Title
A Pipeline for Tagging snRNP Associated Proteins in HeLa

Permalink
https://escholarship.org/uc/item/02s0x3ns

Author
Kim, John

Publication Date
2019

License
CC BY-NC-SA 4.0

Peer reviewed|Thesis/dissertation
A PIPELINE FOR TAGGING snRNP ASSOCIATED PROTEINS IN HeLa

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

MASTER OF ARTS

in

MOLECULAR, CELLULAR, DEVELOPMENTAL BIOLOGY

by

John Kim

June 2019

The Thesis of John Kim is approved:

________________________________________________________________________

Professor Melissa Jurica

________________________________________________________________________

Professor Manuel Ares Jr.

________________________________________________________________________

Professor Jeremy Sanford

________________________________________________________________________

Lori Kletzer

Vice Provost and Dean of Graduate Studies
# Table of Contents

List of Figures v  
List of Tables vi  
Abstract vii  
Dedication ix  
Acknowledgments x  

## Chapter 1: Introduction

1.1 Eukaryotic Pre-mRNA Splicing 1  
1.2 Splicing Chemistry and Intronic Elements 2  
1.3 The Splice Cycle 3  
1.4 Current Spliceosome Structures and Purification Methods 4  
1.5 Spliceosome Assembly and Substrate Recognition 5  
1.6 Recombination Mediated Cassette Exchange 7  

## Chapter 2: Materials and Methods

2.1 Site Directed Mutagenesis PCR of Donor Plasmid 10  
2.2 Two-step PCR, Restriction Digest, and purification of V5-tagged gene insert 11  
2.3 Ligation and Transformation 12  
2.4 Cell Culture 13  
2.5 Stable Cell Line Construction 13  
2.6 Genomic DNA Extraction and PCR 14  
2.7 RNA Extraction and RT-PCR 14
2.8 Whole Cell Extraction
2.9 Mini Nuclear Extraction
2.10 Western Blot
2.11 Co-Immunoprecipitation
2.12 In-vitro Splicing Reactions
2.13 Native Gel Analysis
2.14 Denaturing Gel Analysis
2.15 Mass Spectrometry Prep

Chapter 3: Results
3.1 Establishing RMCE pipeline for V5-snRPB2
3.2 Validation of donor cassette integration and mRNA expression
3.3 Validation of tagged protein expression
3.4 Immuno-purification of tagged protein from HeLa extract
3.5 Mass spectrometry of immuno-purified sample
3.6 Validation of assembly and catalytic activity in tagged protein nuclear extract

Chapter 4: Discussion
4.1 Results Overview
4.2 Protocol Optimizations
4.3 Future Directions

Appendix
References
List of Figures

1.1 Splicing chemistry and representation of conserved intron sequences in yeast and humans
1.2 U2 snRNP components interact with the branch point sequence
1.3 RMCE donor plasmid acceptor locus construct in HeLa
3.1 RMCE pipeline for tagged snRPB2 expression and protein purification from HeLa
3.2 PCR and RT-PCR validation of donor cassette integration and trans-gene mRNA expression
3.3 Presence of V5- tagged protein in whole cell extracts of induced HeLa
3.4 Purification of tagged core spliceosome proteins from induced HeLa extracts
3.5 Mass Spectrometry results of V5-snRPB2 nuclear extract Co-IP indicate presence of U2snRNP components
3.6 Assembly and catalytic activity of spliceosome complexes on full length substrate with mini-prepped nuclear extracts
List of Tables

2.1 Primer sequences for mutagenesis and two-step PCR

2.2 Primer sequences for Sequencing, Genomic PCR, and RT-PCR
Abstract

A PIPELINE FOR TAGGING snRNP ASSOCIATED PROTEINS IN HeLa

by

John Kim

The spliceosome is a dynamic eukaryotic macromolecular complex responsible for catalysis of pre-mRNA splicing. High quality processing of a pre-mRNA by removal of non-coding introns and ligation of exons is essential to preserve the accuracy of a mature RNA transcript that will undergo eventual translation into a functional protein. Given the irresolute nature of the spliceosome, much of the machinery underlying the structural rearrangements that occur throughout the splice cycle have yet to be fully characterized. Thus far, biochemical characteristics of spliceosome assembly in humans have been studied by complexes purified by in-vitro reconstitution with HeLa nuclear extract and a transcribed radiolabeled pre-mRNA substrate. While in-vitro reconstitution has been a standard for purification of both human and yeast spliceosomes, studies in yeast have also been able to utilize endogenous protein tags to study the physical nature of spliceosome intra-interactions. Here we have re-purposed a RMCE protocol from Khandelia et al. to endogenously express V5-tagged snRPB2, a core component of the U2 snRNP, in HeLa cells. Our tagged construct was integrated into the HeLa cell genome by Cre-Lox recombination and construct expression is regulated by a TET-ON system. Selected HeLa cells were then induced to express V5-tagged snRPB2 for eventual purification of the tagged
protein and its endogenous interactions by co-immunoprecipitation (Co-IP). Mass spectrometry data revealed the Co-IP eluent contains all the components of an endogenously assembled U2snRNP from active HeLa nuclear extract. Altogether, these preliminary results open the door for a pipeline to study biochemical characteristics of mammalian spliceosomal complexes purified via endogenously tagged proteins.
Dedication

To my mom and dad who have given up so much to raise me and my sister. To my sister, for her unconditional support and love. To my friends, who mean everything to me. To Rankin and Segre, my two homes away from home. To Anna, my love. To Roux, my dog.
Acknowledgements

First, I would like to acknowledge my faculty advisor Melissa. She has been everything I could ever hope for in a mentor. I’d like to thank her for her constant encouragement and support throughout my graduate career. Her efforts to be understanding, aware, and honest are qualities I will strive to uphold moving forward in my career and life.

Tonio Schutze, a former postdoc in the lab, who gave me my first start. His guidance when I first began were absolutely crucial for my development. It was his idea to utilize this system and tasked me with the responsibility to complete this project. I would also like to thank Piyush Khandelia and the Makeyev group for providing groundwork for this project.

Hannah Maul-Newby, a current member of our lab. Hannah had a critical role in this project. Not only was she generous enough to provide her own Western blot and Co-IP protocols, but also took a personal interest in this project. Hannah has been a great mentor and friend over these last couple years.

Beth Prichard, our lab manager, who taught me how to handle cells and tissue culture. Beth was patient and generous with me when I transitioned into the lab. She has been irreplaceable for the upkeep and organization of our lab.

I would like to thank Jeremy Sanford and Manny Ares, my committee members, for their time and input towards this project. And finally, my lab mates (current and past), the Sanford lab, the fourth floor of Sinsheimer, and the UCSF Mass Spectrometry Facility.
Chapter 1

Introduction

1.1 Eukaryotic Pre-mRNA Splicing

Nascent pre-mRNA transcripts in eukaryotes consist of both protein coding (exons) and non-coding (introns) sequence. The removal of intronic stretches of sequence interspersed within the transcript and subsequent ligation of exons is a process known as pre-mRNA splicing. The resulting mature mRNA goes on to be translated into a functional polypeptide by the ribosome. Splicing at the proper exon-intron junctions of a pre-mRNA substrate is key in order to maintain the correct sequence for translation. Defects in splicing and its corresponding machinery have the potential to alter the amino acid sequence of a polypeptide thus altering a protein’s structure and function in the cell.

The spliceosome is a macromolecular RNA-protein complex responsible for catalysis of splicing on pre-mRNA substrates. Spliceosomes consist of uridine rich small nuclear ribonucleoprotein particles (snRNPs) and non-snRNP auxiliary factors that assemble \textit{de novo} on pre-mRNA substrates (Staley and Woolford Jr. 2009). The cycle in which \textit{de novo} assembly and subsequent rearrangement/removal of core and auxiliary factors, that result in the various complexes of the spliceosome, is known as the splice cycle.
1.2 Splicing Chemistry and Intronic Elements

Introns are excised from pre-mRNA via the splice cycle in two transesterification steps: branching and exon ligation. Branching occurs when the phosphorus of the 5’ exon-intron junction (5’ splice site) undergoes a nucleophilic attack by the 2’-hydroxyl group of an adenosine within the intron (branch point adenosine). The result produces a free upstream exon and a branched intron (lariat intron) that remains bound to the 3’ exon. During second step chemistry, the last nucleotide of the free 5’ exon performs a nucleophilic attack at the 3’ splice site, effectively releasing the lariat intron and ligating the exons together (Figure 1.1A).

![Figure 1.1: Schematic diagram of splicing chemistry and representation of conserved intron sequences in yeast and humans.](image)

(A) Introns are removed via two transesterification reactions in succession. The 5’ splice site, 3’ splice site and branch point adenosine are denoted as 5’SS, 3’SS and BrA, respectively. Red arrows denote nucleophilic attacks. (B) Intron motifs implicated in recognition by the spliceosome are highly conserved in yeast compared to that of humans. Y denotes any pyrimidine, N denotes any nucleotide, R denotes any purine. Adapted from Nagai et al. 2018.
Within introns lie conserved sequences that function to help define intron boundaries for the spliceosome: the 5’ splice site, the 3’ splice site, the branch point sequence, and the polypyrimidine tract. Both 5’ and 3’ splice sites, GU and AG respectively, are invariant dinucleotides that delineate the start and end of the intron. Although these dinucleotides are highly conserved across species, the consensus sequences surrounding them can vary across and even within a species (Figure 1.1B).

The branch point sequence (BPS) is a stretch of seven nucleotides, upstream of the 3’ splice site, that houses the branch point adenosine. In yeast, the BPS is highly conserved, UACUAAC (Newman et al. 1985) while in humans the sequence is mostly degenerate; YNYURAY (Ruskin et al. 1985) (Figure 1.1B).

In higher eukaryotes, the polypyrimidine tract (PYT) is a stretch of 15-20 pyrimidines adjacent to the 3’ splice site. The PYT that has been shown to recruit splicing factors, such as U2AF, required for binding of the U2snRNP (Ruskin et al. 1988) (Figure 1.1B).

1.3 The Splice Cycle

The splice cycle is divided into multiple stages: assembly, catalysis, and disassembly. In summary, initiation of assembly begins when the 5’ splice site is recognized and bound by the U1 snRNP along with binding of non-snRNP factors SF1 and U2AF to the BPS and PPT, respectively (Liu et al. 2001; Ruskin et al. 1988).
The commitment complex, E complex, (Reed et al. 1990) transitions to A complex when the U2 snRNP binds the BPS via base pair interactions with the U2 snRNA (Query et al. 1994). Prior to catalysis, a pre-assembled U4/U6.U5 tri-snRNP associates with A complex to form pre-B complex, a fully assembled spliceosome (Stevens and Abelson 1999; Boseler et al. 2016). Transfer of the 5’ splice site from U1 to U6 snRNA via Prp28, followed by unwinding of the U4/U6 snRNA duplex by Brr2, results in dissociation of U1, U4/U6 proteins, and the U4 snRNA (Staley and Guthrie 1999; Raghunathan and Guthrie 1998; Laggerbauer et al. 1998). This allows for the formation of the catalytic core of the spliceosome by base pair interactions of U2 and U6 snRNA resulting in B$^{\text{act}}$ complex which in turn is converted to B* by a DEAH-box helicase, Prp2 (Madhani and Guthrie 1992; Warkocki et al. 2009). This B* complex catalyzes the first transesterification step of splicing (nucleophilic attack of the 5’ splice site by the BrA) followed by C and C* complexes which catalyze the second step (nucleophilic attack of the 3’ exon by the 5’ exon). Finally, exon ligation generates the mRNA, in P complex, which is released by the helicase Prp22 and snRNPs are recycled for subsequent rounds of splicing (Company, Arenas and Abelson 1991; Schwer 2008).

1.4 Current Spliceosome Structures and Purification Methods

Three-dimensional modeling of the spliceosome via X-ray crystallography and cryo-electron microscopy have provided much insight behind the potential mechanics that govern the various conformations of the spliceosome. Advances in
cryo-EM technology have led to high resolution cryo-EM structures of the spliceosome. In 2015, the intron lariat spliceosome (ILS) complex in *S. pombe* was solved at a resolution of 3.6 Å (Yan et al. 2015). Furthermore, over the past three years (2015-2018), the Shi, Nagai, and Luhrmann labs have provided 17 additional cryo-EM structures of both yeast and human spliceosomes ranging from 3.3 to 9.9 Å resolutions.

Of the 17 structures, the human structures include that of B, B\textsuperscript{act}, C, C*, and the U4/U6.U5 tri-snRNP (reviewed by Shi 2019). Human spliceosomes thus far have been purified by *in-vitro* assembly on a synthetic pre-mRNA substrate with an MS2-tag (Jurica et al 2002; Zhou et al. 2002) compared to direct purification by affinity tags of endogenously assembled yeast structures (Ohi et al. 2007; Yan et al. 2015). Compared to yeast, human spliceosomes share many attributes such as active site conformation, structural organization, and intra-interactions of RNA and proteins. However, additional proteins are present in human structures that are not found in yeast (reviewed by Shi 2019). Furthermore, degenerate sequence motifs along with multiple introns per transcript present themselves as challenges to direct purification stable human spliceosome structures.

### 1.5 Spliceosome Assembly and Substrate Recognition

Higher eukaryotic transcripts possess multiple introns (Michael and Manyuan 1999) compared to a single intron per transcript ratio widely found in yeast; therefore, higher eukaryotic cells have a selection of branch and splice sites to choose from. However, the structural mechanics that define a branch point
selection by the U2 snRNP is not well understood. Commitments to the splicing pathway and to specific splice sites have been shown to be separate steps in the assembly pathway implying U1 and U2 snRNP binding are not the final factors behind splice site choice (Lim and Hertel 2004).

U2 snRNP is comprised of core components such as the U2 snRNA, proteins SNRPA and SNRPB2, along with SF3b and SF3a splicing factors that assemble on to the aforementioned core components (Will et al. 2002) (Figure 1.2). Proper recognition of the BPS by U2 has been shown to be crucial for maintenance of splicing fidelity thus avoiding deleterious consequences of mis-splicing. SF3B1, a subunit of the SF3b splicing factor found in 17S U2 snRNPs, has been shown to be implicated in myelodysplastic syndromes when mutated; resulting in aberrant splicing through usage of alternate branch point sequences and 3’ splice sites (Darman et al. 2015). Wild-type SF3B1 interacts with U2AF2 at the N-terminus while the C-terminus of SF3B1 interacts with SF3B6, another U2 snRNP protein; together SF3B1 and SF3B6 have been shown to contact nucleotides flanking the BPS potentially stabilizing the U2/BS interaction (Gozani et al. 1998). Despite crystal structure data of recombinant human SF3b complexes (Cretu et al. 2016), the structural mechanism for SF3b subunit roles in BPS recognition and the purpose of the SF3b complex as a whole is unclear. Thus, a need to clarify the structural basis underlying the U2 snRNP’s decision making.
Figure 1.2: U2 snRNP components interact with the branch point sequence. Interaction of SF3B1 and the BPS implicates a role in branch point selection. The U2 snRNP is comprised of the U2 snRNA and 20 known proteins. Adapted from Wan and Wu 2013.

1.6 Recombination Mediated Cassette Exchange

Site specific recombination has been a key tool for molecular biologists to modify the genetic code of an organism by recombination of DNA. Recombinases such as Cre, rearrange two pieces of DNA by recognition and cleavage of specific sites that flank the segments of interest. Recombination mediated cassette exchange (RMCE) involves targeted integration of genes or whole gene cassettes, at a pre-defined locus containing recombination sites, that will allow for endogenous expression of a transgene. Donor gene cassettes, flanked by corresponding recombination sites, may hold a variety of genetic elements that contain not only a gene of interest, but also and not limited to: genes that confer antibiotic resistance for selection, over-expression or gene knockdown/silencing constructs, and inducible systems that regulate transgene expression.

Considering that recombination results in assimilation of the donor DNA into the genome, it follows that integrated genes face challenges of stable expression and susceptibility to cis-elements in its chromosomal surroundings. To
circumvent this issue, RMCE protocols have established acceptor locus screens for favorable loci. FLP293 is a HEK293 cell line developed to produce retroviruses via Flippase RMCE in which clones with high and stable virus production were screened for and selected (Schucht et al. 2006).

Khandelia et al. 2011 developed a similar protocol which entailed high-efficiency and low-background RMCE (HILO-RMCE) by Cre recombination, resulting in inducible shRNA expression in various human and mouse cell lines. The authors established multiple acceptor cell lines by lentiviral transduction of a cassette containing a strong EF-1α promoter and a blastocidin resistance gene flanked by mutually incompatible Cre-specific recombination sites, Lox 2272 and LoxP. These cell lines were then co-transfected with a Cre-encoding plasmid and a donor plasmid/cassette, pRD-RIPE, containing a variety of genetic elements for puromycin selection, an inducible shRNA expression system, and an EGFP expression marker, “floxed” by Lox 2272 and LoxP sites (Figure 1.3).

Transfected acceptor cell lines were then screened for stable integration and optimal expression of the trans-gene construct by FACS and RT-qPCR
(Khandelia et al. 2011).

**Figure 1.3: RMCE donor plasmid acceptor locus construct in HeLa.**

In the presence of Cre recombinase, the transfected donor plasmid exchanges gene cassettes at the acceptor locus flanked by corresponding Lox 2272 and LoxP sites. Cassette exchange results in removal of the blastocidin resistance gene followed by integration of puromycin resistance and an EGFP trans-gene regulated by a tetracycline induction module. Adapted from Khandelia et al. 2011
Chapter 2

Materials and Methods

2.1 Site Directed Mutagenesis PCR of Donor Plasmid

To generate a pRD-RIPE donor plasmid with a 3’ NgoMIV site, pRD-RIPE (Khandelia et al. 2011) was mutated following the NEB “Q5 site directed mutagenesis” PCR protocol with the following primers NgoMIV Forward and NgoMIV Reverse. Primer sequences are as indicated in Table 2.1.

PCR products were then purified using the Macherey-Nagel Gel and PCR purification kit. 2 µl of purified PCR products were then used for a KLD reaction in 10X DNA Ligase Buffer, T4 DNA Ligase, T4 PNK and Dpn1 incubated at 16˚C overnight. 10 µl of KLD reaction was then added to 100 µl of DH5-α competent cells and incubated on ice for 30 minutes. KLD competent cells were heat shocked for 45 seconds in a 42˚C water bath and 900 µl of SOC media was added to the cells immediately. Cells were incubated at 37˚C shaking for 1 hour and 100 µl were plated onto LB ampicillin media plates. Plates were incubated at 37˚C for 16 hours and colonies were counted, picked, and incubated into overnight bacterial cultures containing LB ampicillin media for 16 hours. Plasmid DNA was isolated from overnight bacterial cultures with the Thermo Scientific DNA plasmid mini prep kit. 5 µg of plasmid DNA were digested with AgeI and NgoMIV for 2 hours at 37˚C and ran out on a 1% agarose gel. The band with the
correct size was excised and gel purified using the Macherey-Nagel Gel and PCR purification kit. Purified and digested pRD-RIPE was used for the following ligations.

**PCR parameters:** 98°C for 30s, 30 cycles of 98°C for 10s, 68°C for 20s, and 72°C for 4 minutes, and a final extension of 72°C for 3 minutes.

2.2 *Two-step PCR amplification, Restriction Digest, and purification of V5-tagged SNRPB2 gene insert*

V5-tagged SNRPB2 insert was generated following the NEB Phusion PCR protocol using a snRPB2 plasmid template and the following primers for PCR 1: SNRPB2 V5 Forward and SNRPB2 all Reverse. 1 µl of PCR 1 product was then used as template for PCR 2 following the NEB Phusion PCR protocol using the following primers: V5 Forward and ALL Ngo Reverse. Primer sequences are as indicated in Table 2.1.

PCR 2 products were PCR purified via Macherey-Nagel Gel and PCR purification kit (Ref. #740609). 20 µl of purified PCR products were then digested in 10X CutSmart Buffer with AgeI-HF (NEB) and NgoMIV (NEB) at 37°C overnight. Digests were run on a 1% agarose gel and bands of the correct size were excised and purified using the Macherey-Nagel Gel and PCR purification kit (Ref. #740609). Digested and purified SNRPB2 products were used for following ligations.

**PCR 1 parameters:** 98°C for 30s, 35 cycles of 98°C for 20s, 63°C for 20s, and 72°C for 1 minute and 30s, and a final extension of 72°C for 7 minutes
**PCR 2 parameters:** 98˚C for 30s, 35 cycles of 98˚C for 20s, 65˚C for 20s, and 72˚C for 1 minute and 30s, and a final extension of 72˚C for 7 minutes

**Table 2.1: Primer sequences for mutagenesis and two-step PCR.**

<table>
<thead>
<tr>
<th>Name</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>NgoMIV</td>
<td>5' ATCAGAAGGTGCCGGCTGGGTGTGGCC 3'</td>
</tr>
<tr>
<td>Forward</td>
<td></td>
</tr>
<tr>
<td>NgoMIV</td>
<td>5' AGGCAAGCTGCACCTGAG 3'</td>
</tr>
<tr>
<td>Reverse</td>
<td></td>
</tr>
<tr>
<td>SNRBP2</td>
<td>5' CCTCGGTCTCAGATTCTACGGGATCCATGGATATCATCACAGACAAATCATAC 3'</td>
</tr>
<tr>
<td>V5</td>
<td></td>
</tr>
<tr>
<td>Forward</td>
<td>5' CAGCCTGCACCTGAGGTGTCTATTTCTTTGCGCATAGGTGATC 3'</td>
</tr>
<tr>
<td>SNRBP2</td>
<td></td>
</tr>
<tr>
<td>all Reverse</td>
<td>5' CAGCCCTGCACCTGAGGTGTCTATTTCTTTGCGCATAGGTGATC 3'</td>
</tr>
<tr>
<td>V5</td>
<td></td>
</tr>
<tr>
<td>Forward</td>
<td>5' CGCAACTGGTCGGACACCAGGTAAGCTGCTATCCCTAAACCTTCTCGGTCTCGCGATTCATACAGG 3'</td>
</tr>
<tr>
<td>ALL Ngo</td>
<td></td>
</tr>
<tr>
<td>Reverse</td>
<td>5' CGCGCCCGTGACCTTCTGTAGGCAGCTGCACCTGAGTG 3'</td>
</tr>
</tbody>
</table>

### 2.3 Ligations and Transformation

Ligations were carried out at a 5:1 insert to vector molar ratio. 150 ng of digested pRD-RIPE and 93 ng of digested V5 SNRBP2 insert in a ligation reaction containing 10mM ATP, Ligase, 10X Ligase Buffer, and mQH20. Ligations were incubated at room temperature for 1 hour and the total volume added to DH5-alpha competent cells. Competent cells with the ligations were incubated on ice for 25 minutes and heat shocked in a 42˚C water bath for precisely 45 seconds. 900 µl of SOC media was added immediately after and samples were incubated in 37˚C shaking for 1 hour. Competent cells were then plated onto LB media plates containing Ampicillin and left overnight in the 37˚C
incubator. The plates were removed the following morning and colonies picked and incubated in overnight cultures containing LB media and Ampicillin. Plasmid DNA was purified the following morning using the Thermo Scientific DNA plasmid mini prep kit and plasmids with pRD-ORF primers were sent for sequencing to UC Berkeley Sequencing to validate ligation of the insert into vector.

2.4 **Cell Culture**

Adherent HeLa cells containing an established acceptor locus site flanked or “floxed” by *Lox2272* and *LoxP* sites were provided by (Khandelia). HeLa cells were cultured in high glucose DMEM medium supplemented with 10% FBS and passaged using 1X trypsin (HyClone). After transfection, HeLa cells were subjected to puromycin selection at concentrations ranging from 0.75 µg/mL-1.5 µg/mL. Tet-inducible expression was turned on by addition of Doxycycline to a final concentration of 2 µg/mL. All cells were stored in 90% DMEM and 10% DMSO and placed in -80°C until frozen and eventually placed in liquid nitrogen for long term storage.

2.5 **Stable Cell Line Construction**

Stable cell lines expressing V5-tagged proteins was established in adherent HeLa cells via co-transfection of a donor plasmid pRD-RIPE mixed with a Cre-encoding plasmid, pCAGGS-Cre, (Khandelia) with the latter comprising 1% of the total plasmid concentration. A single well from a 12-well plate was
transfected with according to the Polyplus jetPRIME® in vitro DNA transfection protocol. Transfected cells were left overnight and the media changed the following day and incubated for 24 hours. Cells were detached with trypsin, expanded into 6-well plates the following day, and allowed to re-adhere for an additional 24 hours. Adherent cells were subjected to puromycin selection the following day at a half maximal concentration of 0.75 µg/mL and incubated for 24-48 hours with new puromycin media every 24 hours. Following selection, puromycin resistant cells were maintained at a maximal concentration of 1.5 µg/mL.

2.6 **Genomic DNA Extraction and PCR**

Genomic DNA was extracted from ~1.3 x 10⁷ transfected adherent HeLa using the “Zymogen ZR Genomic DNA Tissue Mini-Prep” kit. Genomic preps were quantified using a nanodrop spectrophotometer. ~100ng of genomic DNA was used as template for PCR with Taq polymerase using pRD-ORF primers: pRD-RIPE Forward and pRD-RIPE Reverse. Primer sequences are as indicated in Table 2.2. PCR products were run on a 1% agarose gel (1:17,000 EtBr).

2.7 **RNA Extraction and RT-PCR**

RNA was extracted using the “Qiagen RNeasy Mini” kit (syringe lysis) and quantified with a nanodrop spectrophotometer. 5 µg of total RNA was DNaseI treated following the NEB DNaseI protocol with the following deviations. 20 µl reactions and 0.5 µl of 0.5M EDTA instead of 100 µl reactions and 1 µl of 0.5M EDTA. 1 µg of DNaseI treated RNA was used as template for
RT-PCR in a reaction with 0.5µg of pRD-ORF Reverse primer (Table 2.2),
10mM dNTPs, 0.5 µl of RNasin, 1 µl of M-MLV Reverse Transcriptase, 5X M-MLV RT Buffer, and mQH2O for a total of 20 µl. RT reactions were incubated in a 42°C water bath for 1 hour and heat inactivated at 95°C for 5 mins. 1 µl of RT reaction was used as template for a Taq Polymerase PCR following the “NEB Taq Polymerase” protocol with pRD-ORF Forward and pRD-ORF Reverse primers. Primer sequences are as indicated in Table 2.2.

Table 2.2: Primer sequences for Sequencing, Genomic PCR, and RT-PCR.

<table>
<thead>
<tr>
<th>Name</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>pRD-RIPE Forward</td>
<td>5’ AAGTTGGTCGTCAGGACTG 3’</td>
</tr>
<tr>
<td>pRD-RIPE Reverse</td>
<td>5’ CAGAAGTCAGATGCTCAAGGG 3’</td>
</tr>
</tbody>
</table>

2.8 Whole Cell Extraction

Approximately 1.3 x 10^7 adherent HeLa cells induced with doxycycline were pelleted and washed with Phosphate Buffer Saline (PBS). Cells were resuspended in RIPA Lysis Buffer containing EDTA and homogenized with a 25G needle. The insoluble material was then pelleted at 14,000 x g and the supernatant collected and centrifuged once more to remove nucleic acid and debris for 15 minutes at 4°C. The supernatant was then collected as whole cell lysate.
2.9 *Mini Nuclear Extraction*

Approximately 1.3 x 10^7 adherent HeLa cells induced with doxycycline were pelleted and washed with ice cold PBS. The packed cell volume (PCV) was noted and the cells were resuspended in 1 volume of Buffer A (10mM Hepes pH 7.9, 1.5mM MgCl\(_2\), 10mM KCl, 1mM DTT) and allow to swell on ice for 15 minutes. A 25G hypodermic needle was used to lyse/homogenize the cell suspension and the cell homogenate was centrifuged for 20 seconds in a microcentrifuge at 14,000 g at 4°C. The supernatant was discarded and the nuclear pellet was resuspended in 2/3 PCV of buffer C (20mM Hepes pH 7.9, 1.5mM MgCl\(_2\), 0.2mM EDTA, 25% glycerol, 0.42M KCl, 0.5mM PMSF, and 1mM DTT) to be incubated on ice stirring for 30 minutes. Resuspended solution was centrifuged at 14,000 x g to pellet the nuclear debris and the supernatant collected. Supernatant was dialyzed against buffer D (20mM Hepes pH 7.9, 0.1M KCl, 0.2mM EDTA, 20% glycerol, 1mM DTT, and 0.5mM PMSF) for 3 hours at 4°C and collected as nuclear extract.

2.10 *Western Blot*

Samples were loaded onto a 10% polyacrylamide gel and ran in SDS-PAGE buffer at 75V for 90 minutes. Gels were transferred onto nitrocellulose membranes in Towbin transfer buffer with SDS at 125 mA for 90 mins. Membranes were blocked in 1X TBST containing 1% nonfat dry milk for 1 hour at room temperature and primary antibody added straight to blocking buffer for
overnight incubation at 4°C rotating. The following day, blots were washed with 1X TBST three times for 5 minutes each and placed in milk. Secondary antibody Licor IRDye 680RD Goat anti-Rabbit (#926-68071) was added directly to milk at a ratio of 1:15,000 and incubated for 1 hour shaking at room temperature. Blots were washed with 1X TBST three times for 5 minutes each following secondary incubation and imaged using the Licor software image scanner.

2.11 Co-Immunoprecipitation

Protein A magnetic beads were washed with PBS three times and equilibrated with 5X bead volume of extraction buffer (RIPA for whole cell extracts and Buffer E for nuclear extracts) in 4°C overnight. 5 µg of V5 mouse monoclonal antibody (Thermo cat. #R960-25; same volume of PBS added for mock treatments) was added to extracts and left in 4°C rotating for precisely 13 hours. The following morning, beads were immobilized on a magnetic rack and the buffer taken off and discarded. Incubated extract was added to equilibrated beads and left rotating end over end in 4°C for 4 hours. Beads were immobilized and the flow-through taken off and saved. Beads were resuspended/washed with IP Wash Buffer (10mM Tris, pH 7.5, 1mM EGTA, 150mM KCl, and 1% NP-40) 3 times with each wash taken off immobilized beads and saved. V5 co-IP was eluted by resuspension/elution with 10 µl of 0.1M Glycine pH 2.5 twice for a total of 20 µl and quenched with 20 µl of 1M Tris pH 7.8. Finally, beads were incubated at 95°C for 3 minutes in 5X SDS buffer for a “hard” elution.
2.12 **In-vitro Splicing Reactions**

$^{32}$P-radiolabeled pre-mRNA substrate (MJ50) created from T7-run off transcription was used for *in-vitro* splicing reactions carried out with the following components and final concentrations. 10nM MJ50, 40mM Potassium Glutamate, 4mM Magnesium Acetate, 2.3mM ATP, 3mM Creatine Phosphate, 0.05 mg/ml tRNA, and 50% v/v HeLa nuclear extract were incubated at 30˚C.

2.13 **Native Gel Analysis**

Splicing reactions described as above were loaded on a 1.9% agarose gel made with low melting point agarose and 20mM Tris-Glycine. Native gels ran at 72 V for 3 hours and 30 minutes, subsequently dried onto Whatman paper at 65˚C for 1 hour, and exposed to phosphor screens. Screens were scanned and imaged on a Typhoon (Molecular Dynamics) scanner.

2.14 **Denaturing Gel Analysis**

RNA was gel-extracted from a 5% w/v denaturing polyacrylamide gel and *in-vitro* spliced. Splicing reactions were run on a 15% w/v denaturing polyacrylamide gel. Radiolabeled RNA was visualized by exposure to phosphor imaging screens. Screens were scanned and imaged on a Typhoon (Molecular Dynamics) scanner.

2.15 **Mass Spectrometry Prep**

Co-IP’d nuclear extract was run on a 10% polyacrylamide gel (cast with plates previously washed in phosphoric acid) for 30 minutes at 100V to run the
sample about an inch into the resolving gel. The gel was then washed briefly in mQH2O to remove residual SDS buffer and incubated in Coomassie G (40 g ammonium sulfate, 8 ml phosphoric acid 85% w/v, 0.4 g Coomassie brilliant blue G, mQH2O up to 400 mL and 100 mL methanol) shaking at room temperature overnight. The following day, the gel was washed in mQH2O to remove residual Coomassie G. The gel slice containing the sample was then excised and placed in a clean Eppendorf tube to be sent to the Mass Spectrometry Facility at UCSF (For mass spectrometry protocol please see UCSF In-Gel Digestion Protocol).
Chapter 3

Results

3.1 Establishing RMCE pipeline for V5-snRPB2

We have adapted the RMCE strategy to express a tagged version of a core component of the U2 snRNP, snRPB2, in one of the HeLa acceptor lines from Khandelia et al. 2011. The EGFP expression marker in pRD-RIPE was replaced with V5-tagged snRPB2 via molecular cloning and co-transfected into HeLa with a plasmid encoding Cre. HeLa cells that integrated the tagged gene cassette were selected for by puromycin selection and tagged gene expression was induced via doxycycline induction for 48 hours. Figure 1.4 shows our pipeline for stable inducible expression and purification of a tagged core protein of U2 snRNP. Stable integration and expression of the gene construct were validated by genomic and RT-PCR while tagged protein expression was validated by Western blot. Furthermore, Co-IP purification of V5-tagged snRPB2 from active nuclear extract of induced HeLa and mass spectrometry of the eluent indicates presence of all 21 documented proteins of the U2 snRNP suggesting this is a valid pipeline for purification and subsequent biochemical studies of human snRNPs.
Figure 3.1: Schematic diagram of RMCE pipeline for tagged snRPB2 expression and subsequent protein purification from HeLa.

Established acceptor S3-HeLa are co-transfected and selected for by puromycin selection. Integration of the donor cassette is validated by PCR of genomic DNA with cassette specific primers. Selected cells are grown to confluency and induced via addition of doxycycline to the growth media. Tagged V5-snRPB2 expression is then validated via RT-PCR of extracted RNA and Western Blot of whole cell or nuclear extract of induced HeLa. Following validation, V5-tagged snRPB2 and its endogenous interactions are purified via Co-IP which is validated by Western blot and mass spectrometry.
3.2 Validation of donor cassette integration and mRNA expression

Figure 3.2: PCR and RT-PCR validate donor cassette integration and transgene mRNA expression, respectively.

(A.) Amplified genomic DNA extracted from transfected HeLa cells alongside a 1Kb marker in an agarose gel.

(B.) Agarose gel loaded with amplified cDNA from total RNA extracted from induced HeLa cells alongside a 1Kb marker. Each lane loaded with RT-PCR product, with the exception of STREP-tagged SF3B3, shows a distinct band in the 1Kb region. Bands in the tagged EGFP samples depict subtle differences in size which correspond to the differences in tags. Of the tagged SF3B3 and snRPB2 samples, only snRPB2 contains a band running at the expected size of 1Kb.

(C.) Schematic representation of primer (purple) location relative to the donor cassette.

In order to validate genomic integration and expression of our tagged gene construct, genomic DNA and total RNA extracted from transfected HeLa cells were amplified with primers specific to regions within the cassette and just
upstream and downstream of the trans-gene. **Figure 3.2A** depicts amplification of fragments of expected size for both V5-tagged snRPB2 and V5-tagged EGFP samples. Both lanes show a single distinct band in the 1.3Kb region with a subtle difference in size. V5 snRPB2 and V5 EGFP both run at the expected sizes of ~1,300 bp indicating genomic DNA was amplified with region specific primers (pRD-ORF Forward and Reverse). In **Figure 3.2B**, each lane that was loaded with RT-PCR product, with the exception of STREP-tagged SF3B3, shows a distinct band in the 1Kb region. Bands in the tagged EGFP samples exhibit subtle differences in size which correspond to the differences of size between tags. Of the tagged SF3B3 and snRPB2 samples, only snRPB2 contains a band running at the expected size of 1Kb.

### 3.3 Validation of tagged protein expression

![Image of immunoblot analysis]

**Figure 3.3:** Presence of V5-tagged protein in whole cell extracts of induced HeLa.

(A.) Immunoblot analysis of whole cell extract from two induced HeLa cell lines transfected with V5-tagged snRPB2 and V5-tagged EGFP, respectively. Whole cell extracts ran on a 10% SDS-PAGE gel and probed with V5 mouse monoclonal antibody (Thermo, cat. #R960-25). Numbers on the side denote protein marker sizes in kDa.
(B.) Immunoblot analysis of whole cell extract from un-transfected and transfected (induced) HeLa cells. Whole cell extracts were run on a 10% SDS-PAGE gel and probed with α-snRPB2 rabbit polyclonal antibody (Proteintech, cat. #13512-1-AP). Numbers on the side denote protein marker sizes in kDa.

Transfected and induced HeLa expressing our tagged gene constructs were whole cell extracted in order to assay for expression of V5-tagged protein. **Figure 3.3A** and **3.3B** depict immunoblots of whole cell extracts from induced HeLa cells. In **Figure 3.3A**, snRPB2 and EGFP extracts were probed for the V5 epitope tagged protein, with expected sizes of 27.5 and 28 kDa respectively, and titrated alongside a protein ladder. Lanes loaded with snRPB2 extract show detection of V5 at 28 kDa with increasing intensity corresponding to the volume loaded in each lane. Bands detected in the EGFP lanes exhibit a similar trend to the snRPB2 lanes, at 28 kDa (**Figure 3.3A**). **Figure 3.3B** compares whole cell extracts probed for snRPB2 from un-transfected and induced HeLa cells. Three bands detected in the un-transfected lane are also seen in the induced lane in the 28-33 kDa region with the exception of a fourth band only present in the induced sample indicative of the presence of a V5-tagged protein in the transfected sample.
3.4 Immuno-purification of tagged protein from HeLa extract

**Figure 3.4: Purification of tagged core spliceosome proteins from induced HeLa extracts.**

(A.) Western blot of co-immunoprecipitated sample from snRPB2 whole cell extract alongside an input sample. Whole cell extract was immunoprecipitated with V5 and blots were probed with α-snRPB2 rabbit polyclonal antibody (Proteintech, cat. #13512-1-AP). Initial “Elution” was done to prevent IgG contamination of the sample. “Mock Elution” is derived from extract + PBS (instead of extract + antibody). Hard Elution was done to elute residual proteins from the beads in SDS buffer. Unrelated sample was loaded from a separate protocol that did not affect the results. Numbers on the side denote protein marker sizes in kDa.

(B.) Western blot of co-immunoprecipitated sample from snRPB2 nuclear extract. V5 immunoprecipitated samples were ran alongside a large-scale prepared HeLa nuclear extract sample (not a true input but a positive control for the antibody) and probed with α-snRPB2 rabbit polyclonal antibody (Proteintech, cat. #13512-1-AP). Initial “Elution” was done to prevent IgG contamination of the sample. Mock Elution is derived from extract + PBS (instead of extract + antibody). Numbers on the side denote protein marker sizes in kDa.

(C.) Western blot of co-immunoprecipitated sample from nuclear extract. V5 immunoprecipitated samples were ran alongside a snRPB2 whole cell extract sample (not a true input) and probed with α-SF3B3 rabbit polyclonal antibody (Proteintech, cat. #14577-1-AP). Initial “Elution” was done to prevent IgG contamination of the sample. Mock Elution is derived from extract + PBS (instead of extract + antibody). Numbers on the side denote protein marker sizes in kDa.
Once detection of V5-tagged snRPB2 had been established, we wanted to determine if other core components of the U2 snRNP would co-immunoprecipitate with V5-tagged snRPB2. First, we immunoprecipitated snRPB2 WCE with V5 antibody and found we can indeed detect purified V5-tagged snRPB2 from WCE (Figure 3.4A). Given the whole cell extraction protocol was done under denaturing conditions (SDS in the RIPA buffer), a mini nuclear extract protocol was done under native conditions to preserve endogenous interactions with the V5-tagged snRPB2. Using snRPB2 nuclear extract, we were able to recapitulate the purification of V5-tagged snRPB2 done in Figure 3.4A (Figure 3.4B) as well as detect another core protein, SF3B3, in our snRPB2 nuclear extract Co-IP sample (Figure 3.4C).

### 3.5 Mass spectrometry of immuno-purified sample

**Table 2.3: Mass Spectrometry results of V5-snRPB2 nuclear extract Co-IP indicate presence of U2snRNP components.**

<table>
<thead>
<tr>
<th>Rank</th>
<th>Gene</th>
<th>Num Unique</th>
<th>% Cov</th>
<th>Protein MW</th>
<th>Protein Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>SNRPA1</td>
<td>11</td>
<td>42.7</td>
<td>28415.8</td>
<td>U2 small nuclear ribonucleoprotein A'</td>
</tr>
<tr>
<td>2</td>
<td>SNRPB2</td>
<td>13</td>
<td>41.3</td>
<td>25486.6</td>
<td>U2 small nuclear ribonucleoprotein B''</td>
</tr>
<tr>
<td>3</td>
<td>SNRPF</td>
<td>2</td>
<td>40.7</td>
<td>9725.3</td>
<td>Small nuclear ribonucleoprotein F</td>
</tr>
<tr>
<td>4</td>
<td>SF3B2</td>
<td>36</td>
<td>35.1</td>
<td>100228.4</td>
<td>Splicing factor 3B subunit 2</td>
</tr>
<tr>
<td>5</td>
<td>PHF5A</td>
<td>4</td>
<td>34.5</td>
<td>12405.6</td>
<td>PHD finger-like domain-containing protein 5A</td>
</tr>
<tr>
<td>6</td>
<td>SF3A1</td>
<td>29</td>
<td>33.3</td>
<td>88886.9</td>
<td>Splicing factor 3A subunit 1</td>
</tr>
<tr>
<td>7</td>
<td>SNRPE</td>
<td>4</td>
<td>29.3</td>
<td>10803.7</td>
<td>Small nuclear ribonucleoprotein E</td>
</tr>
<tr>
<td>8</td>
<td>SF3B6</td>
<td>3</td>
<td>28</td>
<td>14585</td>
<td>Splicing factor 3B subunit 6</td>
</tr>
<tr>
<td>9</td>
<td>SF3A3</td>
<td>15</td>
<td>27.5</td>
<td>58849.4</td>
<td>Splicing factor 3A subunit 3</td>
</tr>
<tr>
<td>10</td>
<td>SNRPG</td>
<td>3</td>
<td>26.3</td>
<td>8496.2</td>
<td>Small nuclear ribonucleoprotein G</td>
</tr>
<tr>
<td>11</td>
<td>SNRPD2</td>
<td>4</td>
<td>24.6</td>
<td>13527</td>
<td>Small nuclear ribonucleoprotein Sm D2</td>
</tr>
<tr>
<td>12</td>
<td>SF3B1</td>
<td>24</td>
<td>22.7</td>
<td>145831.6</td>
<td>Splicing factor 3B subunit 1</td>
</tr>
<tr>
<td>13</td>
<td>SF3B5</td>
<td>1</td>
<td>19.8</td>
<td>10135.5</td>
<td>Splicing factor 3B subunit 5</td>
</tr>
<tr>
<td>14</td>
<td>SNRPD3</td>
<td>6</td>
<td>19</td>
<td>13916.4</td>
<td>Small nuclear ribonucleoprotein Sm D3</td>
</tr>
<tr>
<td>15</td>
<td>SF3B3</td>
<td>21</td>
<td>18.6</td>
<td>135578.5</td>
<td>Splicing factor 3B subunit 3</td>
</tr>
<tr>
<td>16</td>
<td>SF3A2</td>
<td>10</td>
<td>14.2</td>
<td>49256</td>
<td>Splicing factor 3A subunit 2</td>
</tr>
<tr>
<td>17</td>
<td>SNRPD1</td>
<td>1</td>
<td>10.9</td>
<td>13281.7</td>
<td>Small nuclear ribonucleoprotein Sm D1</td>
</tr>
<tr>
<td>18</td>
<td>DDX46</td>
<td>8</td>
<td>7.9</td>
<td>117363</td>
<td>Probable ATP-dependent RNA helicase DDX46</td>
</tr>
<tr>
<td>19</td>
<td>SF3B4</td>
<td>3</td>
<td>7.8</td>
<td>44386.1</td>
<td>Splicing factor 3B subunit 4</td>
</tr>
<tr>
<td>20</td>
<td>SNRPB</td>
<td>4</td>
<td>7.1</td>
<td>24610.3</td>
<td>Small nuclear ribonucleoprotein-associated proteins B and B'</td>
</tr>
<tr>
<td>21</td>
<td>HTATSF1</td>
<td>4</td>
<td>5.4</td>
<td>85853.5</td>
<td>HIV Tat-specific factor 1</td>
</tr>
</tbody>
</table>
**Figure 3.5:** snRPB2 nuclear extract Co-IP samples was run on a PAGE gel. The gel slice containing sample was excised and sent to the UCSF Mass Spectrometry facility (Raw data in appendix). The raw data was organized into Table X by the components known to comprise the U2snRNP and ranked by % coverage.

Mass spectrometry of the snRPB2 Co-IP sample returned 183 total hits (raw data in appendix). Of the 21 U2 snRNP protein components listed in the UCSC spliceosome database (reference), all 21 were identified in the snRPB2 Co-IP sample by mass spectrometry and organized by % coverage (**Figure 3.5**).

### 3.6 Validation of assembly and catalytic activity in tagged protein nuclear extract

**Figure 3.6: Assembly and catalytic activity of spliceosome complexes on full length substrate with mini-prepped nuclear extracts.**
Native agarose gel analysis of complex assembly. 1.9% Native gel loaded with splicing reactions spliced for 0, 4, and 15-minutes with nuclear extract from HeLa grown in suspension, un-transfected adherent HeLa, and adherent HeLa transfected with V5-tagged snRPB2. Letters on the side (B, A, and H/E) denote complex identity. Nuclear extract preparation method denoted in parentheses. Schematic of pre-mRNA substrate MJ50 (bottom).

Denaturing polyacrylamide gel loaded with radio-labeled RNA from in-vitro splicing reactions. Pre-mRNA substrate MJ50 was incubated for 15, 30, and 45-minute time points with nuclear extract from HeLa grown in suspension, un-transfected adherent HeLa, and adherent HeLa transfected with V5-tagged snRPB2. Splicing products and intermediates identities are denoted on the side and are as follows (top to bottom): lariat 3’ exon intermediate, lariat, pre-mRNA, mRNA, 5’ exon, and linear intron. Nuclear extract preparation method denoted in parentheses. Bottom schematic is identical to (A.).

To assay for spliceosome assembly and activity of mini-prepped nuclear extracts, radio-labeled full-length pre-mRNA was spliced in-vitro with nuclear extracts from un-transfected suspension and adherent HeLa, along with V5 snRPB2 expressing HeLa. We find that mini-prepped nuclear extracts are indeed capable of assembly upon full length pre-mRNA substrates indicated by presence of early complexes A and B in un-transfected and B2” V5 lanes compared to that of large-scale prepared suspension nuclear extract (Figure 3.6A). Furthermore, presence of identical splicing intermediates and products across all lanes in Figure 3.6B indicate mini prepped nuclear extracts present comparable catalytic activity to large-scale prepared nuclear extract.
Chapter 4

Discussion

4.1 Results Overview

Since the rise of high resolution cryo-EM, structures of fully assembled and catalytic spliceosomes in both human and yeast have come to light (reviewed by Shi 2019). The structure of stand-alone human spliceosome subunits such as the U1 snRNP and the U4/U6.U5 tri-snRNP have also been successfully elucidated (Krummel et al. 2009; Agafonov et al. 2016). Currently, a structure-based mechanism behind branch point selection by the U2 snRNP during constitutive and alternative splicing is still an open question in the field.

In this study, we targeted and tagged a core component of the U2 snRNP with the idea that, by preservation of native conditions, purification of the tagged core component would result in co-purification of its endogenous interactions thus the entire snRNP. In order to establish proof of concept, we adapted an established RMCE protocol to express tagged spliceosomal proteins in adherent S3-HeLa for purification via in-vitro assays. Remarkably, our results suggest that we were able to preserve endogenous interactions between our tagged core protein, V5-snRPB2, and all known protein components of the U2 snRNP during the nuclear extraction and immuno-purification process prior to elution (Figure 3.4C and 3.5). These results are also in line with our findings that nuclear extract
derived from adherent HeLa expressing tagged V5-snRPB2 is capable of assembly and catalytic activity upon exposure to a full-length pre-mRNA substrate within in-vitro conditions (Figure 3.6A and 3.6B). Furthermore, we were able to validate cassette integration into the HeLa genome (Figure 3.2A) and endogenous expression of the trans-gene and tagged protein (Figures 3.2B, 3.3A, 3.3B) prior to immuno-purification.

4.2 Protocol Optimization

Given the strong Ef-1α promoter in front of the cassette, we expected overexpression of our tagged construct was likely. This is a possibility we did not assay for due to the fact validation of tagged protein expression was sufficient to move on to the purification step. Protein overexpression can be consequential to the cell by causing stoichiometric imbalances, resource overload, promiscuous interactions, and undesired pathway modulations that lead to growth defects and/or cell toxicity (Moriya 2015). However, no toxicity or growth defects were observed in expression induced HeLa during cell culturing. Despite, no obvious evidence of detrimental effects, we could confirm overexpression by qPCR using tag specific primers for a quantitative analysis of mRNA expression.

Another concern we will address is the lack of proper inputs in Figures 3.4B and 3.4C. Due to the relatively small amounts of extract from the nuclear extract mini-preps, we used positive controls for antibodies as opposed to true inputs for comparison of starting amounts of protein. Regardless, we are able to
observe presence of a single tagged protein associated band in the elution of each of our IPs (Figures 3.4B and 3.4C).

Currently, we are developing a protocol to incorporate a TEV protease cleavage site in between the V5 tag and the snRPB2 core protein. We surmise by incorporation of this TEV site, elution of snRPB2 and its endogenous interactions, without the V5 tag, can be induced by cleavage at the TEV site with TEV protease. In order to validate addition of the TEV site has no influence on tagged protein expression, we have recapitulated the results shown in Figures 3.3A and 3.3B (Appendix Figure 1). Furthermore, our data includes additional validation of doxycycline induction dependent tagged protein expression. (Appendix Figure 1).

4.3 Future Directions

A core complex of the U2 snRNP, SF3b has been of particular interest due to identification of “hot spot mutations” in HEAT domains of the N-terminal region in one of its core proteins, SF3B1, that implicate the branch selection mechanism with blood cancers (Darman et al. 2015). While the role of these mutations in alternate branch selection is unclear, structural data of SF3B1, from yeast Bact structure, overlapped with the hotspot mutations found in humans allude to a mechanism in which the mutations destabilize a conformation of the pre-mRNA bound to SF3B1 mutants (Jenkins and Kielkopf, 2017).

Pladienolide B (PB), a small molecule compound, among other compounds exhibit properties that inhibit or alter splicing. PB in particular binds
the SF3b complex and stalls an A-like complex prior to stable A complex formation (Yokoi et al. 2011 and Effenberger et al. 2014). Recently, X-ray crystallography of a recombinant SF3b complex bound to PB established the PB binding site in a crystal structure of the complex bound to drug (Cretu et al. 2018). Structural data of SF3b bound by PB reveals the binding site of the drug and the branch adenosine overlap; moreover, binding of PB locks SFB31 in an “open” conformation suggesting in the absence of PB this conformation might be “closed” (Cretu et al. 2018). Our lab will utilize the RMCE pipeline to purify U2-associated components in the presence of small molecule inhibitors in order to investigate their binding properties and potentially SF3b-PB structure in the context of the full U2 snRNP.

Future directions of this project will entail developing a pipeline in suspension HeLa containing the RMCE acceptor locus for large scale nuclear extract preparations, optimization of the co-IP protocol for a TEV protease elution and proportional large scale tagged protein purification. Large scale nuclear extract preparations, compared to mini-preps, will be necessary considering the amount of protein necessary for crystallography or microscopy. A TEV protease elution will also be optimal in order to preserve the integrity of endogenous interactions as the current protocol calls for denaturing conditions. Given we are able to move forward with these optimizations our goals would be to further characterize the biochemical profile the U2 snRNP by V5 pull downs. More specifically, we hope to uncover more information on the roles of individual
snRNP-associated components, their influence on U2 snRNA-protein conformation, and their effect on early spliceosome assembly. These studies will be done in both native and non-native assembled complexes (in the presence of various synthetic substrates). Furthermore, through utility of small molecule inhibitors such as PB, V5 pull downs could be used to identify the precise nature of the A-like complex accumulated in the presence of PB.

In summary, the RMCE pipeline offers a protocol to tag components and complexes of the spliceosome purified from HeLa. Our lab was able to utilize RMCE to express tagged snRPB2 for co-immuno purification with the remaining U2 snRNP proteins. We predict this pipeline will be a useful tool to decipher biochemical and structural properties of the U2 snRNP.
**APPENDIX**

**App. Figure 1: Presence of tagged protein in nuclear extracts of induced HeLa**

(A.) Immunoblot analysis of HeLa nuclear extract probed for α-V5. Adherent HeLa cells were transfected with plasmid DNA encoding V5 B2” or V5 TEV B2” (TEV 1 and TEV 2 denote different plasmid DNA preps) and induced. Untransfected and uninduced HeLa were maintained and prepped as a control, as well as extract from suspension HeLa for comparison. (B.) Immunoblot probing for α-snRPB2 (B2”) loaded with identical samples from (A). Addition of TEV site done by Q5 mutagenesis PCR with pRD-RIPE V5-snRPB2.
REFERENCES


