryotes has extensively been studied over the years. However, it was not until the 1970s that the Roberts and Sharp laboratories independently discovered the "split gene" structure while mapping the adenovirus genome [1, 2]. Later findings suggested that precursor messenger RNA (pre-mRNA) found in eukaryotic nuclei consisted of protein coding sequences (exons) that were interrupted by noncoding sequences (introns) [3-5]. Prior to translation, intron excision and exon-exon ligation are required to generate messenger-RNA (mRNA) which interacts with ribosomes in the cytoplasm initiating protein synthesis. This process, called splicing, is carried out by the spliceosome and is believed to occur in more than 95% of human pre-mRNA gene products [6-8].

The spliceosome is a multi mega-Dalton protein-RNA complex found in the nucleus which recognizes and interacts with key conserved elements within introns. Through these interactions, the spliceosome completes two sequential transesterification reactions that result in the removal of introns and ligation of neighboring exons [9, 10]. While introns can vary greatly in size and sequence, they are defined by specific conserved elements found at both the 5′ and 3′ ends (Figure 1-1A). At the 5′ end, introns are defined by a 5′ splice site (5′ SS) consisting of a GU dinucleotide. At the 3′ end there are several regions including a branch sequence (BS), polypyrimidine tract (PPT), and a YAG trinucleotide sequence [11-15]. The consensus branch point sequence in humans is "yUnAy", which is highly variable but still contains a well conserved branch point adenosine (BPA) which is of great significance and will be discussed later [16].

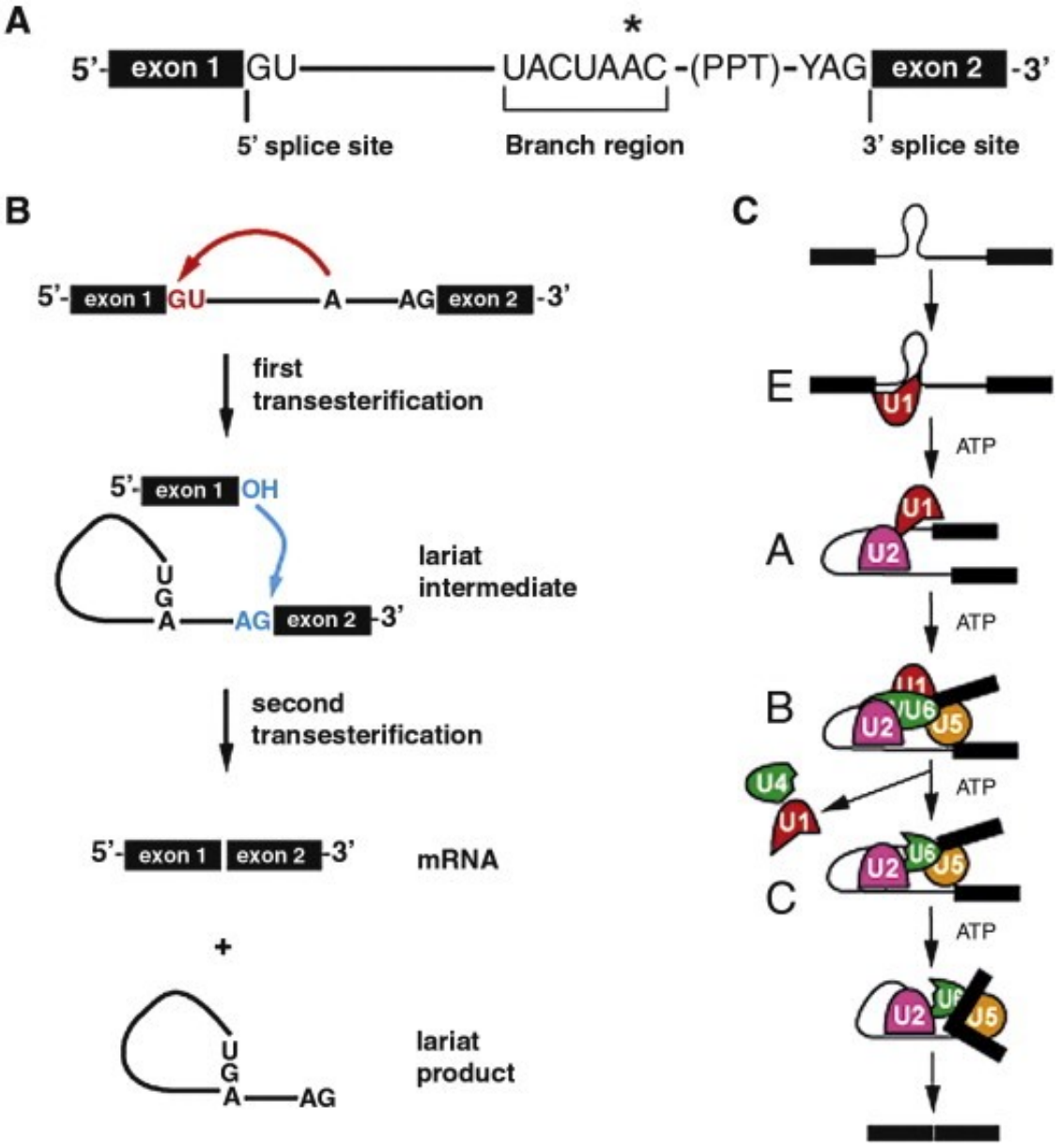


Figure 1-1. Mechanism of pre-mRNA splicing. A) Intron sequence flanked by two exon sequences. Highlighted are the 5′ SS, 3′ SS, and consensus branch sequence followed by the PPT. The conserved branch adenosine is denoted by (*) B) Two step ATP-independent transesterification reaction initiated by the branch adenosine. C) Stepwise assembly of the spliceosome through complexes E-C. The different snRNP particles (U1,2,4,5,6) are seen interacting with the pre-mRNA at different stages of spliceosome assembly. (Adapted from Ritchie *et al*., 2009, *Biochim. Biophys. Acta.* 1789, 624-633.)

The polypyrimidine tract is a string of pyrimidines that spans 4-24 nucleotides downstream of the BS while the "YAG" trinucleotide sequence is found at the terminal end and marks the intron-exon boundary. Splicing occurs in a step-wise fashion when the spliceosome recognizes boundaries between introns and exons. The first step in splicing occurs when the highly reactive 2' hydroxyl group of the branch point adenosine performs a nucleophilic attack at the 5′ splice site forming a free 5′ exon and a lariat intermediate with a 2-5′ phosphodiester linkage. The second step occurs when the hydroxyl group of the 5′ exon performs a nucleophilic attack at the 3′ SS which results in the excision of the lariat intermediate and the subsequent ligation of adjacent exons (Figure 1-1B) [9, 17, 18]. It was originally thought the excised intron had no function and was degraded shortly after; however, recent studies have shown these intron may act as precursors for further processing of other RNAs [19].

I-2.1: Components of the spliceosome.

The spliceosome is an incredibly dynamic macromolecular machine which consists of over 200 individual components which assemble on introns in a step-wise fashion [20]. Many of these components join the spliceosome in protein:RNA sub-complexes called small nuclear ribonucleoproteins (snRNPs) (Figure 1-2). These snRNPs are named: U1, U2, U4, U5, and U6 based on the small nuclear RNA (snRNA) that they contain. As mentioned earlier, successful splicing requires the spliceosome to assemble in a correct order of sequence which is guided by the conserved elements within introns. Since spliceosome vary in different kingdoms, the following discussion will focus specifically on the mechanism by which splicing occurs in metazoans.

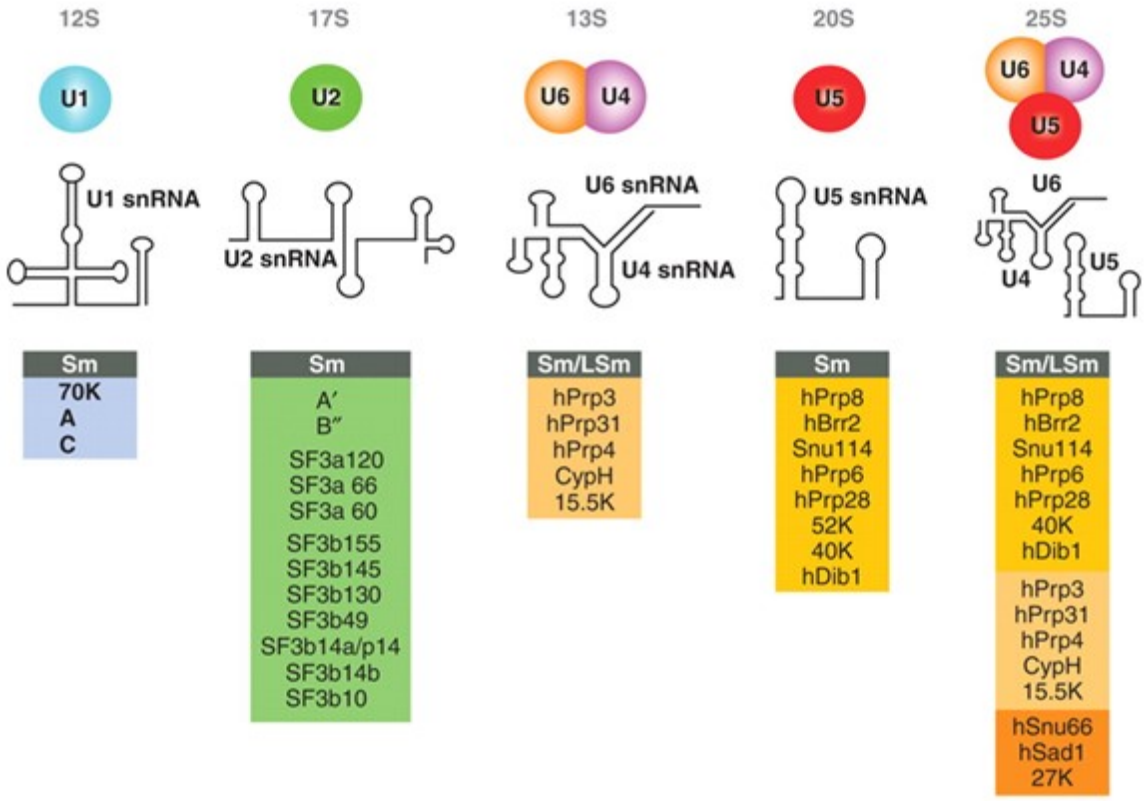


Figure 1-2. Protein composition and RNA secondary structures of spliceosomal snRNPs. Listed below are the proteins that associate with each of the respective snRNPs. Included are the seven Sm core proteins (B/B', D3, D2, D1, E, F, and G) and LSm proteins (Lsm2-8). Note the U4/U6 snRNPs exist associated with one another via base-pairing interactions between their small RNAs. (Adapted from Will *et al*., 2011, *Cold Spring Harb Perspect Biol*;3:a003707)

Following the transcription of genomic DNA by RNA Polymerase II, a pre-mRNA transcript is produced containing both introns and exons While in the nucleus, intron sequences within the immature transcript are recognized by the U1 snRNP via the U1 snRNA which forms base pair interactions at the 5′ SS [21, 22]. This process is ATP independent and results in the formation of the E complex (early). This is followed by the ATP-dependent interaction of U2 snRNP at the 3′ SS forming the A complex. However, prior to this occurring, U2 Auxiliary Factor (U2AF) and splicing factor 1 (SF1) must first bind. These proteins interact at the polypyrimidine tract/3′ splice

site and BS respectively [23]. Important to note, when the U2 snRNA base pairs with the BS a single adenosine from the intron is bulged outwards due to the lack of a complementary base interaction. As a result, this adenosine is selected to carry out the first of two transesterification reactions in the splicing process [24, 25]. The U2 snRNP associated heptameric protein complex SF3b plays a role in stabilizing the interaction of U2 snRNP to the 3′ region of introns, and presents a temporary steric barrier to the branch point adenosine preventing premature splicing [26]. The U4/U6:U5 tri-snRNP is then recruited forming the pre-catalytic B complex. Upon major RNA:RNA and protein:RNA rearrangements, U1 and U4 snRNPs are dissociated from the pre-mRNA giving rise to the activated spliceosome (complex B*). The activated spliceosome carries out the first catalytic step of splicing generating free exon 1 and the intron–exon 2 lariat intermediate, which is now referred to as the catalytic spliceosome (C complex). Further rearrangement of the C complex occurs and the second step of splicing is performed shortly afterwards giving rise to the post spliceosomal complex, which contains the now excised lariat intron and ligated exons (Figure 1-1C) [27-29]. Finally, the U2, U5, and U6 snRNP are recycled for additional rounds of splicing by the DExD/H helicase Prp22 [30].

I-2.2: The SF3b complex.

Following the discovery of splicing in 1977, there has been great interest in characterizing the biochemical and catalytic properties of the spliceosome. A wealth of literature has surfaced aiming to elucidate important proteins and factors required for this complex process. Early studies using chromatographic fractionation procedures in HeLa cells yielded many complexes required for splicing including splicing factors 1-4 (components of HeLa cell nuclear extracts required for pre-

messenger RNA splicing) [31]. The complex SF3 consists of two sub-complexes called SF3a and SF3b, both of which are components of the U2 snRNP and are crucial for splicing [32-34]. The SF3b sub-complex consists of seven subunits: SF3b155, SF3b130, SF3b145, SF3b49, SF3b14a/p14, SF3b14b, and SF3b10 (Figure 1-3A/B) [35-37]. Each of the proteins contain distinct domains which aid in their interaction with other proteins and/or pre-mRNA. It has been shown that SF3b interacts with the pre-mRNA near the BS, stabilizing RNA:RNA interactions [35, 38]. Additional roles of SF3b may include BS selection during the splicing process. The largest protein SF3b155, also called splicing associated protein 155 (sap155), contains an unstructured N-terminal domain (NTD) and a C-terminal 20 HEAT-repeat domain (HD) each consisting of two alpha helices linked by a short loop [39]. Additionally, the NTD contains a short region (373-415) shown to interact with another SF3b protein, p14, and a UHM-ligand motif (ULM) which facilitates interaction with U2AF [40-42]. Together, SF3b155, p14, and U2AF are thought to work jointly in recognizing the BS within pre-mRNA. SF3b145, also referred to as sap145, is the second largest protein in the complex does not contain any predicted structural domains but rather proline rich regions near the N and C termini. SF3b130 is composed of three intertwined seven-bladed β-propeller (BP) domains: BPA, BPB, and BPC followed by an alpha helical C-terminal domain which contacts the other domains. The overall fold of this protein is analogous to that of DDB1, a protein scaffold and core component of the cullin4A-RING E3 ubiquitin ligase complexes (CRL3A) [41, 43, 44]. The similarity in structures translates to how these proteins form interactions with their partners. In the case of SF3b130, its interaction with SF3b10 occurs via a clam-shaped cavity created by the BPA and BPC domains. SF3b49 is comprised of two N-terminal RNA Recognition Motifs and a C-terminal proline rich domain. SF3b10 who's interaction with

SF3b49 was discussed earlier does not have any common structural features. SF3b14a and SF3b14b are two distinct 14 kDa proteins that contain an RRM and a zinc finger domain, respectively (Figure 1-3C) [41].

I-2.3: **Mutations within SF3b155 lead to use of cryptic 3′ SS.**

As of late, SF3b155 has piqued the interest of many researchers as cancer genome sequencing studies have shown mutations in SF3b155 to occur frequently in patients diagnosed with myelodysplastic syndromes (MDS), chronic lymphocytic leukemia (CLL), uveal melanoma, and breast cancer [45-48]. Interestingly, nearly all of these mutations are heterozygous missense substitutions which map to the HEAT domain of SF3b155. Specifically, 33 of them occur between heat repeats H4-H7, and 9 of them between H4-H8, including the most common mutation, K700E [41]. Experiments isolating spliceosomal complexes from HEK293T cells containing K700E mutations revealed significant downregulation of the spliceosomal protein SUGP1 (SURP and G-patch domain containing 1) [49]. This translated to an increase in aberrantly spliced transcripts due to cryptic upstream 3′ SS selection. Interestingly, SUGP1 knockouts resulted in the same splicing errors seen in the K700E mutants; however, this defect could be rescued by overexpression of SUGP1[49, 50].

For adequate recognition of canonical branch sites, SUGP1 must first interact with the HEAT domain of SF3b155. Through this interaction, the U2 snRNP is recruited to the canonical 3′ SS as SUGP1 directly interacts with SF1 and U2AF. Following this interaction, the SUGP1 G-patch domain activates an RNA helicase required for the displacement of SF1 via p14 which allows the

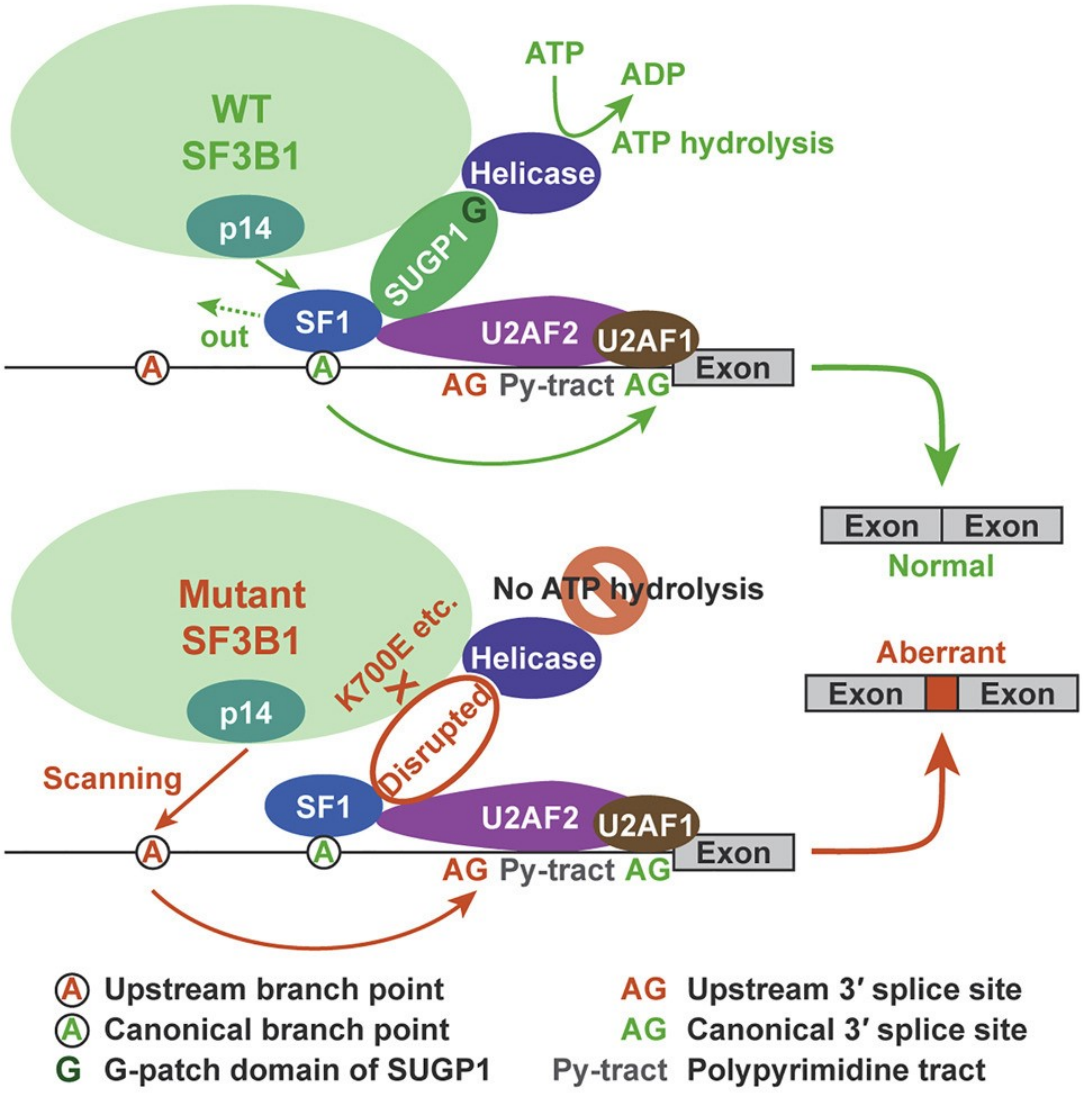


Figure 1-4. Mutation in SF3b155 results in selection of cryptic 3′ SS. Shown on top is the normal mechanism of splicing in which SUGP1 interacts with the HEAT domain of SF3b155. This allows recruitment of the U2 snRNP to the canonical 3′ SS. Thereafter, p14 aids in the displacement of SF1 revealing the adenosine nucleophile required for the first transesterification reaction. Shown below is the mechanism of aberrant splicing resulting from mutations within the HEAT domain of SF3b155 and/or the SUGP1 G-patch domain. As a result, SF1 does not get displaced leading to the U2 snRNP searching for an upstream branch point which leads to a cryptic 3′ SS selection. (Adapted from Zhang *et al*., 2019, *Molecular Cell* 76, 82–95)

U2 snRNA to base-pair with the canonical branch site. In the absence of SUGP1 or mutation within the G-patch domain, SF1 is not displaced and as a result blocks the canonical branch point. This likely leads to U2 snRNP searching for an alternative branch sequence upstream yielding cryptic 3′ SS selection (Figure 1-4) [49]. This mechanism is quite interesting and gives new insight into one possible way by which p14 aids in the selection of canonical BS. More on this idea is discussed in the next section.

I-3.1: Discovery of p14.

The global mechanism of splicing has been heavily studied over the years including attempts to piece together the proteins and factors required at each step. Due to the size and complexity of the spliceosome it has been rather difficult to come up with a precise mechanism. An area that has puzzled researchers over the years revolves around how the branch region is recognized within the pre-mRNA. As mentioned earlier, the branch region is found near the 3′ SS and contains a conserved adenosine residue which acts as a nucleophile for the first transesterification reaction. Furthermore, studies have shown that the U2 snRNA forms base-pairing interactions at the branch region resulting in the bulging and "activation" of the adenosine residue. However, protein interaction at the branch region had not yet been determined. In the early 1990s a set of cross-linking experiments were performed to study potential protein-RNA interactions. To study this, pre-RNAs were synthesized containing chemical modifications in key sequences. The first of which being a benzophenone moiety attached to the N6 position of the branch adenosine via a 15Å tether (Figure 1-5A) [25]. This would allow for the cross-linking of any protein within a 15Å distance from the adenine base. The second modification was a single ^{32}P radio-label 3′ of the

modification. Both wild-type and branch sequence mutant RNA were synthesized, incubated with splicing HeLa nuclear extract for varying times in the presence or absence of ATP, crosslinked on ice by irradiation with a 302-nm lamp and later treated with RNase A and/or RNase T1. The products were run on an SDS-PAGE gel to study protein-RNA interactions.

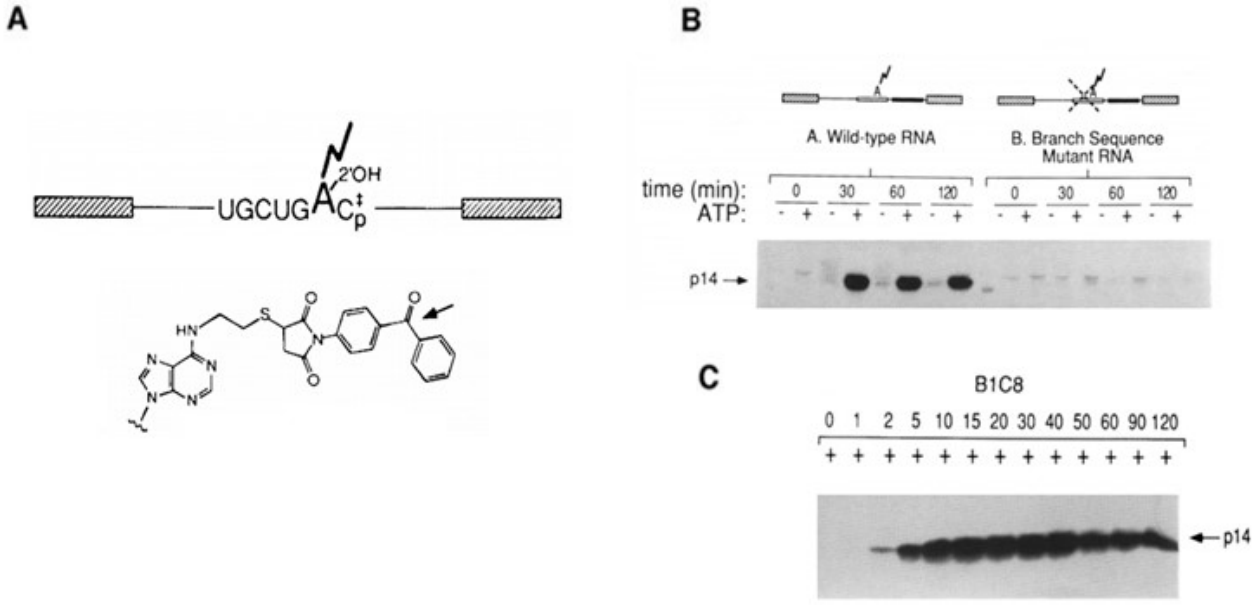


Figure 1-5. Discovery of SF3b14a. A) Synthetic pre-mRNA used to cross-link with SF3b14a. [top] Branch sequence containing conserved adenosine nucleotide tethered to a 15 Å benzophenone modification protruding outward (bold line). ^{32}P-labeled phosphate is attached to the adjacent C nucleotide denoted, **I**. The 2'-OH indicated is the branch nucleophile for the first transesterification reaction. The hatched boxes indicate exon sequences interrupted by the intron. [bottom] Benzophenone maleimide moiety tethered to adenine base via a thiol group. Arrow indicates location where protein:RNA crosslink is expected to take place. B) Time course of cross-linking experiments using WT [lanes 1-8] and branch sequence mutant RNA[lanes 9-16] in the presence of absence of ATP, visualized by autoradiography. C) Time course of cross-linking experiment using WT type RNA later immunoprecipitated by B1C8 antibody to improve specificity of interaction between RNA and spliceosomal proteins. (Adapted from MacMillan *et al*., 1994, *Genes & Dev*. 8: 3008-3020).

The results of these experiments showed that in the presence of ATP the benzophenone-derivatized branch substrate cross-linked to two proteins which ran at 14 and 220 kDa (Figure 1-5B) [25]. To further validate that these interactions occur in the context of splicing, substrates were immunoprecipitated using monoclonal antibodies (mAB) Y12 and B1C8. The Y12 mAB reacts with epitopes on the core SM proteins of the snRNP particles. The B1C8 mAB reacts to members of the SR family of polypeptides [51, 52]. Therefore, immunoprecipitation using these antibodies helped reduce background of cross-links not associated to splicing (Figure 1-5C). Furthermore, treatment with proteinase K resulted in the disruption of protein-RNA interaction indicating that it is not possible that they are solely RNA-RNA in nature. Follow-up cross-linking experiments were performed using an 8-azido adenine modification in the branch position, which is an ~2Å tether [53, 54]. Once again this showed there was a 14 kDa protein associated with the branch adenosine in the A through C complex, specifically at the same stages that the branch adenosine is bulged due to the U2 snRNA base-pairing [55].

To isolate the 14 kDa branch site protein, adenovirus pre-mRNA containing biotin tagged uridines were incubated with spliceosomes. This was followed by the addition of streptavidin-agarose beads that bind biotinylated RNA including any associated proteins. Proteins were subsequently eluted and separated via SDS-PAGE. Four proteins were seen in the 14 kDa size range and each were subjected to peptide sequencing. Three of the proteins had previously been identified (the 20 kDa cap-binding protein, the SmD proteins and a streptavidin contaminant from the column) (Figure 1-6A). However, the fourth protein had not yet been characterized and yielded novel protein sequences: KITAEEMYDIF and KYGPIRQIRVGNTPETRG [56]. Using human

expressed sequence tag EST libraries the researchers matched the peptide sequence to cDNAs found in varying tissues. The protein encoded by the longest EST was found to be 125 amino acids in length and was predicted to be 14.6 kDa in size (Figure 1-6B). Antibodies raised against this cDNA-encoded protein immunoprecipitated spliceosomes, U2 snRNPs, and under denaturing conditions a spliceosome associated protein, later called p14 (SF3b14a) [56].

p14 is an evolutionarily highly conserved protein with orthologs found in many eukaryotic species (Figure 1-6C). Furthermore, as discussed earlier, it is one of the components of the heptameric SF3b complex and has been shown to directly interact with SF3b155, and the pre-mRNA branch adenosine.

To date, the exact role that p14 plays in splicing is not well understood. However, it appears to be an essential in HeLa cells, as siRNA knockout of the gene coding for p14 renders the cells non-viable after a few days (Schellenberg and MacMillan, unpublished). Interestingly, p14 is not essential for survival in the fission yeast *Schizosaccharomyces pombe*, making it an ideal model organism for further study. While most eukaryotes contain a gene encoding for the 14 kDa peptide, some divergent species such as *Saccharomyces cerevisiae* do not. The protein snu17p was originally proposed as a possible homolog for p14 in this organism based on its similarity in sequence, association with the U2 snRNP, and the absence of it's ortholog in humans, however, this was later shown to be untrue [57]. The aim of the research project described here was to characterize the role of p14 in pre-mRNA splicing.

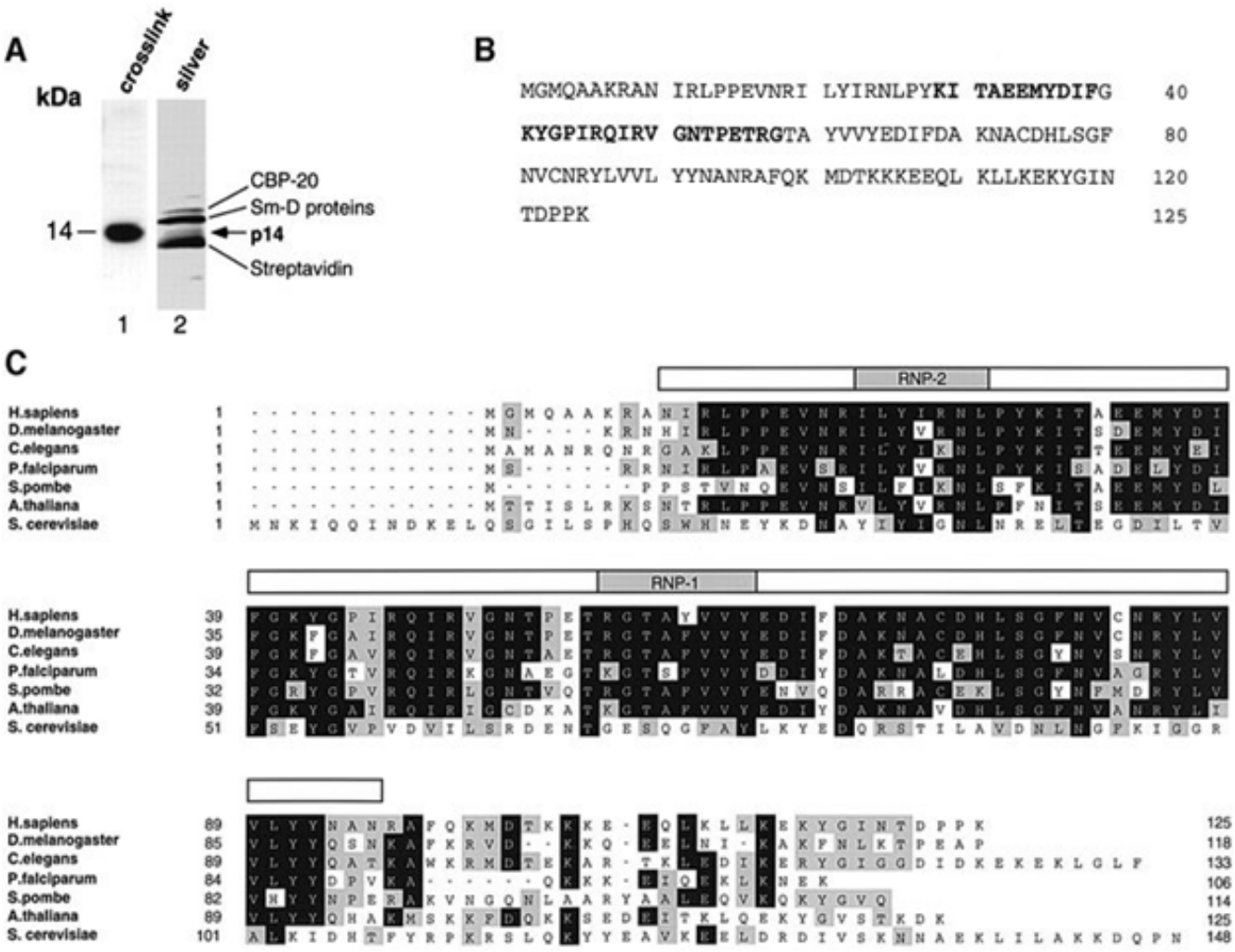


Figure 1-6. p14 contains an RRM that is highly conserved amongst eukaryotes. A) Comparison between branch site cross-linked p14 [lane 1] and purified spliceosome separated on a 20% gel by SDS–PAGE and stained with silver. B) Amino acid sequence of human p14 obtained via microsequencing. C) Amino acid sequence alignment of human p14 against p14 orthologs from different species. Black boxes indicated residues identical in at least four species. The white box overlying the sequences encompasses the RRM which include the RNP-2 and RNP-1 motifs (interrupting grey boxes). (Adapted from Will *et al*., 2001, *EMBO J* 20:4536-4546).

I-3.2: Structure of p14:SF3b.

Comparison of the amino acid sequence of human p14 against putative p14 orthologs from twelve other species provides interesting insight. The central feature of p14 is a highly conserved RRM domain with N- and C-terminal extensions of varying lengths [56]. Furthermore, X-ray protein crystallography structures of p14 bound to a small region of SF3b155 have been determined from several species [42]. Analysis of these structures provides us with clues regarding the function of p14.

Early attempts to express p14 in isolation were unsuccessful. The protein demonstrated poor solubility and aggregated quite easily (Schellenberg and MacMillan, unpublished). Pulldown experiments using several deletion constructs revealed a strong interaction between p14 and the amino acids 373-415 of SF3b155 (Figure 1-7A) [42]. To obtain a soluble complex, SF3b155 peptide was expressed and incubated with MPB tagged p14 followed by cleavage of the fusion protein. The resulting complex was eluted via gel filtration at the volume expected for a 1:1 complex.

Subsequently, the crystal structure was solved at a resolution of 2.5 Å. Analysis of the 3D structure confirms the previously predicted RRM domain of p14 spanning amino acid residues 20-91. This domain contains the classic four anti-parallel β -strands packed against two α-helices (Figure 1-7B). Furthermore, the C-terminus of p14 also contains two additional α-helices: the first being a 5 aa helix (94-98), and a 15 aa helix (103-117).

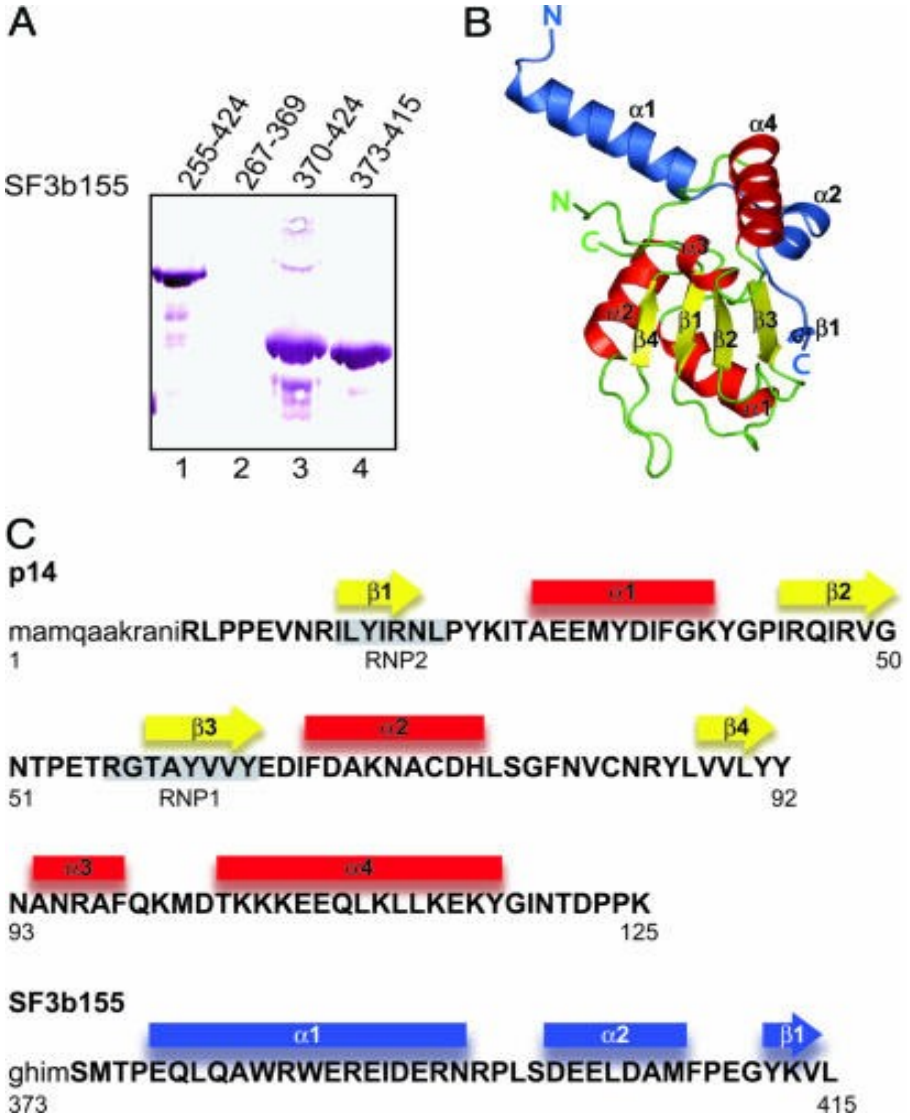


Figure 1-7. p14:SF3b155 peptide complex. A) SDS-PAGE analysis showing the minimal interaction surface between p14 and SF3b155. B) Ribbon diagram of p14:SF3b155 peptide complex. α-helices and β-strands of p14 are indicated by the colors red and yellow, respectively. The loops connecting the secondary structures of p14 are shown in green. SF3b155 peptide (373-415) is colored blue in its entirety. C) Amino acid sequences of p14 and the SF3b155 peptide. Secondary structures found in p14 including α-helices and β-strands are colored yellow and red, respectively, with grey boxes surrounding the amino acids that make up the RNP1 and RNP2 motifs. Secondary structures found in the SF3b155 peptide are shown in blue. (Adapted from Schellenberg *et al*., 2006, *PNAS*, 103 (5) 1266-1271).

The secondary structures found in the SF3b155 peptide include two α-helices: one long N-terminal helix (380-396), and a second shorter helix (401-407). Also found at the C-terminus of the SF3b155 peptide is a short β-strand which interacts with the β-3 strand of the p14 RRM β-sheet. Furthermore, the loop that connects the short β-strand to the α-helix forms contacts with the shorter C-terminal helix and RRM of p14. Perhaps the most remarkable finding revealed by the structure was that the beta-sheet of p14 was highly obstructed by the C-terminal helices of p14, the central helix of the SF3b155 peptide, and the loop connecting SF3b155 to the C-terminal β-strand (Figure 1-7C). This is significant because the 4 β-strands making up the β-sheet represent the canonical RNA-binding surface, which contain the highly conserved RNP1 and RNP2 motifs. These motifs contain residues that form interactions with single-stranded RNA [58-60]. In the context of the p14:peptide complex a small pocket exists on the otherwise occluded surface of the protein where a portion of RNP2 motif is exposed. At the base of the pocket a highly conserved aromatic tyrosine residue (Y22) exists surrounded by four basic amino acid residues: R24, R57, R96, and K100. Interestingly, these residues are highly conserved amongst p14 orthologs but are not conserved between p14 and other RRMs (Figure 1-8) [42].

I-3.3: Interaction of p14 with pre-mRNA and Adenosine.

Based on the structure of the p14: SF3b155 complex, it was apparent that p14 may not be able to interact with RNA in a canonical fashion given that the RRM domain was highly occluded. Generally, RRM domains act as a surface for RNA to bind across the β-sheet [58, 61]. Experiments comparing the affinity of p14 for single-stranded, duplex, or bulged-duplex RNA showed there was no preference, and in fact all demonstrated a weak association [40]

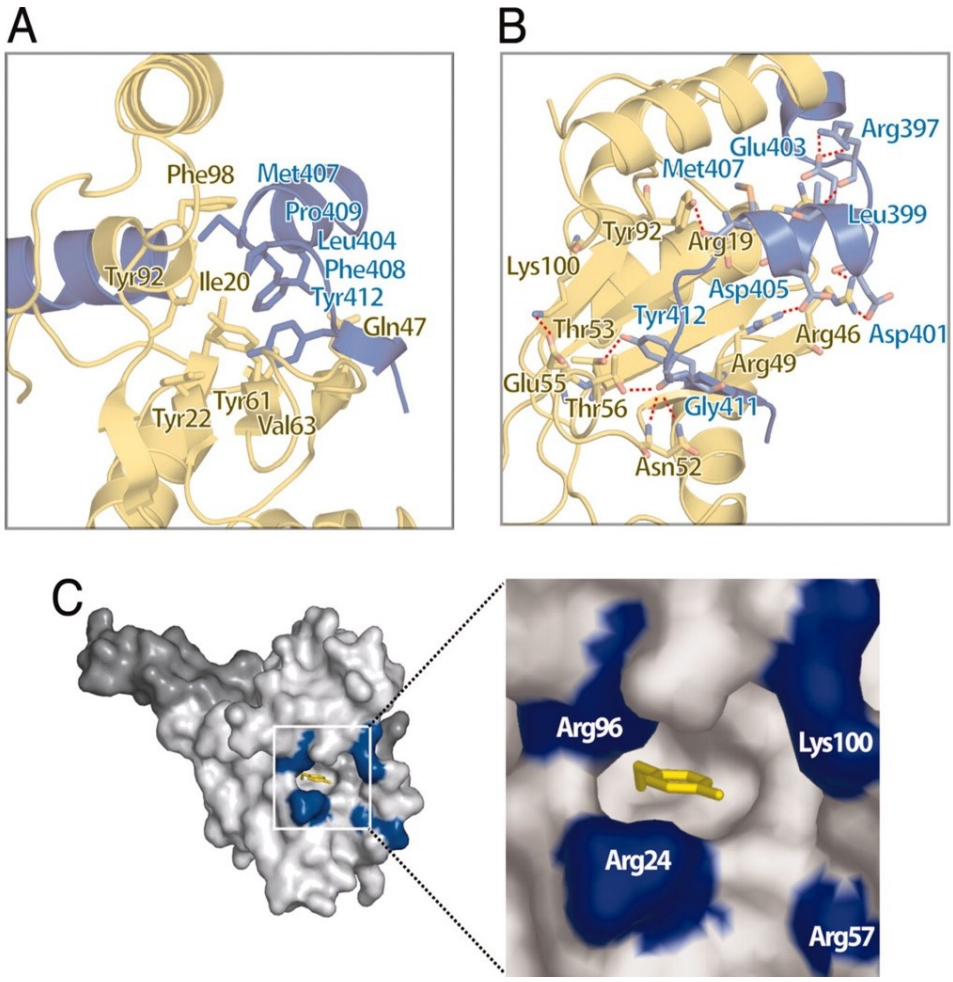


Figure 1-8. Interface between p14:SF3b155 complex. A) Ribbon diagram highlighting the hydrophobic core formed within the p14:SF3b155 complex. The SF3b155 peptide is colored blue, and p14 is colored yellow. B) Hydrogen bonds and electrostatic interactions formed between amino acids surrounding the hydrophobic core. The SF3b155 peptide is colored blue, and p14 is colored yellow. The dotted lines indicate salt bridge interactions between amino acids. C) Surface representation of small pocket formed within the p14:SF3b155 peptide complex. The SF3b155 peptide is colored dark grey, and p14 is colored light grey. The conserved Y22 residue found within the RNP2 motif is colored yellow, and the four basic amino acids surrounding the pocket are colored blue. (Adapted from Schellenberg *et al*., 2006, *PNAS*, 103 (5) 1266-1271).

But, important to note, this data only reflects trials done in the context of the p14:peptide complex and may not translate to what truly happens *in-vivo*. To further study the site of protein: RNA interaction, cross-linking experiments were performed using the p14:SF3b155 complex and pre-mRNA sequences containing a single radiolabeled adenosine at the branch position. Following irradiation at 254 nm, cyanogen bromide and endoproteases Glu-C and Lys-C were used to digest p14 into segments to map the location of the bound adenosine [62]. The radiolabeled adenosine mapped to a location within the β-1-strand of the RRM domain where the RNP2 motif exists (Figure 1-9A). Interestingly, this translates to the same location where the small pocket containing the tyrosine residue was previously observed. To see whether this aromatic residue within p14 was truly interacting with the branch adenosine, a Y22M mutant was created and cross-linking experiments were repeated. Results showed that the M22 residue cross-linked with p14 but this interaction was sensitive to cyanogen bromide treatment (Figure 1-9B). An explanation for this could be that the cross-link was occurring between the branch nucleotide and the terminal methyl group of the side chain of methionine [62]. Therefore, it's suggested that the branch nucleotide interacts directly with the aromatic ring of the tyrosine residue (Figure 1-9C).

As mentioned earlier, the affinity of the isolated p14:SF3b155 complex for single-stranded/duplex RNA is quite poor. Therefore, to obtain a structural and biochemical characterization of p14:RNA interaction, a unique disulfide tethering approach was utilized . This strategy relied upon complex formation between thiol-derivatized RNA and exposed

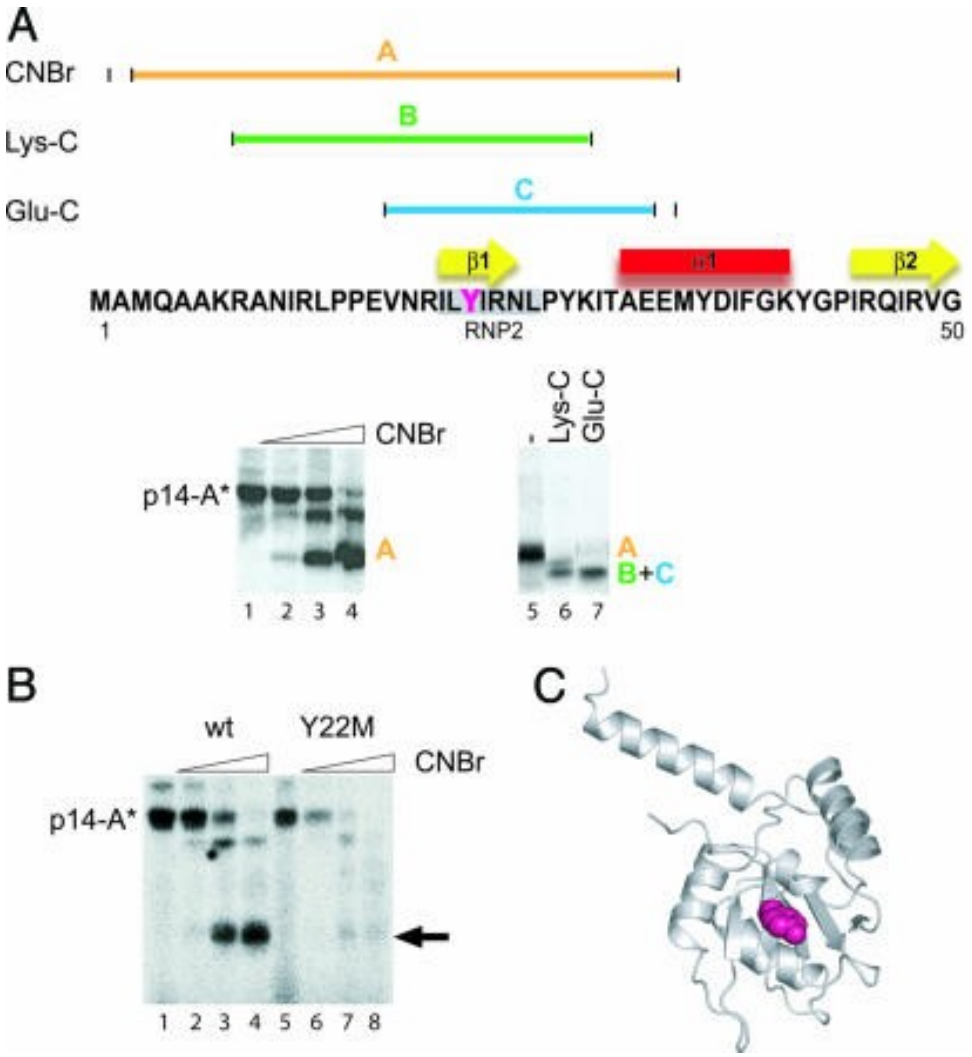


Figure 1-9: Mapping location where the branch nucleotide interacts with p14. A) Treatment of p14:SF3b155 peptide complex:RNA crosslink with cyanogen bromide (CNBr) in increasing concentrations [lane 1-5] leads to formation of fragment A. Follow-up treatment of fragment A with Lys-C [lane 6] leads to the formation of fragment B. Alternatively, treatment of fragment A with Glu-C produces fragment C [lane7] which maps to a location within the RNP2 motif of p14. B) CNBr treatment of WT p14:RNA crosslink [lane 1-4] and Y22M p14:RNA crosslink [lane5-8]. Arrow indicates expected product following protease treatment. C) Ribbon diagram of p14:SF3b155 peptide complex with Y22 residue shown in purple. (Adapted from Schellenberg *et al*., 2006, *PNAS*, 103 (5) 1266-1271).

cysteine (Cys) residues under reducing conditions [63-67]. Based on the aforementioned cross-linking experiments which showed interaction between the bulged adenosine and a pocket on the p14 surface, two models of p14:RNA interactions are possible related by 180° rotation about the

nucleotides glycosidic bond (Figure 1-10A). To elucidate the correct orientation, a Cys-less version of p14 was created (C83S; C74V) and then a panel of single Cys containing p14 mutants was generated [62]. These Cys residues were strategically placed based on the prior X-ray structure to resemble possible areas of interaction between RNA backbone and p14.

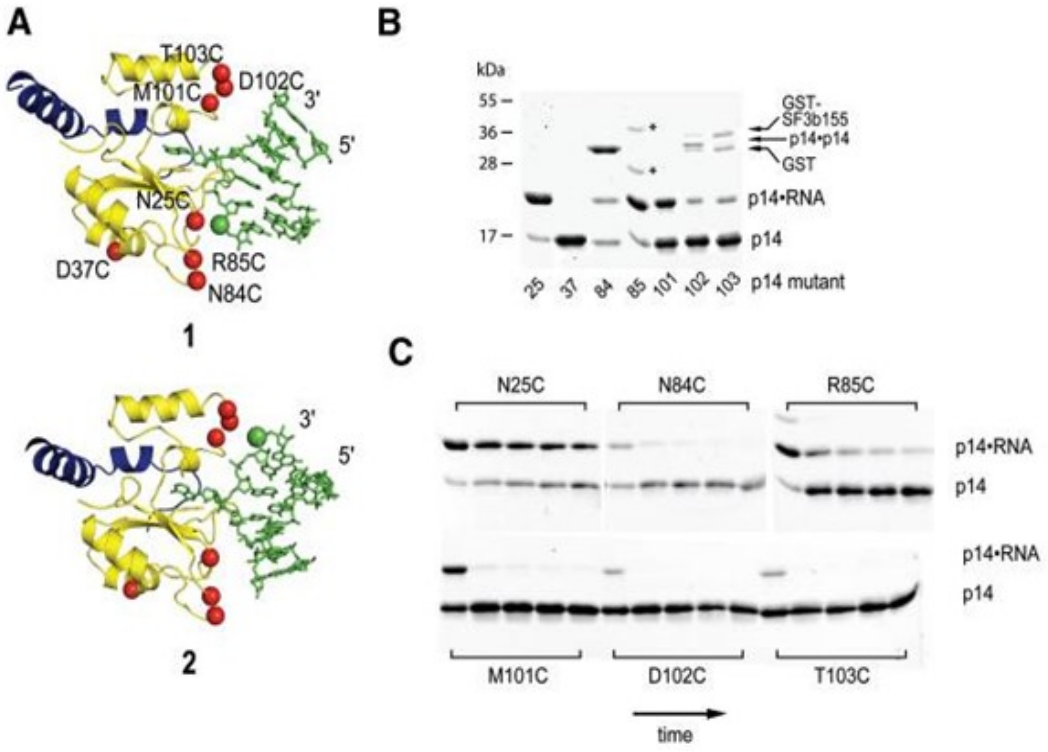


Figure 1-10. Determining the orientation of p14:SF3b155 peptide complex interaction with RNA. A) Ribbon diagrams demonstrating the two possible orientations of RNA related by 180° rotation about the branch adenosine glycosidic bond. SF3b155 peptide is shown in blue, p14 (yellow), RNA (green), single cysteine mutations introduced into Cys-less p14 (red balls), and RNA back bone modification containing C2 thiol functionality (green ball). B) SDS-PAGE gels showing p14:RNA disulfide cross-links under reducing conditions using single Cys-p14 mutants. * indicates uncharacterized contaminants. C) SDS-PAGE analysis showing time course for p14:RNA disulfide formation and stability of single Cys-mutants under reducing conditions. (Adapted from Schellenberg *et al*., 2011*, RNA*. 17(1): 155–165.)

A total of seven single p14-Cys mutants were generated, including a D37C mutant which maps to the "rear" face of the p14:peptide complex acting as a negative control to ensure there weren't

any non-biological interactions occurring. Overnight incubation of Cys mutants with thiol-derivatized RNA under mild reducing conditions (60 μM β-mercaptoethanol) showed the highest disulfide complex formation to be with the N25C and R85C mutants (Figure 1-10B) [62]. The results of these experiments provided further confirmation that the bulged duplex interacts with only one face of the p14/SF3b155 peptide complex (evidence by the instability of disulfide formation with the D37C mutant), and that the most probable orientation places the RNA backbone closest to the N25 residue (higher kinetic stability of N25C compared to R85C mutant) (Figure 1-10C).

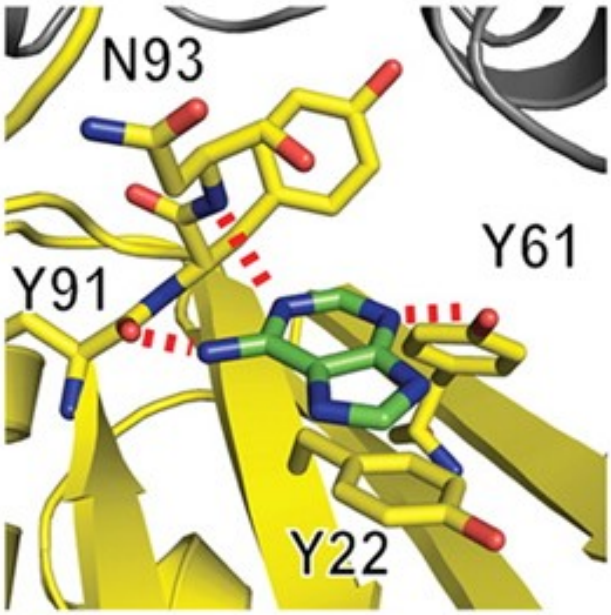


Figure 1-11. Recognition of the branch adenosine by p14. Crystal structure of adenine bound to p14:SF3b155 peptide complex. The adenine base can be seen stacking against the highly conserved Y22 residue of p14. Shown by the red dotted lines are the hydrogen bonds formed between the nucleotide and the surrounding p14 residues (N93, Y22, N91, Y61). The grey structure seen at the top is a portion of the SF3b155 peptide. (Adapted from Schellenberg *et al*., 2011, *RNA*. 17(1): 155–165.)

To further characterize the interaction between the protein complex and the branch adenosine, Cys-less p14:SF3b155 crystals were soaked with adenine. The structure was subsequently solved at a resolution of 2.4 Å. Analysis of the structure confirmed what had previously been shown via cross-linking experiments in which the adenine base directly interacts with the Y22 residue found within the RNP2 motif. The adenine base appears to stack overtop of the aromatic ring of Y22 and form hydrogen bond interactions with neighboring residues (Y91, Y61, N93) (Figure 1-11) [62].

Our lab has recently solved the p14:SF3b155 complex crystal structure from the fission yeast, *S. pombe*, at a resolution of 1.9 Å (Figure 1-12) (Grewal *et al*., unpublished). As predicted, in comparison to the human structure, it was nearly identical. The only difference was attributable to missing density from the N-terminal α-helix of the SF3b155 peptide, and the two C-terminal α-helices of p14, in the *S. pombe* structure.

These results strongly support the use of *S. pombe* to study the function of p14 in pre-mRNA splicing. This thesis aims to answer the aforementioned question by comparing WT and Δ p14 *S. pombe* strains using growth curve assays, stress tests, morphological studies, and mass spectrometry analysis. Given that p14 has been shown to directly interact with the 3′ intron, I hypothesize that it is required for the splicing of a certain subset of introns based on their branch sequence composition.

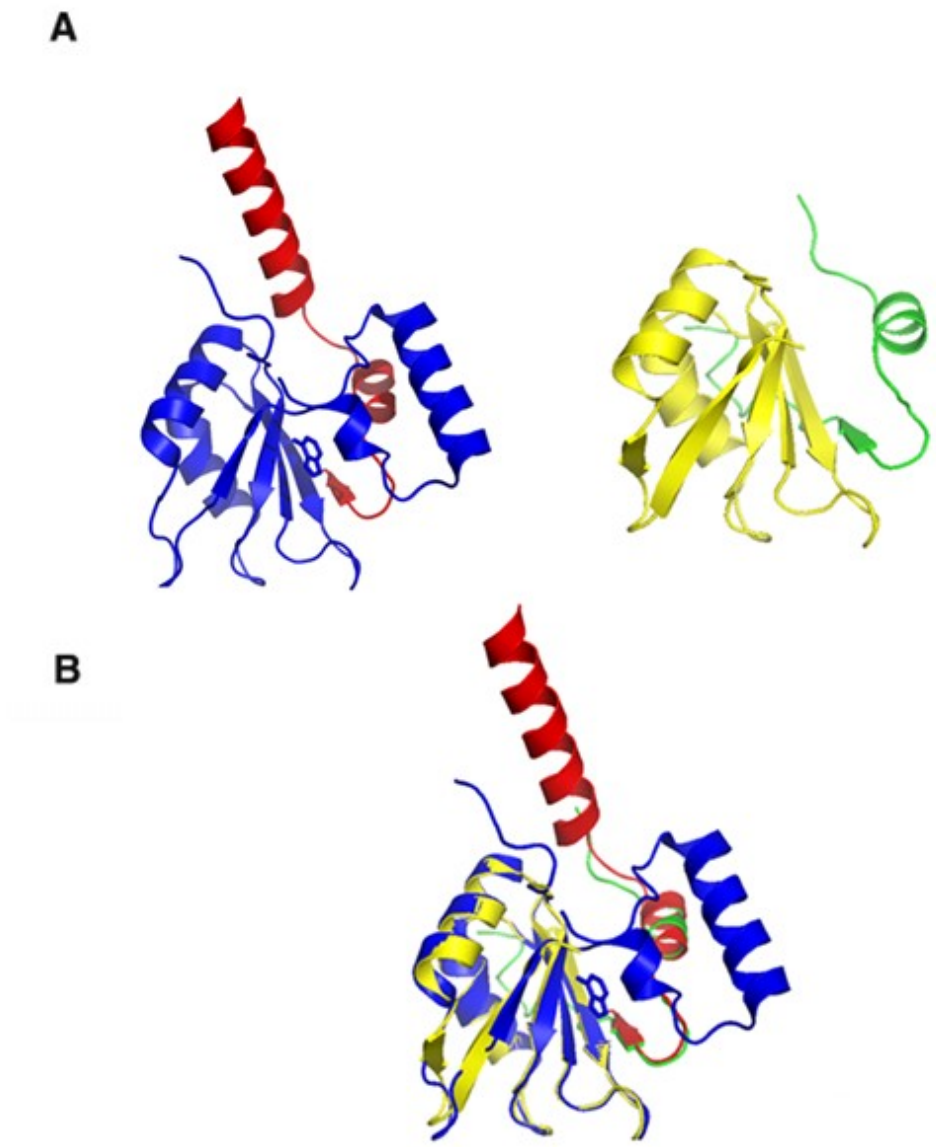


Figure 1-12. Structural comparison of human and *S.pombe* p14:SF3b155 complex. A) [left] Crystal structure of human p14:SF3b155 peptide complex solved at a resolution of 2.5 Å. p14 (blue); SF3b155 peptide (red). [right] Crystal structure of *S.pombe* p14:SF3b155 solved at a resolution of 1.9 Å. p14 (yellow); SF3B155 peptide (green). B) Human and *S.pombe* crystal structures superimposed. Overall, the structures show excellent alignment of the p14 RRM domain a portion of the SF3b155 peptide. The yeast structure lacks density for both the long N-terminal α-helix of the SF3b155 peptide and the two C-terminal α-helices of p14. (Grewal *et al*., unpublished)

Chapter 2
Materials and Methods

2-1.1. Growth Curve Assays:

WT (ura4-D18 leu-32 ade6-216 his3-D1) and Δ p14 (ade6-M210 ura4-D18 leu1-32, generated by targeted mutagenesis from Bioneer) *S. pombe* single colony isolates were collected from YES plates (5 g/L Yeast extract, 20 g/L Dextrose, 1X supplemental stock , 20 g/L agar) and used to inoculate 7 mL of YES liquid media in culture tubes. The tubes were put into a 30°C incubator spinning at 200 rpm (revolutions per minute). The density of the media was periodically checked using a spectrophotometer calibrated at OD_{600}. Once the optical density of media reached between 0.1-0.6 (mid-log phase) or >1.0 (saturation), the cultures were removed from the incubator. To determine the correct volume of pre-culture that should be used to inoculate fresh culture to give a desired optical density within a set time, the formula $vol.of\ cells\ inoculted = (desired\ optical\ denisty/optical\ density\ of\ pre\ culture) \times vol.of\ fresh\ culture/2^n$), where n = the number of generations that the culture is to be grown (see table below) [68]. The calculated volumes of pre-culture were then used to inoculate fresh 25 mL YES media. The media was diluted, and the optical density was checked every hour for 8 hours using spectrophotometry to generate growth curves.

Table 2-1. Generation times at different temperatures for WT *S. pombe*.

Temperature	Generation Time*
20°C	6 hours
25°C	3 hours
30°C	2.5 hours
35°C	2 hours

*denotes the time required for the yeast to double in number.

2-1.2. Stress Test Assays:

YES agar plates containing stress reagents at varying concentrations were made. The stress conditions used were as follows: 5-15 mM hydroxy urea, 1-2 M sorbitol, 0.01-0.05% methyl methanesulfonate, and 1-3 M hydrogen peroxide. JK484 WT and Δ p14 *S. pombe* cells were grown in YES media to an optical density of 0.2. The cells were diluted ten-fold from 10^{-1} to 10^{-4} and 5 µL volumes were plated directly onto YES plates containing the stress reagents. Control plates were also made which contained none of the stress reagents. The plates were stored in a 30°C incubator for 4-5 days.

2-2.1. Preparation of Electrocompetent *S. pombe* Cells:

A culture of 500 mL of *S. pombe* cells were grown at 30°C overnight to an optical density of 0.7 in SD medium (0.67% bacto yeast nitrogen base without amino acids, 2% glucose) supplemented with 1X adenine, histidine, leucine, and uracil. The cultures were decanted into sterile 250 mL centrifuge bottles and chilled in an ice water bath for 15 minutes to supress cell growth. Afterwards, cultures were centrifuged at 1600 xg for 5 minutes at 4°C, and pellets were washed 3x with sterile ice-cold water. The final pellet was resuspended in 1 mL of ice cold 1M sorbitol and 50 uL aliquots were added to chilled microfuge tubes. The tubes containing the electro-competent cells were flash frozen using liquid nitrogen and placed at -80°C for later use up to 6 months. It is very important the cells remain cold at all times or they may lose their electrocompetency [69].

2-2.2. Transformation of *S. pombe* cells via Electroporation:

Microfuge tubes containing 50 µL of electrocompetent cells were collect from the -80°C freezer and thawed quickly using a 30°C water bath. The cells were washed with 750 mL of ice cold 1 M sorbitol via centrifugation. The yeast pellet was resuspended in 50 µL of ice cold 1M sorbitol and 1 µg of pFA6a-3HA-kanMX6 PCR product was added. The sample was then transferred to a pre-chilled 0.2 cm electroporation cuvette and left on ice for 30 minutes. Afterwards, the cuvette was pulsed at 1.5 kV using an Eppendorf Electroporator and 1 mL of ice cold 1.2 M sorbitol was rapidly added. The transformed cells had to be grown in YES media overnight to give them time to express the kanMX6 product making them resistant to the G418 antibiotic. The following day the cultures were centrifuged, and the pellet was resuspended in 150 µL of YES media and plated directly onto YES agar plates containing G418.

2-3.1. Cloning pFA6a-3HA-kanMX6 Integration Vector with SF3b49 Homology Ends:

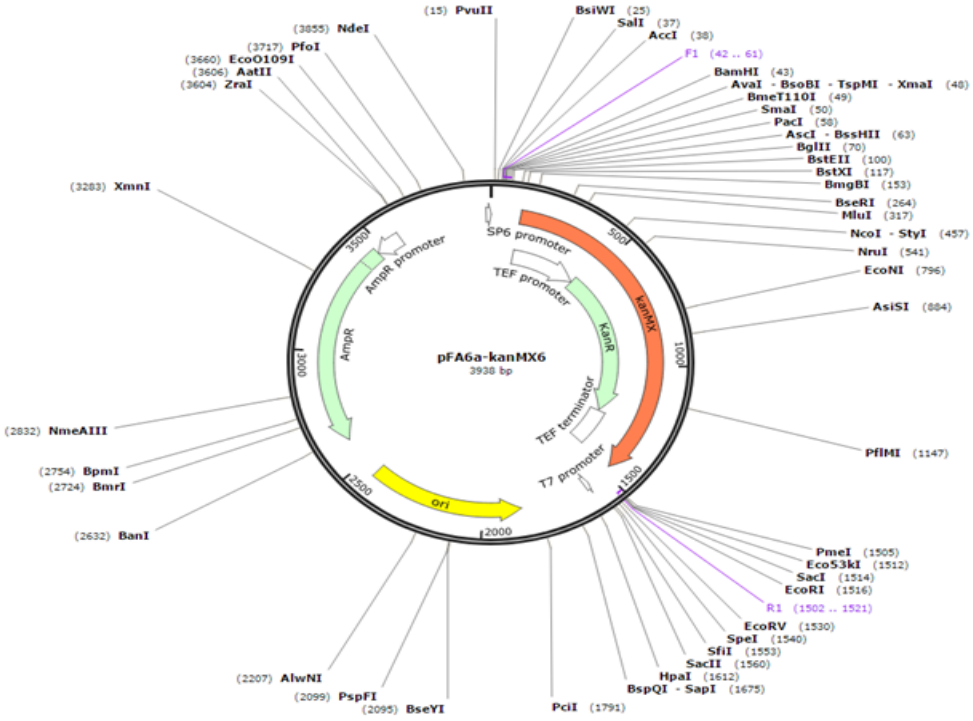
The pFA6a-3HA-kanMX6 is 2.3 kb yeast genomic integration vector containing a selectable marker for G418 antibiotic (Figure 2-1A) [70, 71]. The genomic sequence for SF3b49 was collected from the PomBase website and primers with homology to the sequence were synthesized. The forward primer was 80 nucleotides in length and contained 60 base homology upstream of the stop codon and 20 base homology to the 5′ end of the pFA6a-3HA-kanMX6 cassette 5′ -AACATACCATTTTATCAAACAATCAATGCACAAAATGGATACAGTCAACAACAAAG GAGGCGGATCCCCGGGTTAATTAA-3′ . The reverse primer was also 80 nucleotides in length containing 60 base homology downstream of the stop codon and 20 base homology to the 3′ end

of the pFA6a-3HA-kanMX6 cassette 5′-

TCAATTTTAATTTCTTTTAAATAAGATAAATCTCAGCTTCCACTTTTTTTACTGTACAC

GGAATTCGAGCTCGTTTAAAC-3′.

A



B

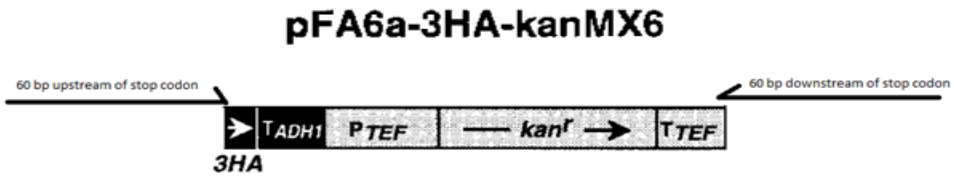


Figure 2-1. Organization and utility of the pFA6a-3HA-KanMX6 vector. A) This vector utilizes homologous recombination to insert a 3 x Hemagglutinin tag into *S. pombe* genomic DNA. Selection markers found within this plasmid include ampicillin resistance (AmpR) and kanamycin resistance (kan^r). B) Cassette region contained within the pFA6a-3HA-KanMX6 vector. The PCR strategy involves using a 60 nt forward primer upstream of the stop codon of your protein of interest and 60nt reverse primer. downstream of the stop codon. Through homologous recombination the PCR product will insert the 3xHA peptide sequence at the C-terminal end of the protein of interest.

Standard PCR protocol using Pfx DNA Polymerase was used to generate linear pFA6a-3HA-kanMX6 vector containing 5′ and 3′ homology upstream and downstream of the SF3b49 gene (Figure 3-6A).

2-3.2. Confirming Integration of the pFA6A-3HA-kanMX6 cassette into the SF3b49 gene.

Following transformation of the linear vector containing homology to SF3b49 (see above), transformants were plated on YES + G418 plates to generate single colony isolates (SCI). SCI were subsequently picked and grown in YES + G418 liquid media overnight to yield enough cells for genomic DNA extraction.

To isolate the yeast genomic DNA, the 10 mL overnight cultures were centrifuged, and the supernatant was discarded. The resulting pellet was washed with 0.5 mL of double distilled water (ddw) and 0.2 mL of extraction buffer (2% Triton X-100, 1% SDS, 100 mM NaCl, 10 mM Tris-HCl (pH 8.0), 1 mM EDTA), 0.2 mL of phenol:chloroform:isoamyl alcohol (25:24:1), and 0.3 g of acid-washed glass beads were added and the microfuge tubes followed by 3-4 minute vortex cycle. Afterwards, 0.2 mL of TE buffer (10 mM Tris pH 8.0, 1 mM EDTA) added the microfuge tubes were centrifuged for 5 minutes at 16000 revolutions per minute (RPM). The aqueous layer was transferred to a fresh tube and 1 mL of 99% ethanol was added to the tubes and mixed by inversion. The tubes were centrifuged for 2 minutes and the pellet was resuspended in 0.4 mL of TE buffer plus 3 μL of a 10 mg/mL solution of RNase A. The mixture was incubated at 37°C for 5 minutes followed by the addition of 10 μL of 4 M ammonium acetate and 1 mL of 99% ethanol. The sample was centrifuged for 2 minutes and the supernatant was discarded. The pellet was air-

dried for 2-3 hours using an air vacuum and the dry pellet was resuspended in 50 μL of TE buffer and quantified.

The extracted genomic DNA was used as PCR template to confirm integration of the pFA6a-3HA-kanMX6 cassette into the SF3b49 gene. The forward and reverse primers used were 100 bases outside of the area of expected integration. FWD: 5′GTCAATACCTAATGTTCTTCCTTTCACTG-3′ REV: 5′ -ATAGACTATATTGCCGATATTTAGGGACG-3′.

2-3.3. Preparation of Yeast Protein Extract.

Cultures consisting of 2 L of S. *pombe* cells were grown to an optical density between 1.5-2.0. The cells were harvested via centrifugation at 3200 rpm for 5 minutes at 4°C, and subsequently washed with 50 mL of ice cold ddw before being transferred to falcon tubes. While in falcon tubes the cells were washed with 50 mL of ice cold AGK buffer (10 mM HEPES-KOH pH 7.9, 400 mM KCl, 1.5 mM $MgCl_2$, 0.5 mM DTT, 1 mM PMSF, 10% glycerol) and resuspended in 7.5 mL of ice cold AGK buffer. The cell suspension was transferred to an 18-gauge needle and slowly dripped into a bucket containing liquid nitrogen. The frozen cell droplets were transferred to a falcon tube and frozen at -80°C overnight. The following day, the cells were poured into a pre-chilled mortar and ground using a pestle for 30 minutes. The ground cells were transferred to a cool beaker and mixed using a magnetic stir bar for 1 hour at 4°C before being transferred to a pre-chilled Oakridge tube and centrifuged at 15,000 rpm for 30 minutes. The supernatant was collected and transferred to a Quick-Seal polyallomer tube and centrifuged at 49,000 rpm for one hour using a 70.1 Ti Rotor. Afterwards, the resulting supernatant was dialyzed using a 7K Slide-A-Lyzer™

cassette in 2L of buffer D (20 mM HEPES-KOH pH 7.9, 0.2 mM EDTA, 100 mM KCl, 0.5 mM DTT, 1 mM PMSF, 20% glycerol) for 1.5 hours. The concentration of the dialyzed protein extract was quantified using the Bradford Assay, aliquots of the protein were made and snap-frozen at -80°C for later use.

2-3.4. Western Analysis to Confirm SF3b49-3HA Tag.

Protein extracts were purified from WT *S. pombe* cells showing PCR-proven pFA6a-3HA-kanMX6 cassette integration into the C-terminus of the SF3b49 gene. Varying amounts of protein exact (5-20 ug) were dissolved in 5x loading dye (250 mM Tris·HCl, pH 6.8, 10% SDS, 30% (v/v) Glycerol, 500mM DTT, 0.05% (w/v) Bromophenol Blue), and run on a standard 16% SDS PAGE resolving gel at 140 V for 1.5 hours. The separated proteins were subsequently transferred to a nitrocellulose membrane using a blotting tank running at 300 mA for 1.5 hours. Afterwards, blocking solution (1X TBST, 10% w/v dry milk powder) was used to cover the nitrocellulose membrane at room temperature for 2 hours while mixing on a rocker. Primary anti-HA mouse antibody was dissolved in 1X TBST + 2% w/v dry milk powder in a 1:10,000 dilution and incubated with the nitrocellulose membrane overnight at 4°C. The next morning the membrane was washed 3x using 1X TBST in ten-minute intervals followed by the incubation of goat anti-mouse (HRP conjugate) secondary antibody (1:2000) at room temperature for 2 hours. A fresh mixture of Immobilon Western Chemiluminescent HRP Substrate was prepared and incubated with the nitrocellulose membrane for 5 minutes. To visualize the signals, x-ray films were exposed to the nitrocellulose membrane for 30-60 seconds in the dark room and subsequently developed using a film imager.

2-3.5. Immunoprecipitation of SF3b49-3HA.

Protein G Sepharose® beads were washed 3x via centrifugation with lysis buffer (25 mM Tris-HCl pH 7.4, 150 mM NaCl, 1 mM EDTA, 1% TritonX-100 and 5% glycerol). A 50% slurry consisting of washed beads and lysis buffer was prepared and stored at 4°C. Anti-HA mouse antibodies were added to the slurry and incubated mixing at 4°C for one hour. Following incubation, 10-50 µL of the slurry was added to yeast protein extract mixing at 4°C for one hour. Afterwards, the mixture was centrifuged at 12,000 rpm for 20 seconds and the supernatant was decanted. The resulting pellet was washed 3x with lysis buffer and one time with wash buffer (50 mM Tris-HCl pH 8.0) and then mixed with 30 µL of sample buffer (50 mM Tris-HCl pH 7.5, 1% SDS, 100 mM DTT) and heated to 95°C for 3 minutes. The supernatant was mixed with 0.1% bromophenol blue and run on 16% 29:1 acrylamide SDS-PAGE and Western analysis was performed (see above).

Alternatively, 40 µL of yeast protein extract (0.385 mg/mL) were mixed with 20 µL of Protein G beads bound to anti-HA antibodies and topped up to 500 µL with extraction buffer (50 mM Tris-HCl pH 8.0, 150 mM NaCl, 1% IGEPAL, 2 mM EDTA, 1 mM PMSF, 2 mM Benzamidine, 50 mM NaF, 0.1 mM Na_3VO_4, 50 mM β-glycophosphate, cocktail protease inhibitor) and incubated at 4°C for 2 hours. The mixture was centrifuged at 12,000 rpm for 20 seconds and the supernatant was collected. The beads were subsequently washed 4x with phosphate buffered saline solution (137 mM NaCl, 10 mM Phosphate, 2.7 mM KCl, and a pH of 7.4) at increasing concentrations of NaCl (137, 400, 600, 800 mM) and each wash fraction was collected for later analysis. Afterwards,

the beads were mixed with sample buffer and incubated in a 95°C heat block for 3 minutes. The wash fractions and the final elution were used for Western analysis and RT PCR.

2-3.6. RNA purification and RT PCR.

Following the immunoprecipitation experiments (see above), 360 µL of each wash solution was added to 40 µL of 3 M NaOAc and mixed by inversion. 120 µL of phenol was added and the mixture was separated by centrifugation. The supernatant was collected and 120 µL of chloroform was added and once again separated by centrifugation. The resulting supernatant was added to a solution containing 950 µL of 95% ethanol and 1 µL of glycogen and frozen at -20°C for one hour. Afterwards, the solution was centrifuged at 16,000 rpm for 20 minutes and the pellet was resuspended in 120 µL of fresh 95% ethanol. The tubes were re-spun and the ethanol was decanted. The final pellet containing the RNA was resuspended in 30 µL of Diethyl Pyrocarbonate (DEPC) -treated water.

Reverse transcription polymerase chain reaction (RT PCR) was performed by using the SuperScript™ III One-Step RT-PCR System with Platinum™ Taq DNA Polymerase (Thermo Fisher). Purified RNA (0.1 µg) was used as template along with primers for the U2 snRNA. The forward primer sequence used was: 5′ -TATTCTCTCTTTGCCTTTTGGCTTAGATCAAG-3′ , and the reverse primer sequence used was: 5′ -TTCGGCGTCGCTTGCCAGTAGTGCAATAGC-3′ . The final products were run on a 1 % Agarose gel and visualized under ultraviolet (UV) light.

2-3.7. Proteomics Mass Spectrometry.

Protein extracts were isolated and purified from WT and Δ p14 *S. pombe* cells (see above) and concentrations were quantified using the Bradford Assay (reagents purchased from Bio-Rad). Samples were run on a 10% sterile filtered SDS-PAGE gel to separate out the proteins and the gel was stained using Staining solution (0.1% Coomassie Brilliant Blue R-250, 50% methanol and 10% acetic acid) for 20 minutes with gentle agitation and subsequently destained using Destaining solution (40% methanol and 10% acetic acid). Protein bands were cut directly from the SDS-PAGE gel using a clean razer blade. Each gel slice was further cut into smaller pieces (~ 1 mm^2) and placed into a clean tube. 100µL of 50 mM Tris pH 8.0 in 50% acetonitrile were added to each tube and the samples were vortexed for 10 minutes. The supernatant was subsequently extracted using a clean pipette tip and discarded. The remaining gel pieces were dried in a Speed Vac for ~5 minutes and then mixed with 100µL of solution containing 10 mM DTT in 50mM Tris pH 8.0 and vortexed before incubating for 15 minutes at 65°C. Afterwards, 15mM (final) iodoacetamide was added to the solution and incubated for 30 minutes at room temperature in dark. The iodoacetamide was quenched by adding 10mM DTT (final) to the sample. The gel pieces were subsequently washed with 100 µL of 50mM Tris pH 8.0 and vortexed for 10 minutes before the supernatant was discarded. The remaining gel pieces were dehydrated using 100µL of 50mM Tris pH 8.0 in 50% acetonitrile, vortexed for 5 minutes and supernatant was discarded. The gel pieces were once again dried using the Speed Vac for 5 minutes. A fresh solution of 12.5ng/µL Trypsin/LysC in 50mM Tris pH 8.0 was prepared and 25µL were added to completely cover the gel pieces before being incubated overnight with agitation at 37°C. The next morning the tubes were centrifuged at 13,000 RPM for one minute and the supernatant was extracted and moved to a new tube. The gel pieces

were washed with 30µL of 5% formic acid in 50% acetonitrile, vortexed for 25 minutes and sonicated for 5 minutes. The gel pieces were spun at 13,000 RPM and the supernatant was collected and added to the supernatant from before. The peptides were purified by reversed phase solid phase extraction and subjected to LC-MS.

To assess changes in protein expression between the WT and Δ p14 *S. pombe* samples, a two-tailed heteroscedastic Student's t-test was completed to proteins in both datasets. The sorted p-values were subjected to the "*q*-value estimation for FDR" web utility (qvalue.princeton.edu) to generate FDRs (*q*-values). Thereafter, the fold changes between protein abundances were determined by applying a log2 function using the standardized average EICs for each protein (*see Appendix A*). The statistical analysis completed with the help of Dr. David Kramer from Dr. Fahlman's lab at the University of Alberta, Department of Biochemistry.

Chapter 3
Results

3-1.1. Comparing Growth Between WT and Δ p14 JK484 *S. pombe*.

Initial experiments were performed to determine whether the U2 snRNP associated protein, p14, had an effect on *S. pombe* growth. Based on the hypothesis that p14 is required for the splicing of a certain subset of introns, knockout of this protein may show differences in cell growth due to poor protein expression. To test this, both WT and Δ p14 yeast cells (generated by Bioneer), either in log phase or saturation, were grown in fresh liquid media for a set time under the same conditions (see Methods). The turbidity of the media was measured by a spectrophotometer every hour and used as a surrogate for cell growth. Our results showed that when *S. pombe* cells were in the log phase, the growth curve between WT and Δ p14 was similar (Figure 3-1).

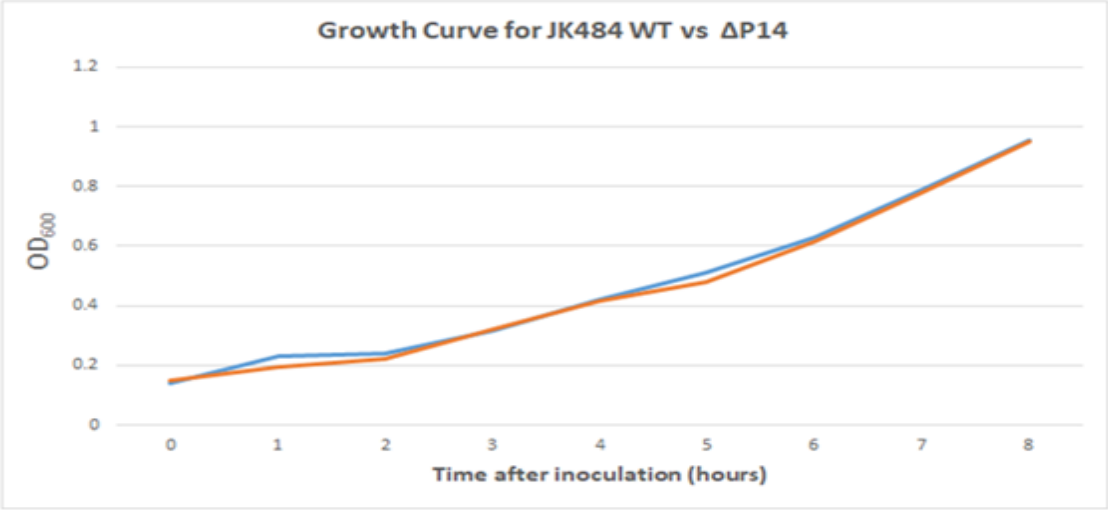


Figure 3-1. Growth curve of WT vs Δ p14 *S. pombe* grown from early log phase. A) Pre-cultures of both strains were grown to an optical density of 0.7 (log phase). 5mL of pre-culture was used to inoculate fresh YES media and incubated in a 30°C centrifuge. Optical densities were checked periodically every hour with a spectrophotometer calibrated at OD_{600}. WT (red); Δp14 (blue). Volume of pre-culture used for inoculation was calculated using the formula: vol. of cells inoculated = (OD required/ OD pre-culture) x (Vol. of culture/ 2n) where n = the number of generations the culture is to be grown.

Table 3-1. Growth curve of WT vs Δ p14 *S. pombe* grown from log phase. Pre-cultures of both strains were grown to an optical density of 0.7 (log phase). 5 mL of pre-culture was used to inoculate fresh YES media and incubated in a 30°C centrifuge. Optical densities were checked at every hour post inoculation with a spectrophotometer calibrated at OD_{600}.

	Time after inoculation (hours)								
Strains:	0	1	2	3	4	5	6	7	8
WT JK484	0.139	0.232	0.240	0.318	0.423	0.510	0.630	0.790	0.956
Δ p14 JK484	0.150	0.197	0.220	0.321	0.415	0.480	0.614	0.778	0.953

Table 3-2. Growth curve of WT vs Δ p14 *S. pombe* grown from saturation. Pre-cultures of both strains were grown to an optical density of 1.2 (saturation phase). 186 uL of pre-culture was used to inoculate fresh YES media and incubated in a 30°C centrifuge. Optical densities were checked at 15- and 17-hours post inoculation with a spectrophotometer calibrated at OD_{600} (n=3).

	Time after inoculation (hours)	
Strains:	15	17
WT JK484	0.639	0.881
Δ p14 JK484	0.188	0.350

When yeast cells were grown from saturation, both strains demonstrated a slower growth rate compared to the log phase cells, however, the Δ p14 strain grew significantly slower than WT (Table 3-1). After 17 hours of incubation, the WT culture had an optical density of 0.881 compared to 0.350 for the knockout. Next, we were interested to see whether *S. pombe* cells lacking p14 were more sensitive to chemical/ environmental stressors. A literature search was done to determine common stress conditions used in yeast cells [76,78,80]. We came up with the following: hydroxyurea (ribonucleotide reductase inhibitor), sorbitol (osmotic stress), hydrogen peroxide (oxidative stress), methyl methanesulfonate (DNA alkylating agent), heat stress (37-42°C), cold stress (18°C). We added the aforementioned reagents to YES agar plates at select concentrations and plated WT and Δ p14 *S. pombe* cells grown to log phase and saturation.

Amongst the different conditions used, two yielded interesting findings. The first was hydroxyurea, which significantly reduced the knockout's ability to grow.

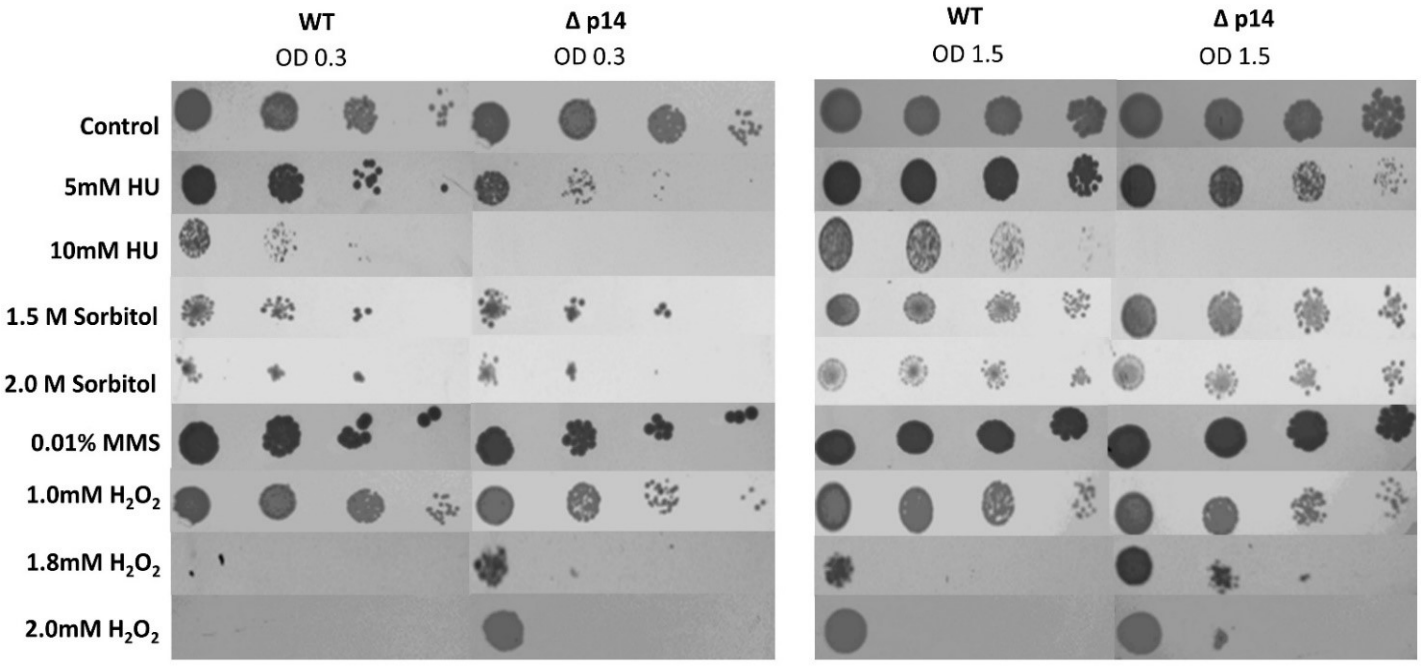


Figure 3-2. Comparing growth of WT vs Δp14 knockout *S.pombe* cells under stress conditions. [left panel] WT and Δp14 knockout cells were grown to an optical density of 0.3 (measured by spectrophotometry) and plated at different dilutions on YES medium (0.5% Yeast Extract, 2.0% Glucose) containing stress reagents. Cells were diluted to (10^{-1},10^{-2},10^{-3},10^{-4}) and 10 µL of each dilution were plated on media. Plates were incubated at 40°C for 5-7 days. [right panel] WT and Δp14 knockout cells were grown to an optical density of 1.5 (measured by spectrophotometry) and a similar protocol to the left panel was performed.

Specifically, Δ p14 cells grown on 5 mM hydroxyurea plates showed a noticeable decrease in colony growth starting from the 10^{-2} dilution, and were completely unable to grow on 10 mM hydroxyurea (Figure 3-2). The WT proved more resistant to hydroxyurea and was able to grow at the 10^{-4} dilution.

The second condition, hydrogen peroxide, showed an interesting trend reversal. The knockout strain was able to grow on 2.0 mM H_2O_2 when grown to log phase, but the WT strain was not. The other stress conditions tested did not yield any significant growth differences between the two strains.

3-2.1. Morphological Differences between WT and Δ p14 *S. pombe*.

We were interested to see whether morphological differences between the two strains could give insight into the physiologic function of p14. For example, if p14 were indeed required for splicing of certain genes then the phenotype of knockout cells could give clues as to which proteins may be affected. To do this, bright field images were taken using the BD LSRFortessa™ flow cytometry instrument of yeast cells grown to either log phase (OD 0.3) or saturation (OD 1.2). When grown to log phase, the WT and Δ p14 *S. pombe* cells demonstrated the expected mitotic division pattern, where the cells grow exclusively from the tips and divide by medial fission giving rise to two identical daughter cells. However, the Δ p14 *S. pombe* cells appeared to stick together and congregate in large clusters of up to ten cells at a time (Figure 3-3A). Interestingly, these clusters were largely resistant to disruption by mechanical means (ie. vortex). To gain a global representation of the population of cells, we used flow cytometry to generate a scatter plot measuring cell area against aspect ratio. Upon analysis of each individual point on the scatter plot (indicating a single cell or a cluster of cells), we crudely correlated area to represent an increasing size of cell groupings (Figure 3-3B). For example, points that plotted a higher area meant they were likely a group of cells stuck together rather than individual cells. In doing so, we concluded that the Δ p14 *S. pombe* strain showed a much greater number of grouped cells than WT. Similarly, the same pattern existed when cells were grown to saturation (Figure 3-4).

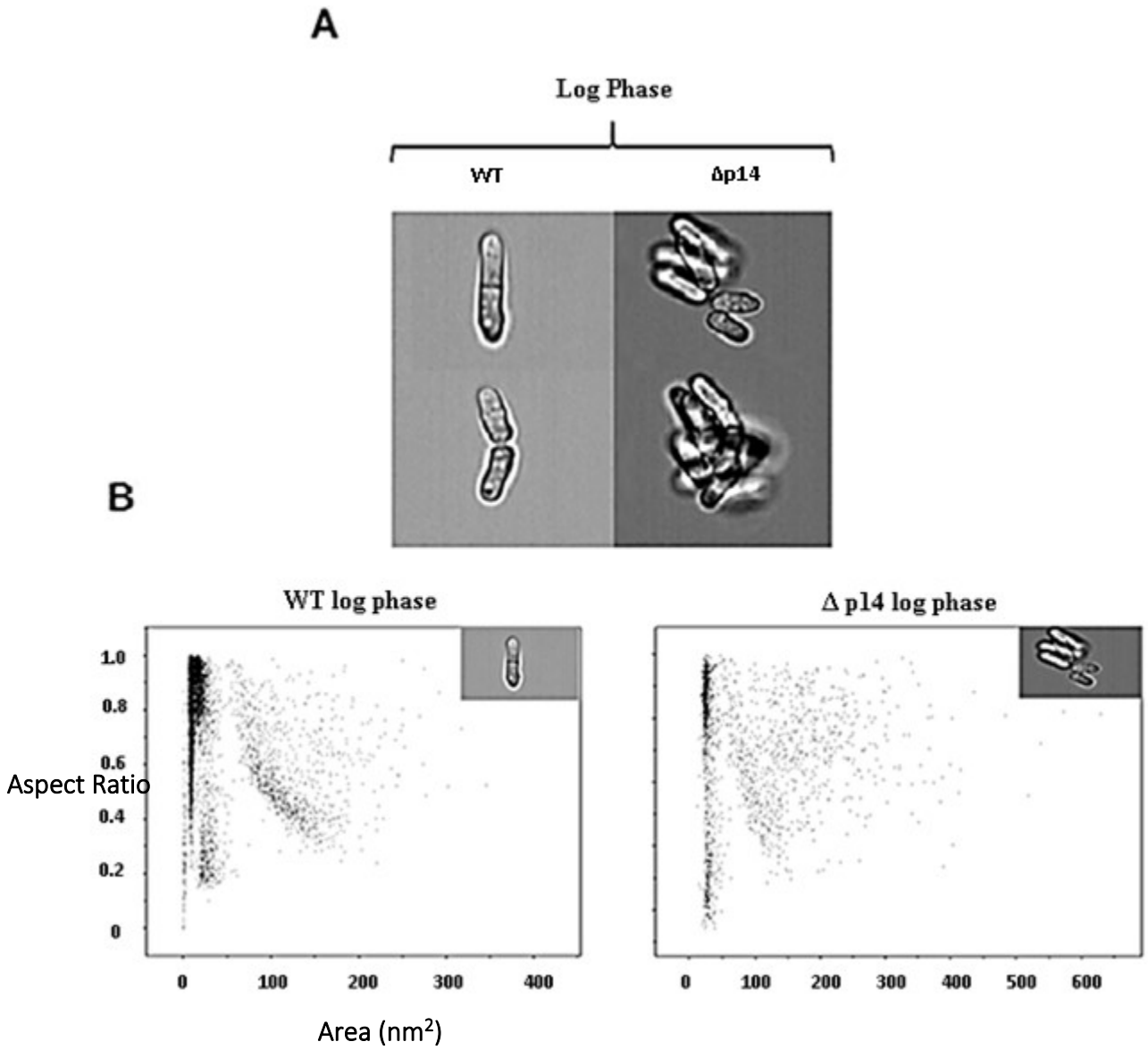


Figure 3-3. Morphological comparison of WT and Δp14 cells grown to log phase. A) Bright field images of *S. pombe* strains grown to an optical density of 0.3 (measured by spectrophotometry). B) Scatter plots generated by the BD LSRFortessa™ flow cytometry instrument measuring cell area against aspect ratio. Each point on the graph indicates one individual measurement which may include one or more cells.

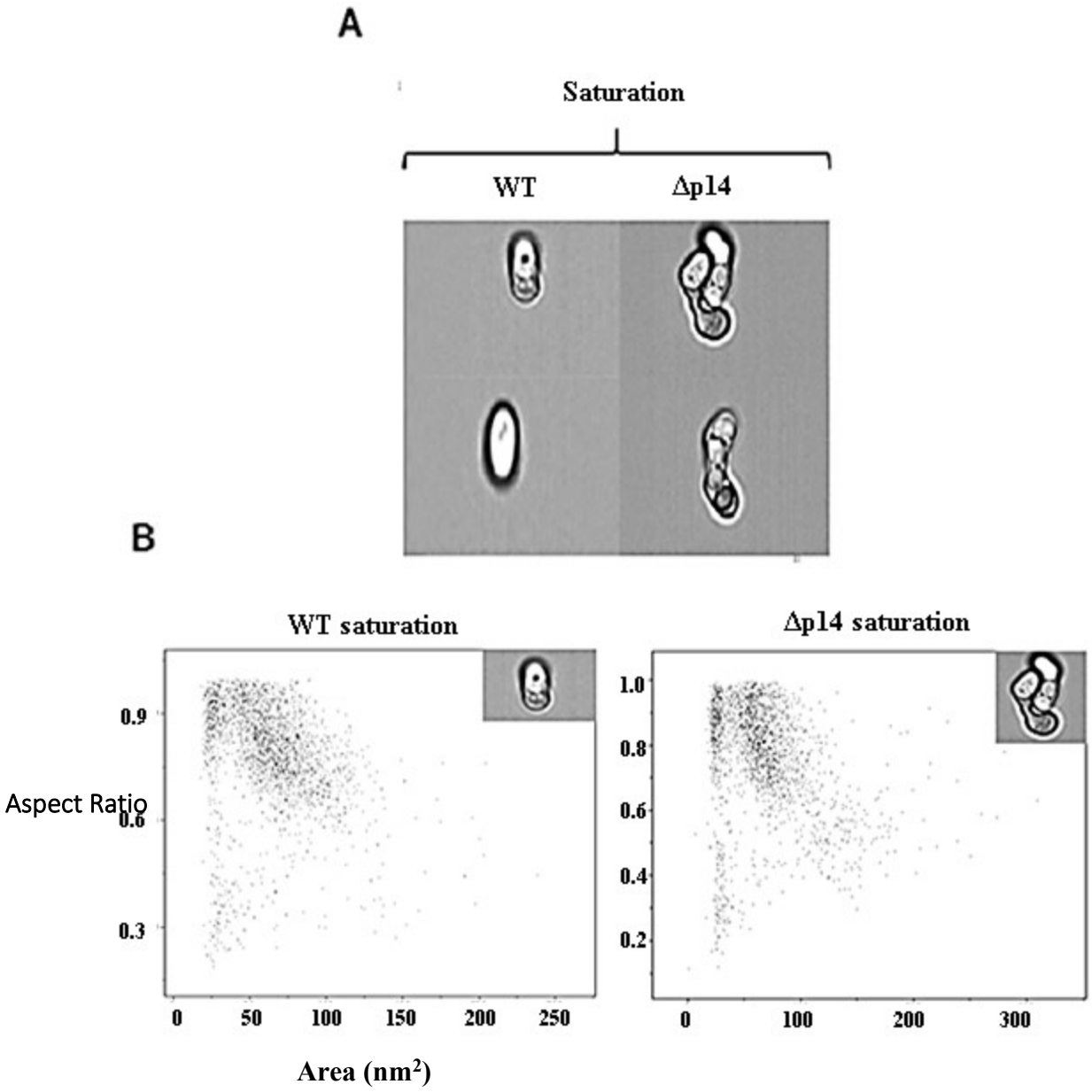


Figure 3-4. Morphological comparison of WT and Δp14 cells grown to saturation phase. A) Bright field images of *S. pombe* strains grown to an optical density of 1.2 (measured by spectrophotometry). B) Scatter plots generated by the BD LSRFortessa™ flow cytometry instrument measuring cell area against aspect ratio. Each point on the graph indicates one individual measurement which may include one or more cells.

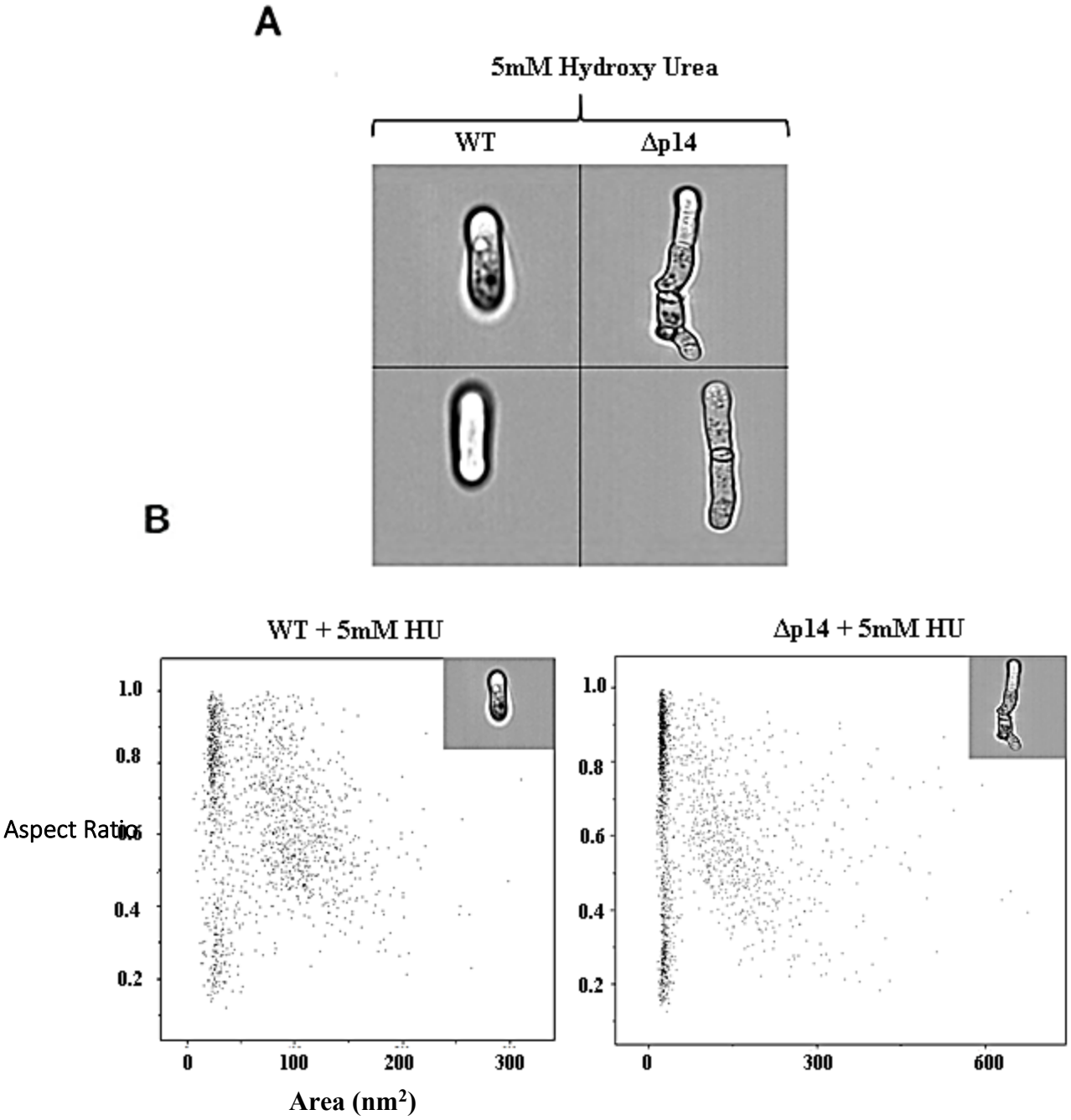


Figure 3-5. Morphological comparison of WT and Δ p14 cells grown to saturation phase. A) Bright field images of *S. pombe* strains grown on YES plates containing 5 mM hydroxyurea. B) Scatter plots generated by the BD LSRFortessa™ flow cytometry instrument measuring cell area against aspect ratio. Each point on the graph indicates one individual measurement which may include one or more cells.

Lastly, we were interested to see whether morphological differences existed amongst the two strains grown in hydroxyurea. As mentioned earlier, hydroxyurea was one of the stress conditions that the knockout strain had difficulties growing in. Intriguingly, we observed such differences in that yeast cells lacking p14 were abnormally elongated and clustered in a chain-like fashion when picked directly from plates and visualized (Figure 3-5). This was reflected in a lower aspect ratio seen on the scatter plot compared to WT.

3-3.1. Isolating the SF3b Complex for Cryo-Electron Microscopy Analysis.

From a structural perspective, we were interested to see whether differences existed within the SF3b complex, with or without p14. To do this, we utilized a genomic DNA tagging strategy to incorporate a nucleotide sequence coding for a protein tag at the terminal end of one of the components of the SF3b complex. This would allow us to later immunoprecipitate and isolate the complex from yeast protein extract and study it using single particle cryo-electron microscopy (cryo-EM). The pFA6a-3HA-kanMX6 yeast integration vector utilizes homologous recombination to insert a 3x HA tag at the terminal end of a protein of interest (see Methods). We created several constructs in an attempt to tag SF3b155, SF3b145, and SF3b130, however, none of these were successful. This was likely a result of the tag interrupting SF3b complex folding, rendering the yeast cells non-viable.

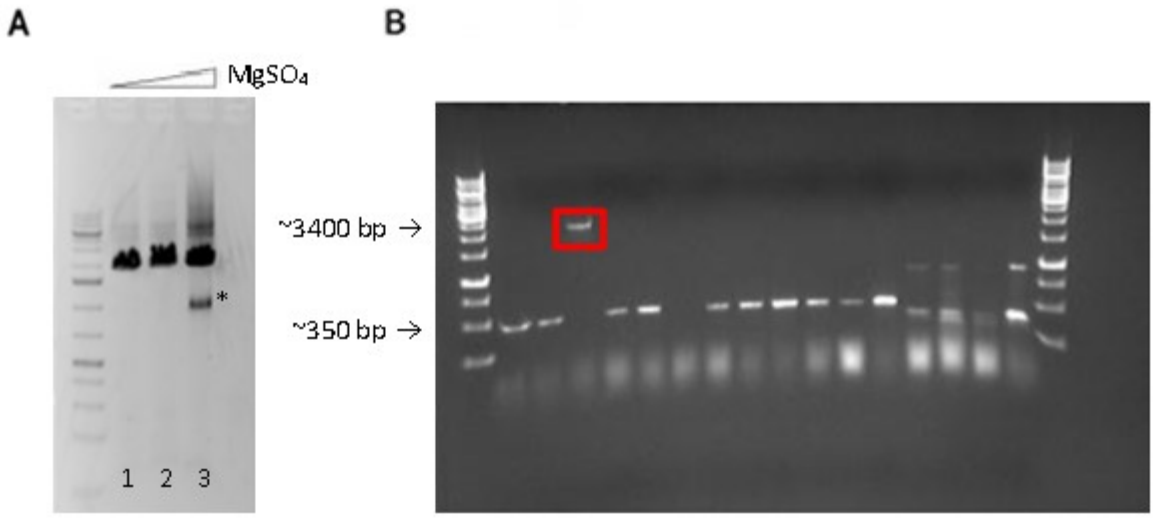


Figure 3-6. Cloning and integration of pFA6a-3HA-kanMX6 cassette into SF3b4. A) Gel electrophoresis of PCR products containing the pFA6a-3HA-kanMX6 cassette flanked by 60 bp upstream and downstream of the SF3B1 gene stop codon. Correct product size observed in lanes 1 and 2. Lane 3 contained an additional uncharacterized contaminant product denoted by *. Reactions run at increasing concentrations of Mg^{2+}. B) Colony PCR of *S. pombe* transformed with products from lane 1 [panel A]. The expected PCR product for the SF3b4 gene containing the 3HA-kanMX6 cassette is ~3400 bp. One positive colony was isolated, red box surrounding the correct product.

However, when we transformed the pFA6a-3HA-kanMX6 vector containing sequence homology to the C-terminus of SF3b49 in WT JK484 *S. pombe* cells, many colonies grew on G418 selection plates. To confirm that the vector had integrated at the 3′ end of the SF3b49 gene, and not elsewhere, we isolated genomic DNA from transformants and performed PCR using primers flanking the anticipated site of integration. After screening nearly 100 colonies, we isolated one positive colony which gave back an amplified fragment at ~3.4 kb (Figure 3-6B).

Given that the Δ p14 *S. pombe* strain we were using already contained a G418 resistance marker (used to knockout p14), we chose an alternative integration vector, pFA6a-3HA-natMX6, which when integrated at the C-terminus of SF3b49, gives the Δ p14 *S. pombe* strain immunity to the

antibiotic nourseothricin [72]. However, despite following the exact same protocol as before, we were unable to isolate any positive colonies containing tagged SF3b49 protein (Figure 3-7).

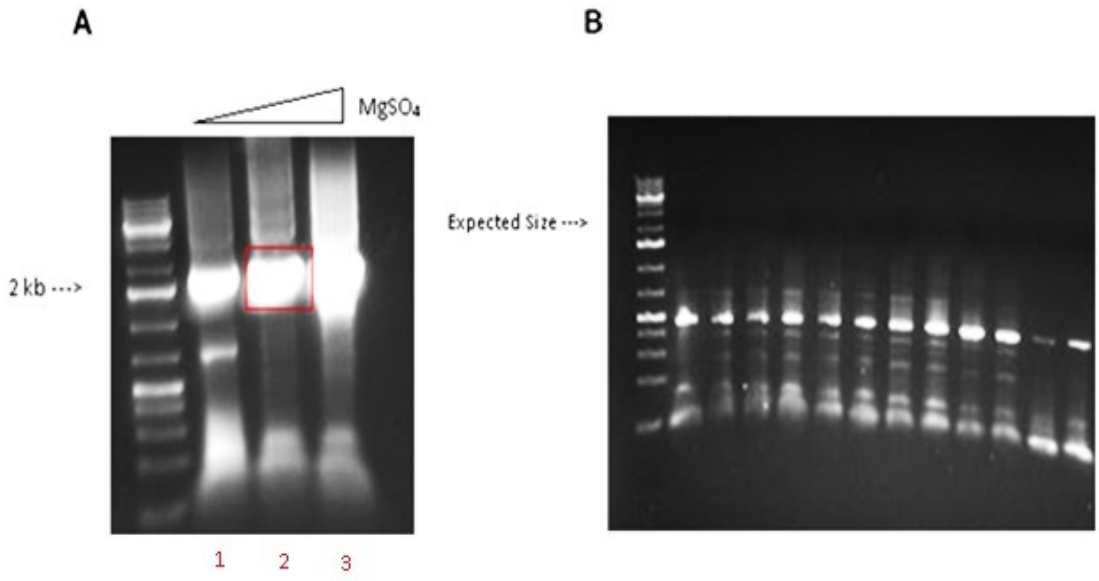


Figure 3-7. Attempt at tagging SF3b4 in Δp14 *S. pombe* with pFA6a-3HA-natMX6 cassette. A) Agarose gel of PCR products containing the pFA6a-3HA-natMX6 cassette flanked by 60 bp upstream and downstream of the SF3B4 gene stop codon. Correct product seen in lane 2 indicated by the red box. B) Colony PCR of *S. pombe* transformed with products from lane 2 [panel A]. The expected PCR product for the SF3b4 gene containing the 3HA-natMX6 cassette is ~3.4kb. All colonies appear to lack integration of the pFA6a-3HA-natMX6 cassette.

Follow-up experiments were performed to confirm expression of the 3x HA tag at the C-terminus of the SF3b49 protein in WT *S. pombe*. To do this, we first purified yeast protein extract from the positive transformant and then performed Western analysis. We used anti-HA mouse primary antibody followed by HRP Goat anti-Mouse secondary antibody to detect the HA epitope. In comparison with the control (untagged WT), we observed a distinct band at 36 kDa which correlates with the expected size of SF3b49 (Figure 3-8). Furthermore, the light and heavy chains of the antibodies we used could also be visualized at 25 and 50 kDa, respectively.

Now that we had confirmed that the HA tag was indeed being expressed at the C-terminus of SF3b49, we wanted to see whether we could immunoprecipitate and isolate the SF3b complex. We did this by incubating Protein G Sepharose beads bound to anti-HA mouse antibodies with yeast protein extract. Given our prior knowledge that the SF3b complex is salt dissociable, PBS buffer consisting of increasing concentrations of sodium chloride (137, 400, 600, 800 mM) was used to wash the protein bound Sepharose beads in an attempt to displace the complex. Aliquots of each wash solution, the original protein extract, and the final eluted product were run on an SDS-PAGE gel and subjected to Western analysis. Our data showed that we were unable to successfully dissociate the SF3b complex using even the highest concentration of NaCl, 800 mM (Figure 3-9A). Given that the SF3b complex is a component of the U2 snRNP, we wanted to see whether we could amplify U2 small RNA sequence from any of the washes. To do this, we purified RNA from each of the wash solutions, the original protein extract, and the final eluted product and used them as template for reverse transcription PCR (RT-PCR). Interestingly, we saw bands corresponding to the correct size of the U2 small RNA in three sequential washes that contained 137, 400, and 600 mM NaCl, and also the final elution (Figure 3-9B).

At this time, we are unable to confirm isolation of the SF3b complex which we had hoped to analyze using cryo-EM.

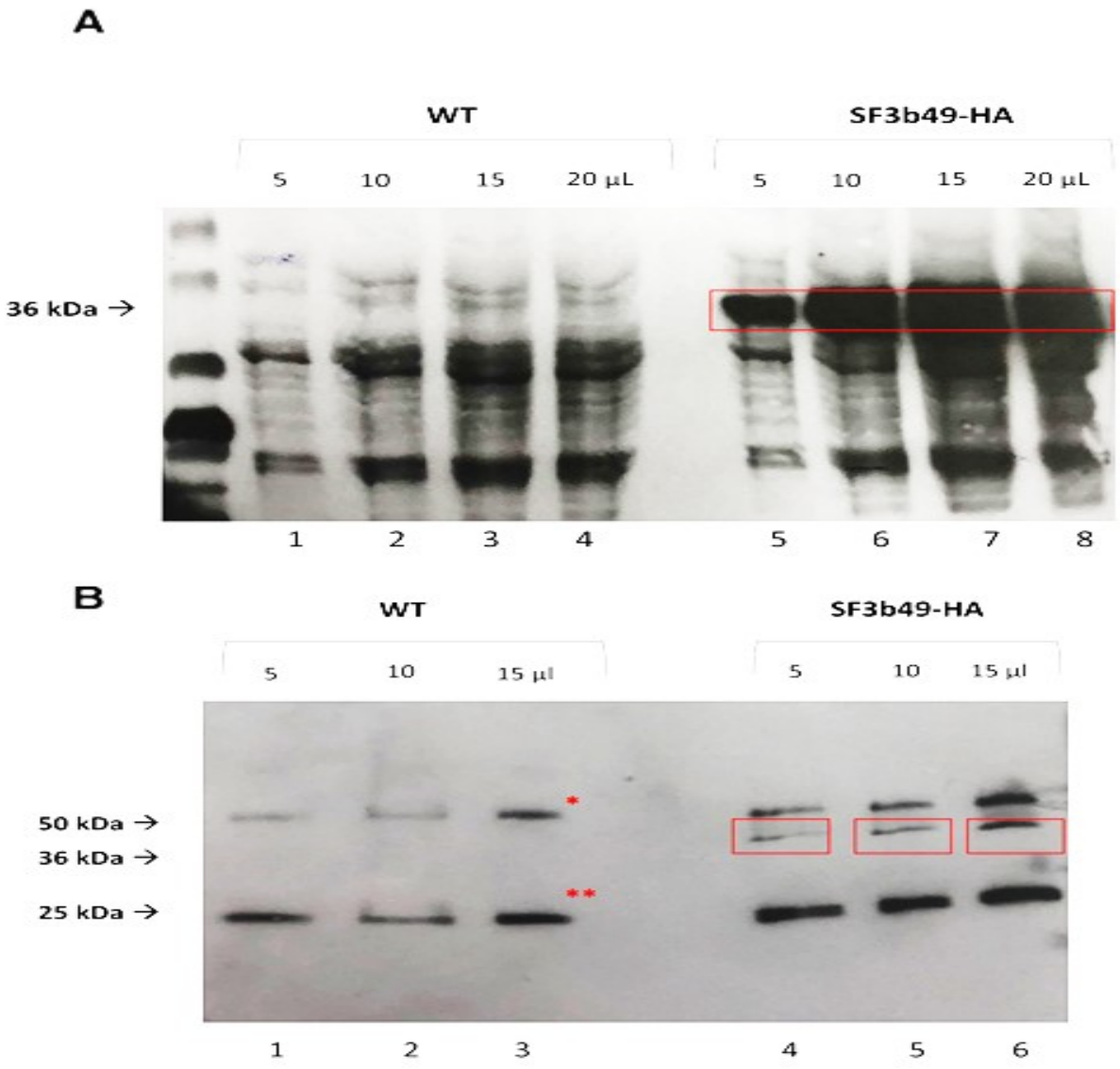


Figure 3-8. Western analysis confirming SF3b49-3HA tag. A) Protein extract from WT and PCR confirmed SF3b49-HA tagged *S. pombe* cells were purified. Different volumes of extract at a concentration of 3.85 mg/mL were run on an SDS-PAGE gel followed by the addition of anti-HA mouse antibody. Secondary goat anti-mouse (HRP conjugate) antibody was used and blot was treated with chemiluminescent HRP substrate and visualised via autoradiography. The expected size for SF3b49 is ~36kDa which is seen in red boxes [lane5-8]. Other bands indicate uncharacterized off-site interactions. B) Same protocol as in panel A with a longer duration of BSA membrane block. * denote non-specific interactions. Expected 36 kDa SF3b4-HA seen in lanes [4-6].

3-4.1. Mass Spectrometry Analysis of Proteins Expressed in WT and Δ p14 *S. pombe*.

Given that p14 is a splicing protein which has been shown to directly contact the branch adenosine, another possible way to study its function is to analyze the protein profiles of WT vs p14 knockout cells. Differences in protein expression could perhaps point towards p14's role in selecting a subset of branch sequences within introns to be spliced. To study this, we performed mass spectrometry analysis on protein extracts isolated from the two strains. The data showed a total of 115 proteins that were significantly downregulated ($p<0.05$) in the knockout strain compared to WT (Figure 3-10). Using the STRING database, which currently covers nearly 25,000,000 proteins from over 5090 organisms, we were able to look for any protein-protein interactions amongst the downregulated proteins. Interestingly, we found that 37 proteins played a role in organo-nitrogen compound metabolic processes, and 9 proteins functioning as positive regulators of mitotic metaphase/anaphase transition (Figure 3-11). Furthermore, functions of other downregulated proteins included: proteasomal ubiquitin-independent protein catabolic processes, ribose phosphate metabolic processes, nucleotide metabolic processes, and peptide metabolic processes.

.

Conversely, a total of 51 proteins were upregulated ($p<0.05$) in the knockout strain compared to WT. Once again, these proteins were subjected to STRING database analysis to look for protein-protein interactions. The results showed that there were less interactions between these proteins as compared with the downregulated proteins. However, some of these protein had similar functions to those that had been downregulated. For example, upregulated proteins included those that played roles in organo-nitrogen compound metabolic processes, nucleobase-containing small molecule metabolic processes, and sulfur compound metabolic processes.

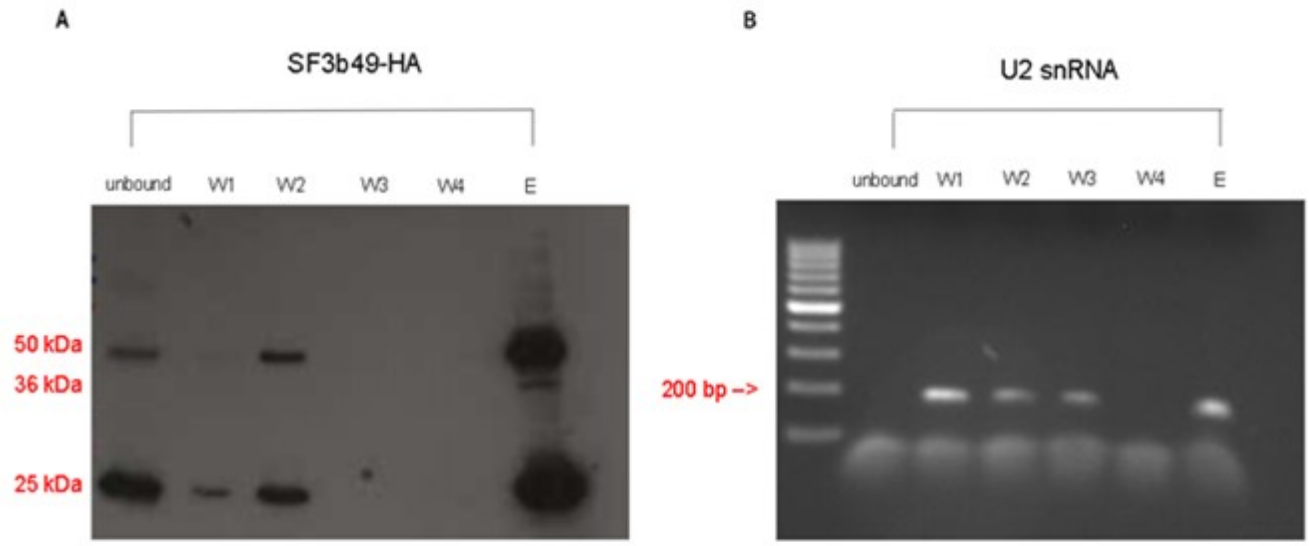


Figure 3-9. RT-PCR and Western Analysis Following SF3b49 Immunoprecipitation. A) Western analysis was performed on wash solutions (W 1-4) of increasing salt concentration (137,400,600,800 mM NaCl), Protein G Sepharose® beads bound to anti-HA mouse antibodies but not incubated with yeast extract (unbound), and the final elution using denaturing conditions (E). The heavy and light chains of the anti-HA mouse antibody can be seen at 50 and 25 kDa, respectively. The expected size of the SF3b49 protein is 36 kDa. B) Similarly, RNA was purified from the aforementioned solutions and subjected to RT-PCR using primers for the U2 snRNA. The expected size of the RNA is ~200bp

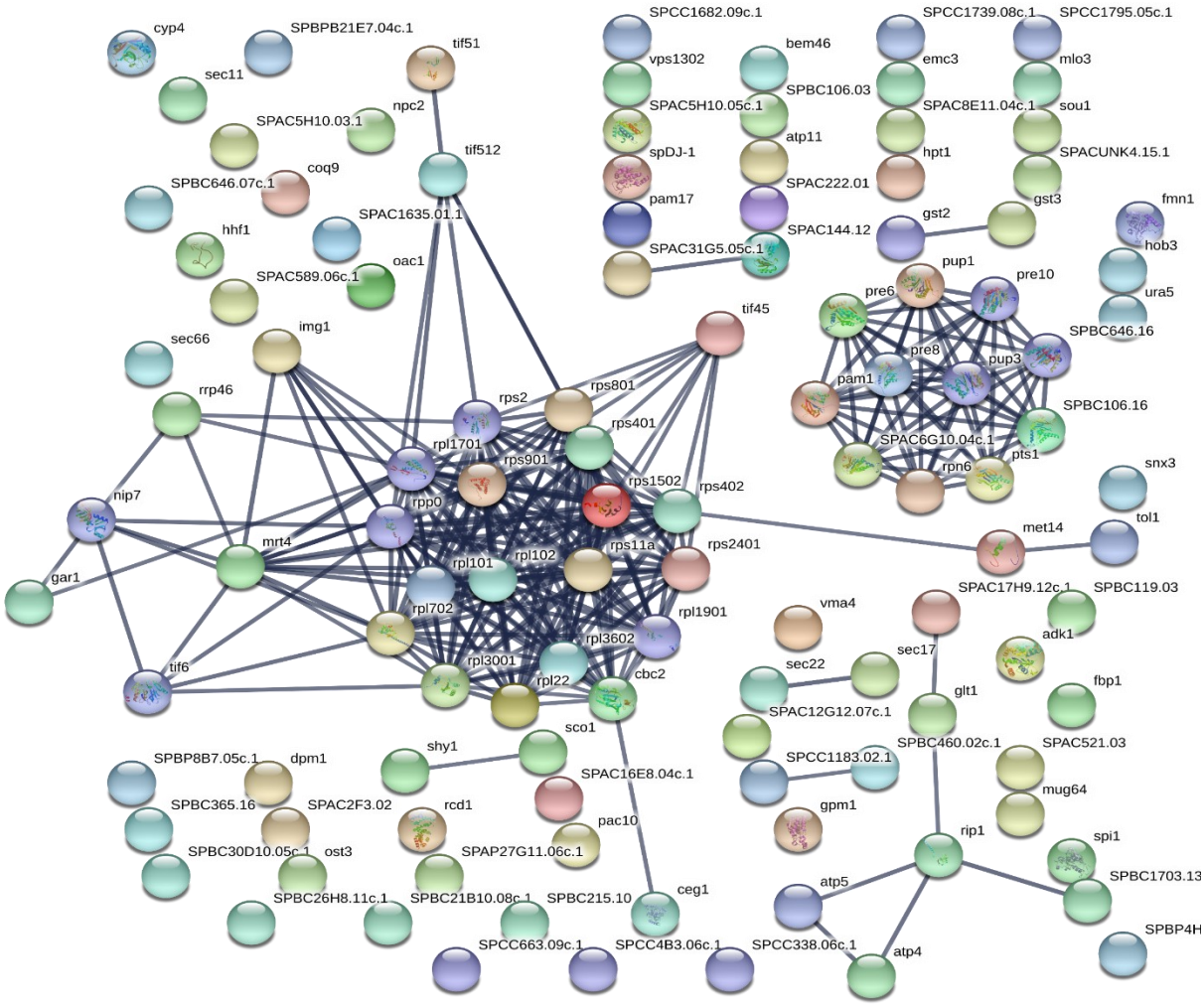


Figure 3-10. STRING Database generated protein-protein interactions between downregulated proteins in the Δp14 *S. pombe* strain. Cell extracts from WT and Δ p14 *S. pombe* were purified and run on SDS-PAGE gel followed by trypsinization. Mass Spectrometry analysis was subsequently performed comparing the protein expression profiles between the strains. A total of 115 proteins were found to be significantly downregulated in the Δp14 strain compared to WT. These proteins were run through the STRING database to look for interactions and the following diagram was generated. Each sphere represents an independent protein, lines between two spheres indicate biological interactions (>90% confidence) predicted by the STRING database. The large cluster denotes proteins important in organonitrogen compound metabolic processes required for the synthesis of amino acids, amines, amino carboxylic acids. The small cluster shows Proteins that play a critical role in the positive regulation of mitotic metaphase to anaphase transition.

Chapter 4
Discussion and Future Direction

4-1. DISCUSSION AND FUTURE DIRECTION:

Several years ago, the U2 snRNP associated protein, p14, was discovered, yet the exact role it plays in splicing remains unknown. However, with the help of recent structural and biochemical studies we have gained new insights and inched closer to cracking the mystery.

Perhaps one of the most intriguing aspects of p14 is that it appears to be essential for the viability of human cells. Furthermore, orthologs of this protein exist in other mammalian species as well as the fission yeast, *Schizosaccharomyces pombe*. However, a key difference which has allowed us to study its function has been that it is not required for the survival of *S. pombe,* our model organism. Therefore, comparative experiments between WT and p14 knockout in this species has been made possible.

While we knew that p14 was not essential for the viability of *S. pombe*, we wondered whether the absence of it would impair cell growth. To test this, we did growth curve assays comparing WT and Δ p14 *S. pombe* cells. Interestingly, we discovered that when cells grown to log phase (OD 0.1-0.6) were inoculated in fresh media, there were no difference in growth rates between the two strains. However, when cells were inoculated from saturation (OD >1.0), the knockout strain grew much slower. The reason for this remains unclear, however, it may have to do with the downregulation of certain proteins. Our mass spectrometry analysis showed a total of 115 downregulated protein in the knockout strain compared to WT, with many of those proteins being implicated in the regulation of metabolic and mitotic processes. While it isn't entirely clear which protein(s) may be involved, it does warrant future investigation as a possible reason for growth

suppression. Alternatively, and perhaps more informative, it would behoove us to do experiments which focus on RNA transcript levels within cells. Two possible experiments include: RNA-seq, and intron reporter assays. With the former, we would be able to compare mRNA transcript levels between WT and Δ p14 *S. pombe* cells to determine whether p14 is required to splice certain intron sequences. The latter experiment involves cloning introns with degenerate branch sequences into a vector containing the open reading frame (ORF) of a fluorophore protein. After transformation, if p14 were indeed required for the splicing of certain introns, we would expect to see the absence of fluorescence in Δ p14 *S. pombe* cells transformed with certain branch sequences [73, 74].

Another strategy to study the function of p14 is look for morphological differences between WT and knockout. From literature, we know that *S. pombe* cells are rod-shaped measuring 4-5 μm in diameter and 8-15 μm in length [75]. Interestingly, using bright field imaging we observed a few differences between the strains. For example, when the Δ p14 strain was grown to log phase, the cells commonly congregated in clusters of up to 10 cells at time. Furthermore, these cells were unusually elongated compared to the WT. A possible explanation for this points to the downregulation of proteins observed by mass spectrometry analysis (see above). Could it be that these cells lack expression of key proteins causing them to stick to one another? Differences also existed when the two strains were grown to saturation. Similarly, we could see that cells lacking p14 were more elongated than WT, but that they also contained numerous spores. The significance, if any, is not quite clear at this moment. I think it would be important to in future studies to rigorously analyze

Stress tests have historically been used as a way to study protein function [76-81]. We decided to compare the resilience of WT vs Δ p14 *S. pombe* cells to various stress reagents/ conditions. As mentioned in the results section, we used the following: hydroxyurea (ribonucleotide reductase inhibitor), sorbitol (osmotic stress), hydrogen peroxide (oxidative stress), methyl methanesulfonate (DNA alkylating agent), heat stress (37- 42°C), cold stress (18°C). Amongst these conditions, the knockout strain was sensitive to hydroxyurea at concentrations of 5 mM or greater. This reagent works by blocking the conversion of ribonucleotides to deoxyribonucleotides, which are required for DNA synthesis and repair [82, 83]. The underlying reason for this sensitivity is not well understood, but could perhaps be related to the downregulation of the 37 proteins that play a role in organo-nitrogen compound metabolic processes. Another finding that has no clear explanation is that the knockout strain is more resilient than WT to hydrogen peroxide. We wondered whether the knockout strain was more resistant to taking up the reagent, or perhaps the upregulation or downregulation of certain proteins was contributory.

To gain a closer look at the SF3b complex, with and without p14, we utilized a tagging strategy to isolate the particle for future study via cryo-EM [84-87]. This was made possible in the WT strain by using the yeast genomic integration vector, pFA6a-3HA-natMX6, to insert a 3x HA coding sequence at the 3′ end of the SF3b49 gene. Using Western analysis, we confirmed expression of the tag and proceeded to immunoprecipitate the complex from yeast protein extract. While we found success binding the complex to Protein G Sepharose beads, we had difficulties eluting it under non-denaturing conditions. Perhaps the reason for this is because we did not use harsh enough detergents or used sub-optimal salt concentrations in our wash buffer. We were however

able to amplify U2 small RNA from the wash buffer using RT PCR. In the future, different combinations of the aforementioned conditions could be tested to see whether dissociation of the complex can be achieved.

Unfortunately, despite ongoing efforts we were unable to tag the Δ p14 *S. pombe* strain. First, we had to use a different vector, pFA6a-3HA-natMX6, because the knockout strain already contained a G418 antibiotic selection marker. Although it is a very similar vector, it may be less efficient in recombining into the genome. Alternatively, the lack of p14 in the SF3b complex could result in the structural rearrangement of the SF3b proteins, therefore changing the location of SF3b49 C-terminus. If this were true, then having a C-terminal tag on this protein may affect folding of the SF3b complex, rendering cells non-viable. Thus, it may be worth trying to tag different components of the SF3b complex, which had previously been unsuccessful in the WT. If doable, we could one day perhaps purify WT and Δ p14 SF3b complexes and structurally study them using single particle cryo-EM.

Chapter 5
References

5-1. REFERENCES.

1. Berget, S.M., C. Moore, and P.A. Sharp, *Spliced segments at the 5' terminus of adenovirus 2 late mRNA.* Proc Natl Acad Sci U S A, 1977. **74**(8): p. 3171-5.

2. Chow, L.T., et al., *An amazing sequence arrangement at the 5' ends of adenovirus 2 messenger RNA.* Cell, 1977. **12**(1): p. 1-8.

3. Padgett, R.A., et al., *Splicing of messenger RNA precursors.* Annu Rev Biochem, 1986. **55**: p. 1119-50.

4. Rio, D.C., *Splicing of pre-mRNA: mechanism, regulation and role in development.* Curr Opin Genet Dev, 1993. **3**(4): p. 574-84.

5. Lopez, A.J., *Alternative splicing of pre-mRNA: developmental consequences and mechanisms of regulation.* Annu Rev Genet, 1998. **32**: p. 279-305.

6. Pan, Q., et al., *Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing.* Nat Genet, 2008. **40**(12): p. 1413-5.

7. Wang, E.T., et al., *Alternative isoform regulation in human tissue transcriptomes.* Nature, 2008. **456**(7221): p. 470-6.

8. Wang, H., et al., *Genome-wide analysis of alternative splicing during human heart development.* Sci Rep, 2016. **6**: p. 35520.

9. Kim, S.H. and R.J. Lin, *Spliceosome activation by PRP2 ATPase prior to the first transesterification reaction of pre-mRNA splicing.* Mol Cell Biol, 1996. **16**(12): p. 6810-9.

10. Dib-Hajj, S.D., et al., *Domain 5 interacts with domain 6 and influences the second transesterification reaction of group II intron self-splicing.* Nucleic Acids Res, 1993. **21**(8): p. 1797-804.

11. Qin, D., et al., *Sequencing of lariat termini in S. cerevisiae reveals 5′ splice sites, branch points, and novel splicing events.* RNA, 2016. **22**(2): p. 237-53.

12. Spingola, M., et al., *Genome-wide bioinformatic and molecular analysis of introns in Saccharomyces cerevisiae.* RNA, 1999. **5**(2): p. 221-34.

13. Castanotto, D. and J.J. Rossi, *Small Sequence Insertions within the Branch Point Region Dictate Alternative Sites of Lariat Formation in a Yeast Intron.* Nucleic Acids Research, 1992. **20**(24): p. 6649-6655.

14. Liu, X.D. and J.E. Mertz, *Sequence of the polypyrimidine tract of the 3′ -terminal 3′ splicing signal can affect intron-dependent pre-mRNA processing in vivo.* Nucleic Acids Research, 1996. **24**(9): p. 1765-1773.

15. Kramer, A., *The structure and function of proteins involved in mammalian pre-mRNA splicing.* Annu Rev Biochem, 1996. **65**: p. 367-409.

16. Gao, K., et al., *Human branch point consensus sequence is yUnAy.* Nucleic Acids Res, 2008. **36**(7): p. 2257-67.

17. Sharp, P.A., *On the origin of RNA splicing and introns.* Cell, 1985. **42**(2): p. 397-400.

18. Query, C.C., M.J. Moore, and P.A. Sharp, *Branch nucleophile selection in pre-mRNA splicing: evidence for the bulged duplex model.* Genes Dev, 1994. **8**(5): p. 587-97.

19. Hesselberth, J.R., *Lives that introns lead after splicing.* WIREs RNA, 2013. **4**(6): p. 677-691.

20. Matlin, A.J. and M.J. Moore, *Spliceosome assembly and composition.* Adv Exp Med Biol, 2007. **623**: p. 14-35.

21. Seraphin, B., L. Kretzner, and M. Rosbash, *A U1 snRNA:pre-mRNA base pairing interaction is required early in yeast spliceosome assembly but does not uniquely define the 5′ cleavage site.* EMBO J, 1988. **7**(8): p. 2533-8.

22. Mount, S.M., et al., *The U1 Small Nuclear Rna-Protein Complex Selectively Binds a 5′ Splice Site Invitro.* Cell, 1983. **33**(2): p. 509-518.

23. Berglund, J.A., N. Abovich, and M. Rosbash, *A cooperative interaction between U2AF65 and mBBP/SF1 facilitates branchpoint region recognition.* Genes Dev, 1998. **12**(6): p. 858-67.

24. Berglund, J.A., M. Rosbash, and S.C. Schultz, *Crystal structure of a model branchpoint-U2 snRNA duplex containing bulged adenosines.* RNA, 2001. **7**(5): p. 682-91.

25. MacMillan, A.M., et al., *Dynamic association of proteins with the pre-mRNA branch region.* Genes Dev, 1994. **8**(24): p. 3008-20.

26. Lardelli, R.M., et al., *Release of SF3 from the intron branchpoint activates the first step of pre-mRNA splicing.* RNA, 2010. **16**(3): p. 516-28.

27. Gozani, O., R. Feld, and R. Reed, *Evidence that sequence-independent binding of highly conserved U2 snRNP proteins upstream of the branch site is required for assembly of spliceosomal complex A.* Genes Dev, 1996. **10**(2): p. 233-43.

28. Dybkov, O., et al., *U2 snRNA-protein contacts in purified human 17S U2 snRNPs and in spliceosomal A and B complexes.* Mol Cell Biol, 2006. **26**(7): p. 2803-16.

29. Ilagan, J., et al., *The role of exon sequences in C complex spliceosome structure.* J Mol Biol, 2009. **394**(2): p. 363-75.

30. Matera, A.G. and Z. Wang, *A day in the life of the spliceosome.* Nat Rev Mol Cell Biol, 2014. **15**(2): p. 108-21.

31. Kramer, A., M. Frick, and W. Keller, *Separation of multiple components of HeLa cell nuclear extracts required for pre-messenger RNA splicing.* J Biol Chem, 1987. **262**(36): p. 17630-40.

32. Zhang, X., et al., *Structure of the human activated spliceosome in three conformational states.* Cell Res, 2018. **28**(3): p. 307-322.

33. Kramer, A. and U. Utans, *Three protein factors (SF1, SF3 and U2AF) function in pre-splicing complex formation in addition to snRNPs.* EMBO J, 1991. **10**(6): p. 1503-9.

34. Yan, D. and M. Ares, Jr., *Invariant U2 RNA sequences bordering the branchpoint recognition region are essential for interaction with yeast SF3a and SF3b subunits.* Mol Cell Biol, 1996. **16**(3): p. 818-28.

35. Will, C.L., et al., *Characterization of novel SF3b and 17S U2 snRNP proteins, including a human Prp5p homologue and an SF3b DEAD-box protein.* EMBO J, 2002. **21**(18): p. 4978-88.

36. Wang, Q., et al., *Interactions of the yeast SF3b splicing factor.* Mol Cell Biol, 2005. **25**(24): p. 10745-54.

37. Sun, C., *The SF3b complex: splicing and beyond.* Cell Mol Life Sci, 2020.

38. Query, C.C., P.S. McCaw, and P.A. Sharp, *A minimal spliceosomal complex A recognizes the branch site and polypyrimidine tract.* Mol Cell Biol, 1997. **17**(5): p. 2944-53.

39. Cass, D.M. and J.A. Berglund, *The SF3b155 N-Terminal Domain Is a Scaffold Important for Splicing.* Biochemistry, 2006. **45**(33): p. 10092-10101.

40. Spadaccini, R., et al., *Biochemical and NMR analyses of an SF3b155-p14-U2AF-RNA interaction network involved in branch point definition during pre-mRNA splicing.* RNA, 2006. **12**(3): p. 410-25.

41. Cretu, C., et al., *Molecular Architecture of SF3b and Structural Consequences of Its Cancer-Related Mutations.* Mol Cell, 2016. **64**(2): p. 307-319.

42. Schellenberg, M.J., et al., *Crystal structure of a core spliceosomal protein interface.* Proc Natl Acad Sci U S A, 2006. **103**(5): p. 1266-71.

43. Li, T., et al., *Structure of DDB1 in complex with a paramyxovirus V protein: viral hijack of a propeller cluster in ubiquitin ligase.* Cell, 2006. **124**(1): p. 105-17.

44. Fischer, E.S., et al., *Structure of the DDB1-CRBN E3 ubiquitin ligase in complex with thalidomide.* Nature, 2014. **512**(7512): p. 49-53.

45. Furney, S.J., et al., *SF3B1 mutations are associated with alternative splicing in uveal melanoma.* Cancer Discov, 2013. **3**(10): p. 1122-1129.

46. Scott, L.M. and V.I. Rebel, *Acquired mutations that affect pre-mRNA splicing in hematologic malignancies and solid tumors.* J Natl Cancer Inst, 2013. **105**(20): p. 1540-9.

47. Malcovati, L., et al., *Clinical significance of SF3B1 mutations in myelodysplastic syndromes and myelodysplastic/myeloproliferative neoplasms.* Blood, 2011. **118**(24): p. 6239-46.

48. Malcovati, L., et al., *SF3B1 mutation identifies a distinct subset of myelodysplastic syndrome with ring sideroblasts.* Blood, 2015. **126**(2): p. 233-41.

49. Zhang, J., et al., *Disease-Causing Mutations in SF3B1 Alter Splicing by Disrupting Interaction with SUGP1.* Mol Cell, 2019. **76**(1): p. 82-95 e7.

50. Liu, Z., et al., *Pan-cancer analysis identifies mutations in SUGP1 that recapitulate mutant SF3B1 splicing dysregulation.* Proc Natl Acad Sci U S A, 2020.

51. Prusty, A.B., et al., *Impaired spliceosomal UsnRNP assembly leads to Sm mRNA down-regulation and Sm protein degradation.* J Cell Biol, 2017. **216**(8): p. 2391-2407.

52. Blencowe, B.J., et al., *A coactivator of pre-mRNA splicing.* Genes Dev, 1998. **12**(7): p. 996-1009.

53. Meffert, R., et al., *Ultraviolet crosslinking of DNA-protein complexes via 8-azidoadenine.* Methods Mol Biol, 1994. **30**: p. 237-50.

54. Meffert, R. and K. Dose, *UV-induced cross-linking of proteins to plasmid pBR322 containing 8-azidoadenine 2'-deoxyribonucleotides.* FEBS Lett, 1988. **239**(2): p. 190-4.

55. Query, C.C., S.A. Strobel, and P.A. Sharp, *Three recognition events at the branch-site adenine.* EMBO J, 1996. **15**(6): p. 1392-402.

56. Will, C.L., et al., *A novel U2 and U11/U12 snRNP protein that associates with the pre-mRNA branch site.* EMBO J, 2001. **20**(16): p. 4536-46.

57. Gottschalk, A., et al., *A novel yeast U2 snRNP protein, Snu17p, is required for the first catalytic step of splicing and for progression of spliceosome assembly.* Mol Cell Biol, 2001. **21**(9): p. 3037-46.

58. Kenan, D.J., C.C. Query, and J.D. Keene, *RNA recognition: towards identifying determinants of specificity.* Trends Biochem Sci, 1991. **16**(6): p. 214-20.

59. Deo, R.C., et al., *Recognition of polyadenylate RNA by the poly(A)-binding protein.* Cell, 1999. **98**(6): p. 835-45.

60. Oubridge, C., et al., *Crystal structure at 1.92 Å resolution of the RNA-binding domain of the U1A spliceosomal protein complexed with an RNA hairpin.* Nature, 1994. **372**(6505): p. 432-438.

61. Birney, E., S. Kumar, and A.R. Krainer, *Analysis of the RNA-recognition motif and RS and RGG domains: conservation in metazoan pre-mRNA splicing factors.* Nucleic Acids Res, 1993. **21**(25): p. 5803-16.

62. Schellenberg, M.J., E.L. Dul, and A.M. MacMillan, *Structural model of the p14/SF3b155 . branch duplex complex.* RNA, 2011. **17**(1): p. 155-65.

63. Cohen, S.B. and T.R. Cech, *Engineering disulfide cross-links in RNA using thiol-disulfide interchange chemistry.* Curr Protoc Nucleic Acid Chem, 2001. **Chapter 5**: p. Unit 5 1.

64. Maglott, E.J. and G.D. Glick, *Probing structural elements in RNA using engineered disulfide cross-links.* Nucleic Acids Res, 1998. **26**(5): p. 1301-8.

65. Gilbert, H.F., *[2] Thiol/disulfide exchange equilibria and disulfidebond stability*, in *Methods in Enzymology*. 1995, Academic Press. p. 8-28.

66. He, C. and G.L. Verdine, *Trapping Distinct Structural States of a Protein/DNA Interaction through Disulfide Crosslinking.* Chemistry & Biology, 2002. **9**(12): p. 1297-1303.

67. Huang, H., S.C. Harrison, and G.L. Verdine, *Trapping of a catalytic HIV reverse transcriptase·template:primer complex through a disulfide bond.* Chemistry & Biology, 2000. **7**(5): p. 355-364.

68. Forsburg, S.L. and N. Rhind, *Basic methods for fission yeast.* Yeast, 2006. **23**(3): p. 173-83.

69. Suga, M., M. Isobe, and T. Hatakeyama, *Cryopreservation of competent intact yeast cells for efficient electroporation.* Yeast, 2000. **16**(10): p. 889-96.

70. Sato, M., S. Dhut, and T. Toda, *New drug-resistant cassettes for gene disruption and epitope tagging in Schizosaccharomyces pombe.* Yeast, 2005. **22**(7): p. 583-91.

71. Gadaleta, M.C., et al., *New vectors for epitope tagging and gene disruption in Schizosaccharomyces pombe.* Biotechniques, 2013. **55**(5): p. 257-63.

72. Krugel, H., et al., *Analysis of the nourseothricin-resistance gene (nat) of Streptomyces noursei.* Gene, 1988. **62**(2): p. 209-17.

73. Nasim, M.T. and I.C. Eperon, *A double-reporter splicing assay for determining splicing efficiency in mammalian cells.* Nat Protoc, 2006. **1**(2): p. 1022-8.

74. Carle-Urioste, J.C., et al., *In vivo analysis of intron processing using splicing-dependent reporter gene assays.* Plant Mol Biol, 1994. **26**(6): p. 1785-95.

75. Roque, H. and C. Antony, *Fission yeast a cellular model well suited for electron microscopy investigations.* Methods Cell Biol, 2010. **96**: p. 235-58.

76. Gonzalez, R., et al., *New Genes Involved in Osmotic Stress Tolerance in Saccharomyces cerevisiae.* Front Microbiol, 2016. **7**: p. 1545.

77. Tang, H.-M.V., et al., *Loss of APD1 in Yeast Confers Hydroxyurea Sensitivity Suppressed by Yap1p Transcription Factor.* Scientific Reports, 2015. **5**(1): p. 7897.

78. Brennan, R.J. and R.H. Schiestl, *Cadmium is an inducer of oxidative stress in yeast.* Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis, 1996. **356**(2): p. 171-178.

79. Tran, K. and E.M. Green, *Assessing Yeast Cell Survival Following Hydrogen Peroxide Exposure.* Bio Protoc, 2019. **9**(2).

80. Soto, T., et al., *Cold induces stress-activated protein kinase-mediated response in the fission yeast Schizosaccharomyces pombe.* European Journal of Biochemistry, 2002. **269**(20): p. 5056-5065.

81. Chang, M., et al., *A genome-wide screen for methyl methanesulfonate-sensitive mutants reveals genes required for S phase progression in the presence of DNA damage.* Proceedings of the National Academy of Sciences, 2002. **99**(26): p. 16934.

82. Koc, A., et al., *Hydroxyurea arrests DNA replication by a mechanism that preserves basal dNTP pools.* J Biol Chem, 2004. **279**(1): p. 223-30.

83. Chapman, T.R. and T.J. Kinsella, *Ribonucleotide reductase inhibitors: a new look at an old target for radiosensitization.* Frontiers in oncology, 2012. **1**: p. 56-56.

84. Finci, L.I., et al., *The cryo-EM structure of the SF3b spliceosome complex bound to a splicing modulator reveals a pre-mRNA substrate competitive mechanism of action.* Genes Dev, 2018. **32**(3-4): p. 309-320.

85. Nwanochie, E. and V.N. Uversky, *Structure Determination by Single-Particle Cryo-Electron Microscopy: Only the Sky (and Intrinsic Disorder) is the Limit.* International journal of molecular sciences, 2019. **20**(17): p. 4186.

86. Murata, K. and M. Wolf, *Cryo-electron microscopy for structural analysis of dynamic biological macromolecules.* Biochimica et Biophysica Acta (BBA) - General Subjects, 2018. **1862**(2): p. 324-334.

87. Costa, T.R.D., A. Ignatiou, and E.V. Orlova, *Structural Analysis of Protein Complexes by Cryo Electron Microscopy*, in *Bacterial Protein Secretion Systems: Methods and Protocols*, L. Journet and E. Cascales, Editors. 2017, Springer New York: New York, NY. p. 377-413.

Appendix

Gene Name	KO-1: Corrected Area	KO-2: Corrected Area	KO-3: Corrected Area	WT-1: Corrected Area	WT-2: Corrected Area	WT-3: Corrected Area	P-value
ceg1	0	1.94E-05	0	0.000458	0.000864	0.000636	0.031113
SPCC4B3.06c.1	3.44E-05	0	0	0.000449	0.000659	0.000743	0.018771
vma4	0	9.51E-05	0	0.000941	0.001177	0.00191	0.044711
rpl3602	2.24E-05	2.46E-05	1.5E-05	0.000892	0.00053	0.00097	0.029074
mrt4	0	0.000122	0	0.001026	0.001763	0.001146	0.027639
SPAC17H9.12c.1	6.84E-05	3.93E-05	0	0.00085	0.001305	0.001072	0.014014
sec72	6.71E-05	0	0	0.00045	0.000449	0.000352	0.000991
SPAC144.12	8.24E-05	0	0	0.000483	0.00037	0.000468	0.001023
SPAC5H10.03	0.000166	0.00017	0.000278	0.003498	0.003176	0.002862	0.00282
pac10	4.82E-05	0	0	0.000294	0.000242	0.000198	0.004662
SPAC222.01	0.002184	0.00129	0.001608	0.022025	0.02577	0.029173	0.006647
atp5	0.00011	4.86E-05	4.83E-05	0.000911	0.001119	0.000714	0.016315
adk1	0.000451	0.000178	0.000117	0.003129	0.003441	0.00285	0.000439
Q9UST4	0.000163	0.000118	0.000327	0.00191	0.002851	0.002528	0.011723
SPAC13C5.01c.1	0.000184	0.000232	6.35E-05	0.001877	0.002006	0.001828	1.86E-05
tif512	0	0.000169	9.13E-05	0.000712	0.001134	0.001143	0.015589
rcd1	9.12E-05	1.78E-05	0	0.000345	0.000403	0.000471	0.001676
tif51	0.000337	0.000144	0.000312	0.002088	0.003316	0.003258	0.020534
cpy4	0.000173	0.000226	7.17E-05	0.001584	0.001721	0.001754	2.85E-05
rpl3001	9.64E-05	0.000102	0.000121	0.001124	0.000757	0.001473	0.039195
atp4	0.000245	0.000154	0.000654	0.004102	0.00407	0.002574	0.016746
atp11	3.37E-05	0	0	9.15E-05	0.000121	0.000118	0.002786
aim27	0	5.03E-05	0	0.000109	0.000196	0.000177	0.014884
rps1502	0	0.000135	8.9E-05	0.000478	0.000839	0.000754	0.019958
sec17	0	0	0.000103	0.000262	0.000254	0.000373	0.007238
rpl22	0.000799	0.000644	0.000953	0.007	0.004406	0.008753	0.041976
SPAC589.06c.1	0.000217	0.000313	0.000168	0.001719	0.002221	0.001748	0.006482
hob3	0.000188	7.42E-05	0.000147	0.000873	0.001341	0.001074	0.015123
rrp46	6.03E-05	3.17E-05	0	0.000156	0.000278	0.000299	0.028233
snx3	0	7.09E-05	4.4E-05	0.000174	0.000317	0.000378	0.041499
pre8	0.000114	0.000265	0.000113	0.001247	0.001247	0.001161	0.000262
ura5	0.000534	0.000267	0.000254	0.002151	0.002489	0.002562	0.000338
fbp1	0.000157	0.000235	0	0.000954	0.000711	0.000843	0.002021
SPCC338.06c.1	0.000119	0	0	0.000251	0.000216	0.000276	0.020614
SPCC1739.08c.1	0.000224	0.00014	0	0.000634	0.00072	0.000913	0.004449
SPAC5H10.05c.1	0	0.000104	0.000169	0.000466	0.000618	0.000609	0.002459
mlo3	0.000495	0.00021	8.25E-05	0.001229	0.001708	0.001841	0.006169

SPBC21B10.08c.1	0.000886	0.000236	0.000108	0.002495	0.002325	0.002256	0.010094
tol1	0.000136	0	0	0.000193	0.000251	0.000317	0.024383
gst3	0.015745	0.006534	0.002297	0.039332	0.047214	0.05037	0.002172
nip7	0.00016	0	0	0.000322	0.00027	0.000298	0.036811
rpl102	0.000334	0.000422	9.63E-05	0.001245	0.001728	0.001688	0.004142
mug64	0.00014	0	0	0.000214	0.000226	0.000299	0.030586
rpn6	0	0	0.000635	0.000885	0.00102	0.001371	0.032119
pam17	0	6.07E-05	0	7.72E-05	8.95E-05	0.000145	0.044885
SPAC17H9.12c.1	0.000212	0.000135	0.000117	0.000723	0.000957	0.000654	0.014092
img1	0	3.98E-05	4.13E-05	0.000106	0.000173	0.000117	0.017781
vas2	0	7.14E-05	0	7.8E-05	0.000136	0.000127	0.043171
pup1	0.000203	0.000178	0.000146	0.000635	0.001062	0.000802	0.03199
rps2401	1.86E-05	0.000305	0.000248	0.000584	0.001084	0.001033	0.027765
tif6	0.000243	0.000271	9.75E-05	0.000839	0.001221	0.000779	0.021624
dpm1	0.000619	0.000195	0.000127	0.001219	0.001579	0.001526	0.005367
met14	0.000237	0.00017	0.000272	0.001329	0.00085	0.000902	0.030153
SPAC8E11.04c.1	6.43E-05	0	5.08E-05	0.000162	0.000177	0.000179	0.015096
rpl3602	0.000974	0.000733	0.000548	0.003843	0.003071	0.003199	0.002298
spi1	0.002132	0.00206	0.002014	0.007767	0.009904	0.009299	0.008189
unk4	8.41E-05	3.97E-05	0.000112	0.000366	0.000313	0.000336	0.000868
coq9	0.000162	3.94E-05	0	0.000293	0.000282	0.000288	0.045254
tif45	0.000226	0.000144	0.000163	0.000534	0.000829	0.000897	0.030567
SPBC30D10.05c.1	0.00093	0.000874	0.000507	0.00299	0.003538	0.003086	0.000465
SPAC12G12.07c.1	0.000192	9.55E-05	9.34E-05	0.000475	0.000566	0.000539	0.000832
ost3	0	0.000778	0.000304	0.001675	0.001344	0.001288	0.023897
sou1	0.001863	0.001689	0.00108	0.005709	0.006768	0.005837	0.000631
sec66	4.66E-05	0.000129	0	0.000251	0.000187	0.000244	0.028064
gar1	0.000513	0.000189	0.000259	0.001458	0.001066	0.001152	0.004709
SPBC365.16	0.000281	0.000278	0.000358	0.00129	0.000864	0.001282	0.023859
SPBC646.07c.1	0.000646	0.000285	0.000653	0.002121	0.001356	0.002296	0.026372
SPBC215.10	0.000973	0.000638	0.000845	0.002533	0.003259	0.002906	0.003626
pre6	0.000477	0.000573	0.000632	0.001636	0.002264	0.00201	0.012467
pam1	0.000815	0.000568	0.000606	0.002041	0.002574	0.002337	0.002637
SPCC1682.09c.1	0.000231	0.000133	0.000151	0.000714	0.00051	0.000506	0.015607
h4.1	0.00171	0.000545	0.000564	0.003319	0.002526	0.003348	0.013757
vps1302	7.26E-05	4.51E-05	0	0.000145	8.7E-05	0.00015	0.039642
pts1	0.000771	0.000579	0.000312	0.00188	0.002043	0.001351	0.012382
rps402	0.000842	0.001161	0.001199	0.002508	0.003422	0.004092	0.032246
rps401	0.000785	0.000775	0.000849	0.001918	0.002777	0.002782	0.027031

SPCC1739.08c.1	0.004103	0.002802	0.001869	0.010933	0.009053	0.007221	0.013013
SPCC1183.02.1	0.001909	0.001426	0.000905	0.004724	0.004418	0.003943	0.00166
SPAC5H10.03.1	0.000874	0.000505	0.000414	0.001599	0.001948	0.001951	0.002819
sec22	0.000177	0.000149	0.000217	0.000621	0.000558	0.000486	0.003581
SPBC1703.13c.1	0.001977	0.002697	0.001992	0.006083	0.008331	0.005965	0.019753
shy1	0.000218	0.000131	0.000388	0.000606	0.000577	0.000937	0.036057
hpt1	0.000834	0.000463	0.000476	0.001566	0.001672	0.001641	0.009841
SPAC16E8.04c.1	0.000195	0.000126	0.000115	0.000328	0.000462	0.000401	0.008442
rps1101	0.000369	0.001209	0.000676	0.001845	0.002669	0.001649	0.033026
pre3	0.001498	0.000345	0.000561	0.002231	0.002345	0.00189	0.047664
oac1	0.000469	0.000365	0.000277	0.000882	0.001187	0.000908	0.010046
SPCC1795.05c.1	0.001097	0.000354	0.000283	0.001763	0.001684	0.001126	0.048488
SPBC577.10	0.000908	0.001168	0.000655	0.002068	0.002355	0.002617	0.002742
rip1	0.001699	0.000559	0.000588	0.002705	0.002771	0.001814	0.039906
sco1	0.000145	0.000147	0.00018	0.000378	0.000408	0.000412	0.000101
bem46	6.44E-05	8.17E-05	0.000127	0.000199	0.000278	0.000213	0.012257
SPCC1183.02.1	0.00205	0.001734	0.001619	0.005238	0.004409	0.003799	0.016489
Q9P6M1	0.000447	0.000359	0.000563	0.000836	0.001407	0.001086	0.04577
SPBC106.03	0.002152	0.000809	0.001018	0.003137	0.003473	0.00291	0.033965
SPAC31G5.05c.1	0.001732	0.000839	0.000933	0.002635	0.002962	0.002654	0.020198
SPAC13C5.01c.1	0.000344	0.000228	0.000437	0.000721	0.000822	0.000794	0.007643
sec11	0.000136	0.000103	8.38E-05	0.000285	0.000243	0.000206	0.010735
npc2	0.000178	0.000127	0.000142	0.000272	0.00035	0.000358	0.009672
glt1	0.002135	0.001586	0.001368	0.003216	0.004275	0.003667	0.007632
SPBP8B7.05c.1	0.001243	0.000596	0.000555	0.001418	0.00191	0.001862	0.031696
SPAC12G12.07c.1	0.000409	0.000188	0.000265	0.000646	0.000691	0.000487	0.023288
pre10	0.000813	0.000708	0.000652	0.001468	0.00161	0.001483	0.000262
pam1	0.001142	0.000771	0.001558	0.002578	0.00262	0.002061	0.013558
rps801	0.001006	0.000422	0.000585	0.001611	0.001484	0.001109	0.03469
rps901	0.006678	0.004212	0.003434	0.011789	0.010346	0.007193	0.045546
SPCC1739.08c.1	0.003087	0.00235	0.002682	0.006203	0.00561	0.004755	0.009687
SPBC106.03	0.001183	0.000962	0.000707	0.001635	0.002111	0.001833	0.009525
rpl702	0.000345	0.000461	0.000254	0.000681	0.00068	0.000623	0.026686
SPAC1635.01.1	0.007185	0.006677	0.005615	0.009877	0.013145	0.013369	0.024826
SPBPB21E7.04c.1	0.004618	0.006804	0.00555	0.008364	0.010828	0.011843	0.025466
cbc2	0.000188	9.05E-05	0.000137	0.000265	0.000244	0.000243	0.049312
rps2	0.002746	0.00318	0.002794	0.004616	0.004914	0.005503	0.005467
SPBPB21E7.04c.1	0.003209	0.003714	0.003752	0.005063	0.006191	0.006611	0.023843