UNIVERSITY OF CALIFORNIA

SANTA CRUZ

Tagging a Spliceosome Protein and Inhibiting Splicing in Cells

A thesis submitted in partial satisfaction of the

requirements for the degree of

MASTER'S OF ARTS

in

MOLECULAR, CELLULAR AND DEVELOPMENTAL BIOLOGY

by

Sean Bredeson

June 2022

The thesis of Sean Bredeson

is approved:

_____

Professor Melissa Jurica

_____

Professor Josh Arribere

_____

Professor Euiseok Kim

_____

Dean of Graduate Studies Peter Biehl

Table of Contents

List of Figures

List of Tables

Abstract

Sean Bredeson

Tagging a Spliceosome Protein and Inhibiting Splicing in Cells

The spliceosome is an important molecular complex that converts pre-mRNA to mRNA. It removes non-coding regions of RNA known as introns and connects the exons through a two-step chemical reaction. The spliceosome is made up of five snRNPs and dozens of proteins which enter and exit the spliceosome at different points of the splicing cycle. Many of the interactions between spliceosome components remain a mystery. The goal of my research is to create a tool for the lab which will allow us to further investigate the interactions of proteins within the splicing cycle. To do this I adapted a protocol developed by Khandelia et. al. to create adherent HeLa cell lines that express tagged spliceosome proteins. These cells will allow us to perform immunoprecipitations, so we can analyze the sub-complexes that are pulled down with the tagged protein. I successfully tagged was RBM17. It was confirmed that immunoprecipitation of tagged RBM17 pulled down sub-complexes. This will allow the lab to further investigate the interactions that RBM17 makes with other spliceosome proteins. I also tested the SF3B4, but determined that HeLa cells will not express a tagged version of this protein. The other part of my research investigated the effects of splicing inhibitory drugs in-vivo. I focused on the natural product herboxidiene, which targets the SF3B complex of the spliceosome. I tested how various chemical modifications changed the

effectiveness of this drug on inhibition of cell growth. This research builds on a paper recently published by Gamboa Lopez et. al. that focused on the drug's effects on cell-free splicing. I found that different chemical modifications can modulate the strength of the drug and that some of the results contradicted the cell-free data. These studies are important as this drug is a potent inhibitor of cancer cell growth and a further understanding could one day lead to this drug's application as a cancer therapeutic.

Acknowledgements

I want to thank the Jurica lab. The completion of this program would not have been possible without their help and support. I want to thank Beth Prichard, who was the lab manager at the time that I joined the lab. She helped me get started in the lab and taught me tissue culture during the peak of covid. This allowed for a smooth transition into my graduate program. Lastly, I would like to thank Melissa Jurica. She is a terrific mentor and her guidance allowed me to develop as a scientist over the last few years.

Chapter 1

Introduction

1.1 Eukaryotic Pre-mRNA Splicing

The central dogma of biology states that DNA is transcribed into RNA and then RNA is translated into protein, but this is a simplification of the full story. In eukaryotes, RNA must be processed from pre-mRNA to mRNA before is translated, RNA processing includes adding a 5' cap, polyadenylation of the 3' end and splicing. My thesis studies focus on splicing.

Pre-mRNA contains regions known as exons and introns. Introns are non-coding and must be removed and the exons spliced together. Introns are identified by four important conserved sequences, the 5' and 3' splice sites, the branch point sequence that contains an important adenosine and the polypyrimidine tract (Wilkinson et al. 2020) (Fig 1.1). In humans, the nucleotide sequence around the 5' splice site and the branch point sequence are less stringently conserved than in yeast (Wilkinson et al. 2020).

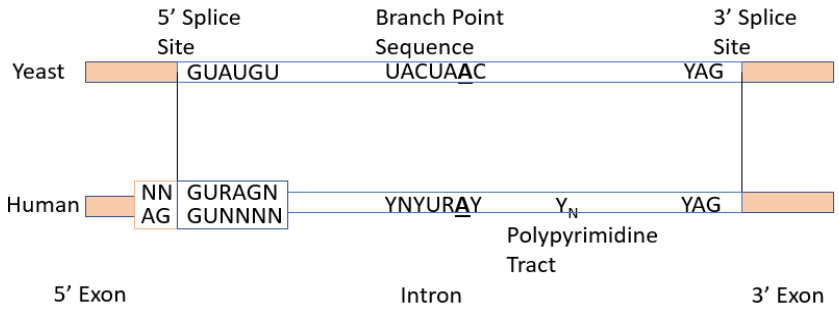

Fig. 1.1: An intron with the conserved sequences in yeast and humans. Conserved sequences in the intron that are important for spliceosome recognition. The Y stand for pyrimidines, the R stand for purine and the N stands for any base.

1.2 The Chemistry of Splicing

The process of removing the intron and splicing together the 5' and 3' exons is completed in two chemical-steps (Fig 1.2). Each step is a transesterification reaction that exchanges one phosphodiester linkage for another (Nilsen 2003). First, is intron lariat formation. After the 5' splice site and the branch point adenosine are positioned in the active site of the spliceosome, the 2' hydroxyl of the adenosine acts as a nucleophile to attacks the 5' splice junction. This breaks the phosphodiester bond and forms a new 2'5' phosphodiester linkage between the branch point adenosine and the 5' terminal nucleotide of the intron, which creates an intron lariat connected to the 3' exon intermediate (Fig 1.3). After a conformational change repositions the 3' end of the intron into the active site, the second reaction proceeds. The newly released 3' hydroxyl of the 5' exon attacks the 3' splice site to break the bond and the end of the intron and form a new one

between the 5’ and 3’ exons and intron (Nilsen 2003) (Fig 1.3). Splicing concludes with the release of the final products: spliced mRNA and an intron lariat.

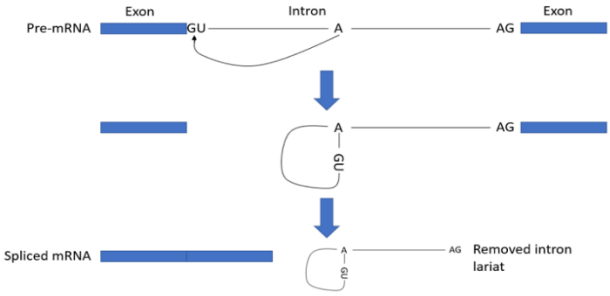

Fig. 1.2 Removal of an intron. An illustration of an intron being removed as a lariat.

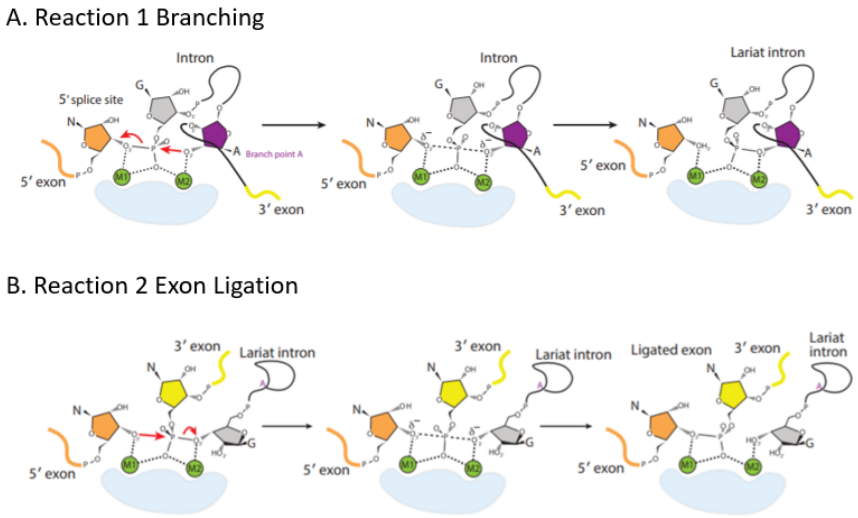

Fig. 1.3 Splicing chemistry. This illustration shows the two steps of chemistry in splicing and how two metal ions stabilize a penta-covalent transition state. (Adapted from Wilkinson et al. 2020)

1.3 Alternative Splicing

Alternative splicing is a mechanism in which some exons can be skipped or included to give multiple mRNA isoforms. Alternative splicing is important as it

allows a large increase in the number of mRNA transcripts that can be generated from a smaller number of genes. The information provided in this section is by no means a comprehensive review of alternative splicing. Instead, I felt that it was important to mention as an example of the complexity of splicing.

Molecular analysis shows that alternative splicing determines the binding properties, intracellular localization, enzymatic activity, protein stability and post-translational modifications of many proteins (Stamm et al. 2005). Alternative splicing also allows organisms to react to their environment, changing gene expression with changing conditions (Kelemen et al. 2013). DNA microarray experiments indicate that 74% of all human genes are alternatively

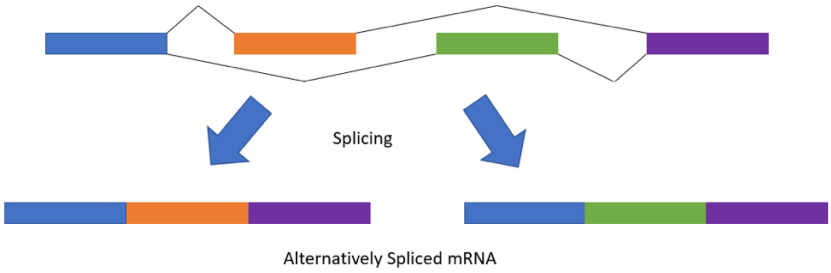

Fig. 1.4 Schematic of alternative splicing. This schematic shows the basic idea of alternative splicing and the two different spliced mRNA's that arise.

spliced (Johnson et al. 2003). On average a human gene produces two to three transcripts (Stamm et al. 2005). A major hurdle in understanding splice site selection is that the splicing regulatory sequences are degenerate. They can be described as consensus sequences that are followed with a lot of variation (Black 2003). The

current understanding is that the splicing sequences for alternatively spliced exons deviate more from the consensus sequence. These are considered weaker and implies that there would be a lower affinity to the spliceosome resulting in reduced recognition. The branch point of consecutive exons is typically within 40 nucleotides of the 3' splice site, but this distance can be much larger in alternative exons. Exon selection can also be regulated by expression of proteins such as splicing factors (Kelemen et al. 2013). There is even evidence that long non-coding RNAs can regulate exon selection (Yu et al. 2008). The big question is how do proteins in the spliceosome control alternative splicing. One of the proteins I attempt to tag is SF3B4 which is part of the spliceosome during branch point recognition.

1.4 The Splicing Cycle

Splicing is accomplished by the molecular complex known as the spliceosome, which functions in a process commonly referred to as the splicing cycle (Wilkinson et al. 2020). The spliceosome is made up of five small nuclear ribonucleoproteins (snRNPs) called U1, U2, U4, U5 and U6. The snRNPs are made up of small nuclear RNAs, which assemble with associated proteins to form the catalytic center of the spliceosome (Matera and Wang 2014). During intron removal, the snRNPs and dozens of additional proteins enter and exit the spliceosome at different points of the splicing cycle (Fig 1.5). This section will cover the interactions as the spliceosome progresses through the splicing cycle. Some of the protein names used in this

section are associated with yeast, but there are highly conserved orthologs in mammalian cells.

In the earliest stages of the splicing cycle, recognition and marking of an intron is done by the U1 and U2 snRNPs within the pre-spliceosome (Plaschka et al. 2018). The U1 snRNP binds to the 5' splice site through base pairing to form E complex (Zhang et al. 1986). In eukaryotes the pre-mRNA branch point sequence is bound by SF1/mBBP and the heterodimer U2AF65-U2AF35 which cooperatively binds to the polypyrimidine tract and the 3' splice site (Zamore et al. 1992). In the next step an A complex forms. The helicases PrP5 and Sub2 interact SF1 and U2AF. At this point the U2 snRNP is recruited to the branch point sequence (Kistler and Guthrie 2001). Then the SF3B complex binds to the 5' side of the U2 snRNA (Kramer et al. 1999). The SF3A complex bridges the SF3B4 and core domains of the U2 snRNP (Wilkinson et al. 2020). When the U2 snRNP is stably integrated into the spliceosome, its base pairing interaction with the branch point sequence forms a branch helix (Parker et al. 1987). The branch point adenosine flips out and interacts with SF3B1, setting up the process of making it available to act as a nucleophile in the first chemical reaction (Plaschka et al. 2017).

Then the next part of the splicing cycle transitions through several forms of B complex: pre-B, B, B^{act} and B^{*}. To go forward to the pre-B complex in humans the 5' splice site is released from the U1 snRNP, by Prp28 leading to events that relocate

the Brr2 helicase setting up activation of the spliceosome (Charenton et al. 2019). Unwinding of U4 and U6 snRNA and folding of U2 and U6 snRNA forms the active site in the B^{act} complex (Fica et al. 2014). There are also two catalytic metal ions that activate nucleophiles and stabilize leaving groups (Steitz and Steitz 1993). At this point the spliceosome is activated but inhibited. The 5' splice site is in the active site, but the 2' OH group of the branch point adenosine, which is the nucleophile in the branching reaction, is outside of the active site in contact with the SF3 complex (Rauhut et al. 2016). These interactions are disrupted by Prp28 as the spliceosome moves into the B^{*}complex. The branch point adenosine is now allowed to enter the active site (Wilkinson et al. 2020). With the two components of the branching reaction in the active site, the spliceosome moves into the C complex. The intron lariat is formed in C complex when the 2' OH of the branch point adenosine performs a nucleophilic attack on the 5' splice site. The spliceosome uses the same active site for the second reaction, but remodeling must occur in which there is a displacement of the branch helix. This remodeling rotates the branch helix, giving space in the active site for the 3' splice site (Fica et al. 2017). At this point the spliceosome is in the C^{*} conformation in which the exon ligation occurs. The 3' splice site is docked in the active site and PrP18 and Slu7 act as exon ligating factors (Semlow et al. 2016).

The spliceosome goes into the final P complex. PrP22 catalyzes the release of the ligated exons (Company et al. 1991). Then the spliceosome is disassembled, and the intron lariat is degraded and recycled.

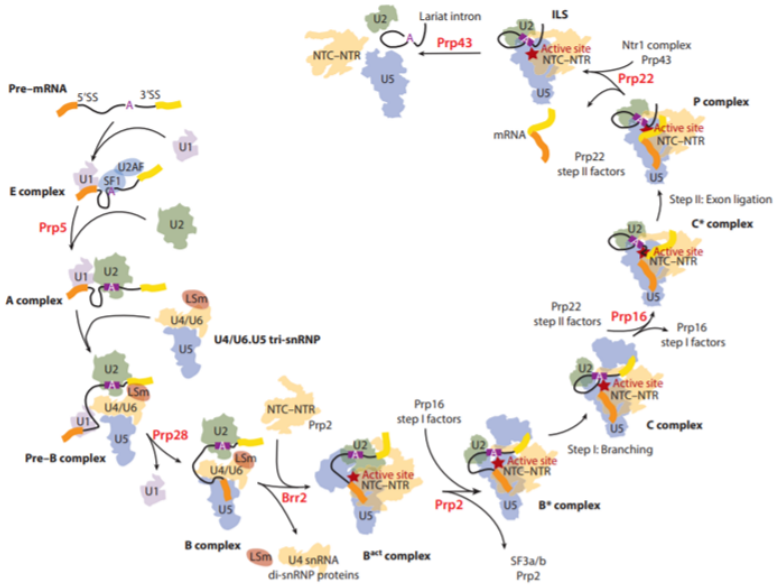

Fig. 1.5 The splicing cycle. This image is adapted from Wilkinson 2020.

1.5 A Protein of Interest: SF3B4

As was stated previously, there are dozens of proteins that enter and exit the spliceosome at different points within the splicing cycle. One of the proteins that my studies focused on is SF3B4, which is part of the SF3B complex. The SF3B complex comes in with the U2 snRNP to associate with the intron in the formation of the A complex and mediates the tethering of U2 snRNP to the branch site of pre-mRNA (Champion-Arnaud and Reed 1994). The sequence and domain architecture

of SF3B4 is well conserved within eukaryotes, indicating its importance (Igel et al. 1998). Evidence shows that SF3B4 is also important for transcription, translation and cell signaling (Xiong and Li 2020). It is unknown whether SF3B4 directly plays a role in these processes or if the effects are related to SF3B4's role in splicing.

SF3B4 is a haplo-insufficient gene and heterozygous loss of function mutation in the gene results in Nager syndrome (Petit et al. 2014). Nager syndrome is characterized as an acrofacial dysostosis, which results in a phenotype displaying craniofacial and limb malformations. There is also evidence that over expression of SF3B4 causes problems. In human cells, over expression of SF3B4 results in mis-spliced Kruppel-like factor 4 (KLF4) (Shen and Nam 2018). The aberrant splicing results in a non-functional transcript of a tumor suppressor encoding gene and promoted tumorigenesis leading to the most common type of primary liver disease: hepatocellular carcinoma (HCC).

A full characterization of SF3B4 and the complexes it forms within the splicing cycle is imperative in gaining a better understanding of the spliceosome and working towards developments of therapeutics for these conditions. I aimed to create a stable HeLa cell line that expresses affinity tagged SF3B4 protein and allow us to isolate SF3B4 associated intermediates from the splicing cycle.

1.6 Another Protein of Interest: RBM17

RBM17 is an RNA binding protein that has a role in splicing regulation (Tan et al. 2016). Also, knockout of RBM17 in mice results in early embryonic lethality and rapid degeneration in Purkinje neurons (Tan et al. 2016). RBM17-dependent splicing changes correlated with the use of cryptic splice sites of genes related to motor coordination and cell survival. Another study postulated that RBM17 plays a role in short intron recognition (Fukumura et al. 2021). These studies will help a lab-mate, Hannah Maul-Newby, in her research into spliceosome assembly. The spliceosome protein DHX15 best known for its role in disassembly of the intron lariat spliceosome at the end of splicing (Fourmann et al. 2016). It is regulated by a G-patch co-factor that direct it to the intron lariat (Wen et al. 2008). Hannah showed that DHX15 may also regulate A complex formation, and RBM17 is a candidate G-patch protein regulator (Agafonov et al. 2011). The development of a cell line that expresses tagged RBM17 protein will allow for further analysis of these protein's role in spliceosome assembly and function.

1.7 Splicing Inhibition Drug Assay

The second part of my research project is investigating the effects of splicing inhibitory drugs in cells. This research builds on a paper recently published from the Jurica lab that investigated these drugs in a cell-free splicing system. The drugs target the SF3B complex of the spliceosome. Recent hypotheses has arisen based on

cryo-EM structural models of the spliceosome, explaining how the SF3B complex may play a role in branch point sequence recognition (Plaschka et al. 2018). SF3B contains a C shaped HEAT repeat domain that can take on an open or closed conformation. Structural models suggest that the SF3B clamp closes around the branch point sequence (Cretu et al. 2016). The splicing drugs bind in a tunnel in SF3B and are predicted to stop SF3B1 from closing and interfere with the formation of the A complex which forms early in the splicing cycle. We are focusing on herboxidiene and recently identified features that differentially influence binding and inhibition (Gamboa Lopez et al. 2021) (Fig 1.6).

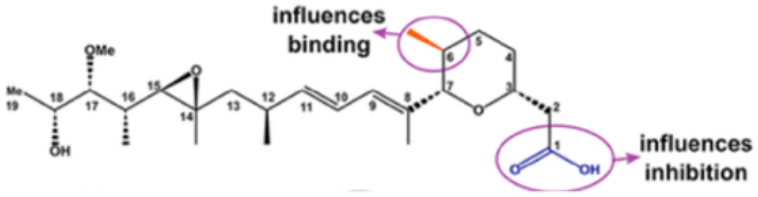

Fig. 1.6 Chemical structure of herboxidiene. This chemical structure illustrates the sites of variation between drug compounds relating to binding and inhibition. Image adapted from Gamboa Lopez 2021.

This study is important because many cancers and dysplasia exhibit mutations in the SF3B complex (Hubert et al. 2013). Also, herboxidiene compounds have been identified as potent inhibitors of tumor cell growth (Mizui et al. 2004). This study will build on our understanding of splicing inhibition and how it could be used in cancer therapeutics.

Chapter 2

2.1 Introduction: Creation of Tagged Cell Lines

My thesis project aimed to create cell lines expressed tagged spliceosome proteins over the endogenous proteins as it is difficult to modify the endogenous proteins. I tested two proteins- SF3B4 and RBM15, and two tags. The first tag is the V5-epitope, which is useful for immunoprecipitation with the V5-antibody. The second tag is the streptavidin binding peptide (SBP) which is recognized by streptavidin for protein purification.

2.2 Recombination Mediated Cassette Exchange

I adapted a protocol developed by Khandelia et.al. to engineer a cell population with an inducible locus that expresses a transgene (Khandelia et al. 2011). The technology is known as recombination mediated cassette exchange (RMCE). RMCE uses the activity of site-specific recombinases to integrate a donor sequence flanked by self-compatible but mutually incompatible recombination sites at a predefined acceptor locus containing a similar pair of recombination sites (Wirth et al. 2007)(Fig 2.1). The donor plasmid is co-transfected with a pCAGGS-Cre plasmid encoding a wild type Cre recombinase. This technology allowed me to engineer a HeLa cell line with puromycin resistance for selection and an inducible transgene locus, all of which was integrated into the genomic DNA to create a stable cell line.

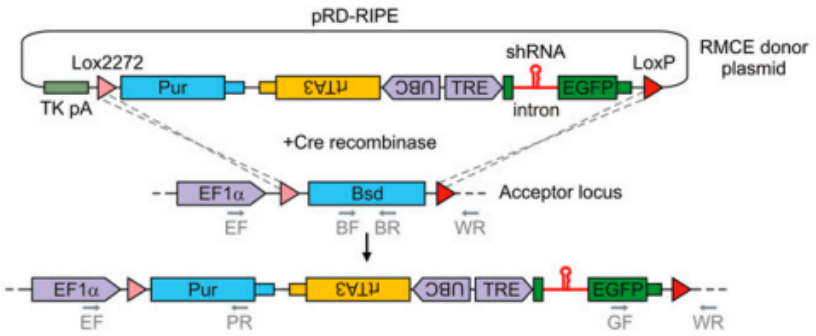

Fig. 2.1 The donor plasmid and acceptor locus. In the presence of Cre recombinase the locus flanked by Lox 2272 and LoxP is integrated into the genomic DNA. (adapted from Khandelia et. al. 2011)

Material and Methods: Creation of Tagged Cell Lines

2.3 Donor Plasmid Assembly

The first step in the assembly of the donor plasmid is PCR. I amplified the plasmid backbone from the pRD-RIPE template plasmid (E.V. Makeyev). The PCR reaction contained the pRD-RIPE plasmid, the correlating primers for the plasmid backbone (Table 2.1) and Q5 DNA polymerase (NEB). **PCR parameters:** 98°C for 30 sec., 30 cycles of 98°C for 10 sec., 65°C for 30 sec., and 72°C for 4 minutes.

I then amplified the insert containing the transgene of interest using PCR. This was also done with Q5 DNA polymerase (NEB), a template plasmid and the correlating primers with the insert (table 1). **PCR parameters:** 98°C for 30 sec., 30 cycles of 98°C for 10 sec., 58°C for 30 sec., and 72°C for 1 min. 10 sec.

I ran the PCR products on a 0.8% agarose gel at 90 volts for 1 hour. Parent plasmid was digested by adding 1µL Dpn1 (Thermo Scientific) to each PCR reaction and incubating at 37°C for 1 hour. After this I purified the reactions using the Macherey-Nagel Gel and PCR purification kit.

Assembly of the plasmid was done using the NEB HiFi DNA Assembly kit. The ratio of vector to insert in the reaction was 1:2 and 0.03-0.2 pmol of total DNA was used. To transform *E. coli*, 2µL of this reaction was added to 150µL of DH5-α competent cells and incubated on ice for 45 minutes. After heat shocked for 1 minute in a 42°C water bath, 500µL of LB was added and the cells were incubated for 1 hour at 37°C. The tubes were inverted every 20 minutes. 100µL of cells were plated on LB ampicillin media plates and incubated overnight at 37°C. Plasmid DNA from colonies on the plates was isolated using Thermo Scientific DNA Plasmid Miniprep kit. For restriction analysis, 200ng of DNA from each reaction was digested using the appropriate restriction digest enzyme, for 1 hour at 37°C. The digest reaction was then run on a 0.8% agarose gel for 1 hour at 90 volts. Samples were sent to UC Berkeley Sequencing Center.

2.4 Cell Culture

Adherent HeLa cells containing an established acceptor locus site flanked by a Lox 2272 and LoxP site was provided by E.V. Makeyev. The HeLa cells were grown

in high glucose DMEM medium supplemented with 10% FBS. They were passaged using 1x trypsin (HyClone).

2.5 Creation of Stable Cell Lines

A 6 well plate was seeded with HeLa Crelox cells at a concentration of $20x10^4$ cells per well. The 6 well plate was incubated overnight at standard conditions. The cells were transfected using a total of 2µg of DNA per well; 99% donor plasmid and 1% Cre-encoding plasmid, pCAGGS-Cre (E.V. Makeyev). The transfections were done with Polyplus jetPRIME buffer and reagent, following the manufacturer's standard protocol. The 6 well plate was made up of two wells for the new cell line, two wells for the positive control, V5_EGFP and two wells as negative controls that received buffer and reagent but no DNA. After an incubation period of 24 hours, puromycin selection was started. The media was removed from the wells and fresh media containing 3µg/mL puromycin was added. The puromycin media was changed every other day for one week. After this, cells were given media with no puromycin and allowed to grow to confluency and expanded into a t75 flask. Cell stocks were frozen in 90% DMEM/10% FBS and 10% DMSO. They were placed in -80°C overnight and then moved to liquid nitrogen for long term storage.

2.6 Western Blot

Cells were grown in a t75 flask to 60% confluency and then induced with 1x doxycycline . They were incubated at standard conditions for 24 and 48 hours. Next, they were trypsinized at 80 to 90% confluency and a 50 µL sample was taken from each cell type. To lyse the cells, 10 µL of 5X SDS-PAGE sample buffer and 5 µL of betamercaptoethanol was added to each sample. The solution was resuspended, boiled at 95°C for 5 minutes, vortexed and boiled again for 2 minutes. The samples were stored at -20°C.

The samples were loaded onto a 10% polyacrylamide gel and ran in SDS-PAGE buffer at 120V for 90 minutes. The gel was transferred onto a PVDF membrane in transfer buffer at 125mA for 90 minutes. The membrane was blocked with 1x TBST containing 1% nonfat dry milk on a shaker at room temperature for 1 hour. The primary antibody was added directly to the blocking buffer and was placed on the shaker at 4°C overnight. The next day the blot was rinsed with 1x TBST for 5 minutes on the shaker 3 times. Blocking buffer (1x TBST containing 1% dry milk) was reapplied. Corresponding Licor IR Dye secondary antibodies (red goat anti-mouse (#926-68070), red donkey anti-rabbit (#926-68073) or green donkey anti-mouse (#926-32212)) were added directly to the blocking buffer at 1:12,500 dilution and incubated on the shaker at room temperature for 1 hour. The blot was

rinsed with 1x TBST for 5 minutes on the shaker 3 times. The blot was imaged using the Licor Odyssey software image scanner.

For SBP tag blotting, Licor IR Dye 800CW Streptavidin (#925-32230) was applied as the primary antibody at a 1:3000 dilution. This was incubated overnight on the shaker at 4°C. The blot was imaged the next day, with no need for application of a secondary antibody.

2.7 Genomic DNA PCR

Cells (~500x10^4 cells) were removed from a confluent t75 flask and centrifuged at 2.5 rcf for 5 minutes. The media was removed, and the cells resuspended in 1 mL of PBS. Genomic DNA was extracted from a 100 µL aliquot of cells using a Zymo Research Genomic DNA Tissue Miniprep kit. The genomic DNA was used as a template in a PCR reaction with primers SQ1 and SQ3 (Table 1). **PCR parameters:** 98°C for 30 seconds, 35 cycles of 98°C for 10 seconds, 62°C for 30 seconds, and 72°C for 1 minute. The entire volume of the PCR reaction was run on an 0.8% agarose gel at 90V for 1 hour. The gel was visualized using the Bio Rad Gel Doc XR+ imaging system.

2.8 Reverse Transcription PCR

Cells were grown in a 6 well plate, with parent HeLa CreLox cells and the recombined cell lines of interest. The cell lines were induced with doxycycline at

~60% confluency and incubated for ~24 hrs. RNA was extracted from the cells at 80-100% confluency using Tri Reagent (Thermo Fisher). The RNA was DNase treated using RQ1 RNase-Free DNase (Promega). No stop solution or heat was used to quench the reaction. After the 30 min. incubation period, the RNA isolated by phenol:chloroform: isoamyl alcohol extraction and ethanol precipitation.

This purified RNA was used for the reverse transcription (RT) PCR. First, 800ng of RNA is mixed with 20µM appropriate reverse primer (Table 1) and 1µL of 10mM dNTPs in a 14µL reaction. The solutions were incubated at 63°C for 5 min. then put on ice for an additional 5 min. To these tubes, 1µL of DTT at 0.1M, 1µL of homemade reverse transcriptase and 4µL of 5x first strand buffer (Thermo Fisher) were added. These reactions were put in the thermocycler and run at 50°C for 50min. then 85°C for 5 min. An aliquot of the RT reaction was used as the template for PCR, with Q5 DNA polymerase (NEB along with no RT negative control, and positive controls for endogenous CCNA2 RNA, and plasmid used for RCME. **PCR parameters:** 98°C for 30 sec., 34 cycles of 98°C for 10 sec., 63°C for experimental, no RT and plasmid or 66°C positive control for 30 sec. and 72°C for 52 sec. 4µL of 6x loading dye was added to each reaction and the entire volume was loaded onto a thick 0.8% agarose gel. The gel was run at 90V for 1 hr. and then imaged.

Primer	sequence	function
G3.1	ATCGTACAGGGACATATGCATggatcccgtagaatcgagac	amplifies plasmid backbone containing V5 tag, goes with RBM17
G4.1	TTTGGCAGAACAAGTTTAAATCGATacactcaggtgcaggc	amplifies plasmid backbone containing V5 tag, goes with RBM17
G1.1	cgattctacgggatccATGCATATGTCCCTGTACGATGACCT	amplifies RBM17 gene
G2.1	ggcagcctgcacctgagtgtATCGATTTAAACTTGTTCTGCCAAAT	amplifies RBM17 gene
G1	cgattctacgggatccATGCATcatggctgccgggcc	amplifies SF3B4 gene
G2	ggcagcctgcacctgagtgtATCGATttactgagggagagggcctc	amplifies SF3B4 gene
G3	ggcccggcagccatgATGCATggatcccgtagaatcgagac	amplifies plasmid backbone containing V5 tag, goes with SF3B4
G4	gaggccctctccctcagtaaATCGATacactcaggtgcaggc	amplifies plasmid backbone containing V5 tag, goes with SF3B4
G1.3	CCCTCAGTAAATCGATACA**ATGGACGAGAAGACCACTGG**	amplifies SBP tag, goes with plasmid backbone containing SF3B4 with no tag
G2.3'	CAGCCTGCACCTGAGTTCGAA**TGGTTCACGTTGACCTTG**	amplifies SBP tag, goes with plasmid backbone containing SF3B4 with no tag
G3.3	CCAGTGGTCTTCTCGTCCATT**GTATCGATTTACTGAGGG**	amplifies plasmid backbone containing SF3B4 with no tag, goes with SBP
G4.3	CACAAGGTCAACGTGAACCATTCGAA**CTCAGGTGCAGGCTGCCTATC**	amplifies plasmid backbone containing SF3B4 with no tag, goes with SBP
SQ1	CTTACTGACATCCACTTTGC	sequences plasmid starting just upstream of V5 tag
SQ3	CAGGAGAGGAGGAAAAATCT	reverse primer sequences plasmid starting just downstream of the inserted gene
RNA_SF3B4_1	ggaaggtaagcctatcccta	forward primer for RT PCR, used for V5_SF3B4 and V5_EGFP
RNA_SF3B4_2	ttactgagggagagggc	reverse primer for RT PCR
RNA_V5EGFP_2	TTACTTGTACAGCTCGTCCA	reverse primer for RT PCR
CCNA2-F	AACTTCAGCTTGT	positive control for RT PCR
CCNA2-R	AAAGGCAGCTCCAGCAATAA	positive control for RT PCR

Table 1. Primer sequences

2.9 Small Volume Preparation of Nuclear Extract from Adherent HeLa Cells

Cells were grown to 40-60% confluency in six 150 mm dishes. I then induced 3 of the plates with doxycycline and incubated for 24-48 hrs., to bring the confluency of cells to ~80%. Next the plates were washed 2x with PBS. Then I added 3mL of PBS to each plate and harvested the cells using a rubber cell scraper. The cells transferred to 15mL falcon tubes and spun for 5 min. at 2000 rpm. Then the supernatant was discarded, and the cells were put on ice. The cells were resuspended in 1 packed cell volume (PCV) of Buffer A (Table 2) and transferred to an Eppendorf tube. The cells were then allowed to swell on ice for 15 min. Next the cells were homogenized using a narrow-gauge hypodermic needle, while avoiding introducing air bubbles. The cell homogenate was centrifuged at top speed for 20

sec. in a table-top centrifuge and the supernatant was discarded. The nuclear pellet was resuspended in 2/3 PCV of Buffer C (Table 2) and incubated on ice with stirring for 30 min. The nuclear debris was pelleted by spinning at 12,000g at 4°C for 30 min. Then the supernatant was dialyzed against Buffer D (Table 2) for 2 hrs. at 4°C. The precipitates were eliminated by spinning at 12000g at 4°C. The nuclear extract was aliquoted and froze in liquid nitrogen, then stored at -80°C.

Buffers	Hepes pH 7.9	MgCl2	KCl	EDTA	Glycerol	NaCl	H2O	DTT	PMSF
Buffer A	10mM	1.5mM	10mM					1mM	
200 mL	2mL 1M	0.3mL 1M	2mL 1M						
10 mL	0.1mL 1M	0.015mL 1M	0.1mL 1M				9.775mL	10uL 1M	
Buffer C	20mM	1.5mM		0.2mM	25%	0.42M		1mM	0.5mM
100 mL	2mL 1M	0.15mL 1M		40uL 0.5M	25mL	8.4mL 5M			
10 mL	0.2mL 1M	0.015mL 1M		4uL 0.5M	2.5mL	0.84mL 5M	6.431mL	10uL 1M	
Buffer D	20mM		0.1M	0.2mM	20%				
1000 mL	20mL 1M		100mL 1M	400uL 0.5M	200mL				
10 mL	0.2mL 1M		1mL 1M	4uL 0.5M	2mL		6.786mL	10uL 1M	

Table 2. Buffer table 1

2.10 Immunoprecipitation

Day 1

The nuclear extract was thawed on ice. (Only nuclear extract previously thawed no more than 1 time can be used). 2-5µg of antibody was mixed with 75 µL of nuclear extract and 1x PBS (same volume as antibody). This was then covered in foil and rotated at 4°C for 13.5 hrs.

Next the beads were prepared. 10µL of protein A magnetic beads were washed 3 times with 50µL of cold 1x PBS. The beads were pelleted by placing them

on a cold magnet for 2 min., then the supernatant was removed. The beads were stored in 50µL of 1x PBS at 4°C until the next day.

Day 2

The rest of the protocol was done on ice. The 1x PBS was removed from the beads and they were resuspended in nuclear extract incubated with antibody. The tubes were covered in foil and rotated at 4°C for 4 hrs. After the samples were rotated, they were placed on a cold magnet for 2 min. The supernatant was removed and saved as the flow through sample. The beads were then washed 3 times with 50µL of ice-cold wash buffer (Table 3). The beads stayed in the wash buffer for 1 min. before they were put on the magnet for 2 min. and then the wash buffer was removed. All the washes were saved in individual tubes.

Next a soft elution (IgG remains on beads) was performed by adding 10µL of 0.1M glycine pH2.5. Then the beads were placed on ice for 5 min. Next, they were placed on the magnet for 2 min. and the elution was transferred to a new tube. This was repeated and the elution's combined. An equal volume of 1M Tris pH8 was added to quench the glycine.

Then a hard elution (IgG released from beads) was performed. 16µL of diH_2O was added directly to the beads plus 4µL of 5x SDS-sample buffer. This was boiled for 5 min. at 95°C and loaded directly onto the PAGE gel. All samples were analyzed by western blot.

Wash Buffer	Soft Elution Buffer	Hard Elution Buffer
100mM Tris pH7.5	0.1M glycine pH2.5	5x SDS sample buffer
1mM EGTA		
120mM KCl		
1% NP40		

Table 3. Buffer table 2

Results: Creation of Tagged Cell Lines

2.11 Verification of Donor Plasmid Assembly

The first step in the creation of the cell lines, is the assembly of donor plasmids used in the transfections. My goal was to assemble the various plasmids shown in Figure 2.2. I would need to add inserts that contained either of the proteins of interest, SF3B4 or RBM17. Also, for SF3B4_SBP I needed to remove the V5 sequence, which was already part of the plasmid backbone and add the SBP sequence to the 3' end of the SF3B4 sequence.

Restriction digest was used to verify that assembly was successful (Fig 2.3). Fig2.3A shows the ClaI digestion of V5_SF3B4. The successful assembly of the donor plasmid is verified by a single band at ~9Kb. A positive control was used as well and showed a single band at ~2.7Kb as predicted. Fig2.3B is the ClaI digestion of V5_RBM17. The successful assembly of the donor plasmid is verified by a single band at ~8.3Kb. A positive control was used on this one as well and displayed the predicted single band. Fig2.3C depicts XmnI digestion of SF3B4_SBP. Successful assembly was verified by the appearance of 4 bands at ~4.5, ~2.6, ~0.7 and ~0.5Kb

as predicted. The last two lanes are an SF3B4 plasmid with no tag which displayed 3 bands as predicted. The donor plasmid was also verified through sequencing after the restriction digest.

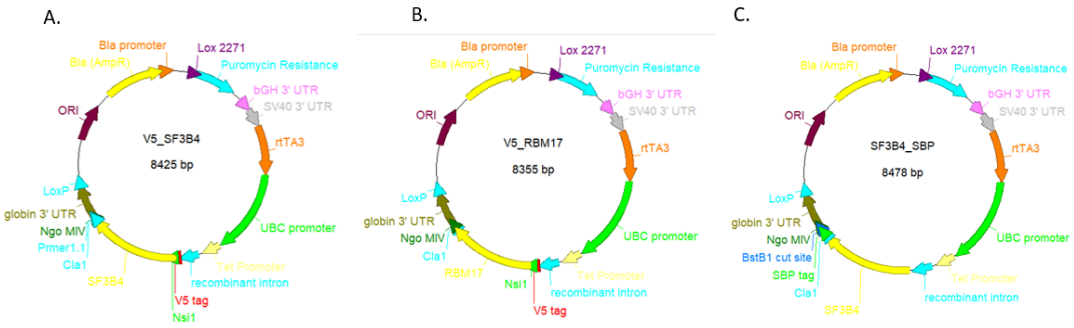

Fig. 2.2 The donor plasmids. These are the different plasmids that were co-transfected into HeLa cells to create the various cell lines.

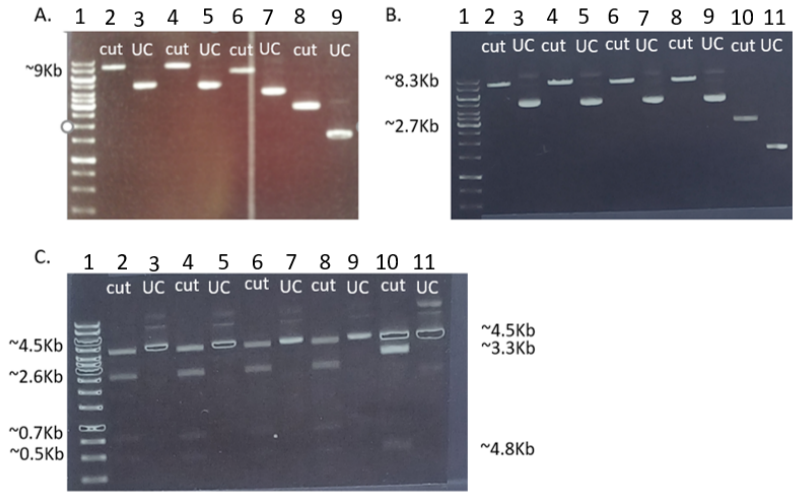

Fig. 2.3 Restriction digest to confirm assembly of donor plasmid. A. Restriction digest of V5_SF3B4 plasmid digested with Cla1. Lanes 2-7 contain V5_SF3B4 with cut and uncut (UC) lanes. Lanes 8 and 9 are a positive control of a plasmid with a

known Cla1 site. B. Restriction digest of V5_RBM17 plasmid digested with Cla1. Lanes 2-9 contain V5_RBM17, lanes 10 and 11 are a positive control. C. Restriction digest of SF3B4_SBP digested with Xmn1. Lanes 2-9 contain SF3B4_SBP and lanes 10 and 11 contain SF3B4_no tag plasmid.

2.12 Transfected Adherent HeLa Cells

Adherent HeLa cells were used in a co-transfection with either V5_SF3B4, V5_RBM17 or SF3B4_SBP plasmids and a plasmid which coded for a wild-type Cre recombinase. I assessed successful transfection and recombination by cell growth after one week of puromycin selection (Fig 2.4A). Fig 2.4B shows a positive control V5_EGFP adherent HeLa cells induced with doxycycline. The functionality of the inducible system was confirmed through the expression of EGFP. Fig 2.4C shows a negative control of un-transfected adherent HeLa Crelox cells after one week of puromycin selection. Selection was determined to be complete upon the observation of only dead cells in the negative control well.

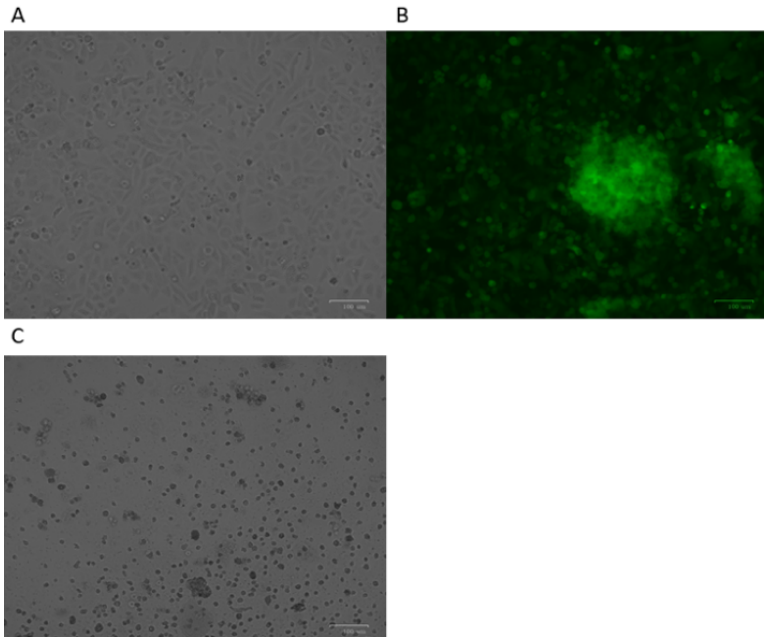

Fig. 2.4 Images from tissue culture. A. picture of successfully transfected V5_RBM17 adherent HeLa cells. This image is representative of all the successfully transfected tagged adherent cell lines. B. Expression of EGFP in the positive control, V5_EGFP adherent HeLa cells. C. Dead cells in the negative control well after puromycin selection.

2.13 Western Blots of V5_RBM17

I tested for express of V5_RBM17 using western blot analysis of whole cell extract. First, I blotted for RBM17, shown in figure 2.5 A. Single bands showed up at ~45KDa for induced V5_RBM17 lanes. This was the appropriate molecular weight of V5_RBM17. Multiple bands showed up in the other lanes at ~45KDa. This could be different isoforms of RBM17. This may be an example of gene regulation. Image B

is a blot for V5. Cre was the negative control and showed no bands as expected. V5_EGFP un-induced showed a single band, and that band was significantly amplified in the V5_EGFP induced lane. The presence of a band in the un-induced lane was unexpected, but this shows that the inducible expression system is a bit leaky. The V5_RBM17 un-induced lane showed a single band at ~45KDa. This band was slightly amplified in the induced lanes. This data confirms that I was able to establish a cell line that expresses a tagged spliceosome protein. The final thing to consider is there was no difference between the intensity of the band found in the 24 hour and 48-hour induction period lanes. This was important for considering the parameters of the immunoprecipitation.

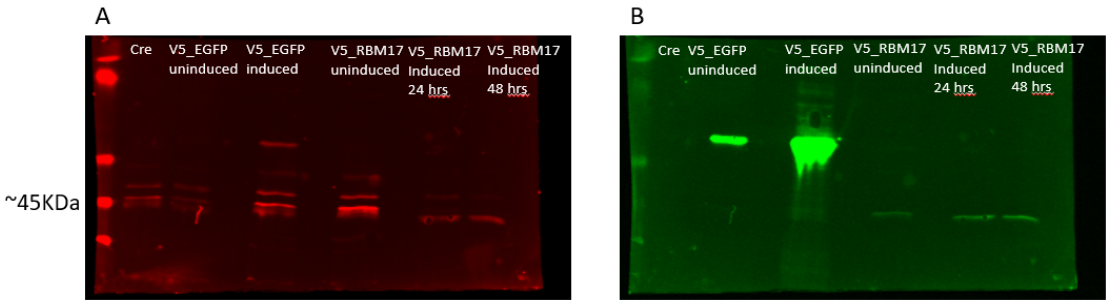

Fig. 2.5 V5_RBM17 western blot. Image A shows a western blot that used anti-RBM17, which recognizes RBM17. Image B is a western blot that used antiV5, which recognizes the V5 tag. Cre is the negative control which is the HeLa Crelox cells with an empty Crelox site.

2.14 Immunoprecipitation of V5_RBM17

Now that I confirmed that I was able to create the V5_RBM17 cell line, the next step was to test if we could isolate the tagged protein using immunoprecipitation. First, we blotted for RBM17 (Fig. 2.6 A). A single band showed up in un-induced and induced input, un-induced flowthrough, and induced elution at ~45KDa. The band in induced input is slightly upshifted, as that is V5_RBM17. The band that is slightly lower in the un0induced lanes is the endogenous RBM17. Lastly, the most amplified band is in the induced elution lane.

Image B shows a blot for V5. The same band shows up in induced input. A very light band shows up in un-induced and induced flowthrough. The most amplified bands show up in the elution lanes. The band in the un-induced elution is the heavy chain of the anti-V5 antibody, which runs at roughly the same molecular weight as RBM17.

Image C is an overlay of image A and B. Yellow bands represent the presence of V5_RBM17. A slight yellow band shows up in the induced input. The most amplified band is in the induced elution. This confirms that we are able to isolate V5_RBM17 via immunoprecipitation.

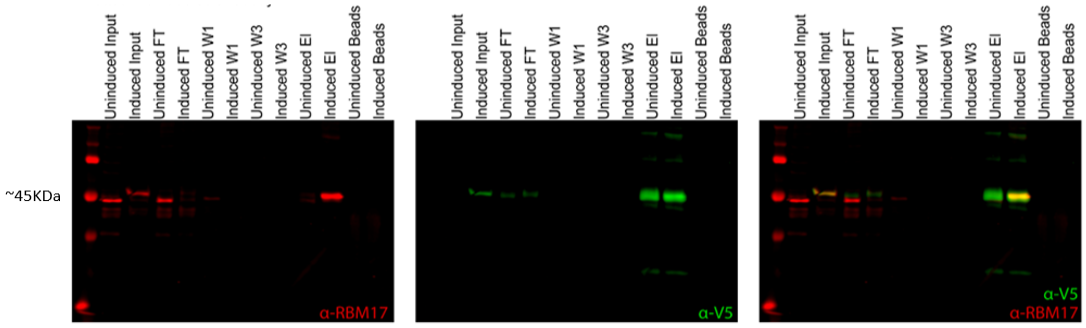

Fig. 2.6 Western blots of immunoprecipitations. These images depict the samples obtained from the immunoprecipitations which were tested by western blot. Images were created by Hannah Maul-Newby.

2.15 Testing for Pulldown of RBM17 Subcomplexes

The final thing we wanted to test is if other proteins were being pulled down with V5_RBM17. Figure 2.7 shows a blot for RBM17 and SF3B1. RBM17 is known to associate with SF3B1 (Fukumura 2021). Un-induced input shows a band for SF3B1 and endogenous RBM17. Induced input shows an amplified band for SF3B1, a dark band for V5_RBM17, and a light band for endogenous RBM17. This shows that inducible expression is functional. Un-induced and induced flowthrough only show a band for SF3B1. Un-induced elution only shows a band for SF3B1. This band is amplified in induced elution. This indicates that we are pulling down more SF3B1 with V5_RBM17. There is also a dark band for V5_RBM17 in the induced elution. Un-induced beads showed a band for SF3B1. Induced beads showed a band for both

SF3B1 and V5_RBM17. This shows that we were not able to fully elute all the protein from the beads.

The difference between the amplification of SF3B1 in the elution lanes shows that we can pulldown RBM17 subcomplexes. This is exciting as we can start testing for other proteins that interact with RBM17 and the conditions required for these interactions.

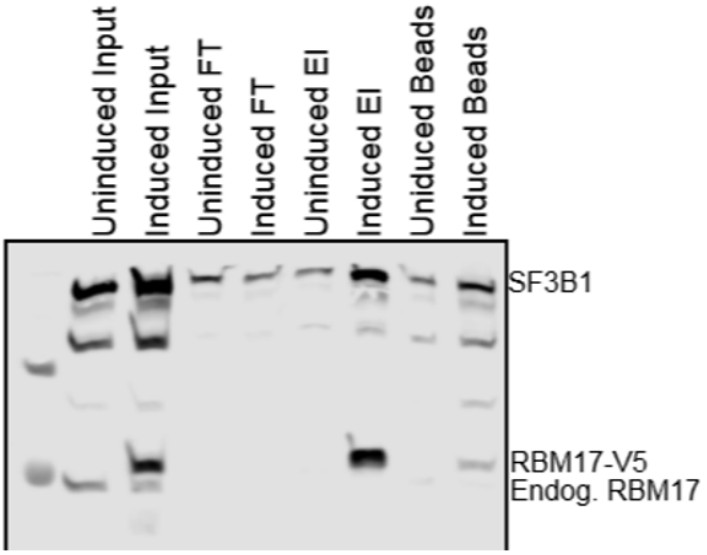

Fig. 2.7 Blotting for SF3B1. Western blot from immunoprecipitation using nuclear extract. This image was produced by Hannah Maul-Newby.

2.16 Western Blot for V5_SF3B4

The other spliceosome protein I was interested in tagging was SF3B4. I followed the same protocols to grow up confluent cells transfected with V5_SF3B4 plasmid. I then ran western blots. Figure 2.8 A shows a blot for SF3B4. All cell lines

both un-induced and induced showed a band at ~44KDa indicating expression of SF3B4. Image B is a blot for V5. There are no bands present. This shows that only endogenous SF3B4 is being expressed and not any V5_SF3B4.

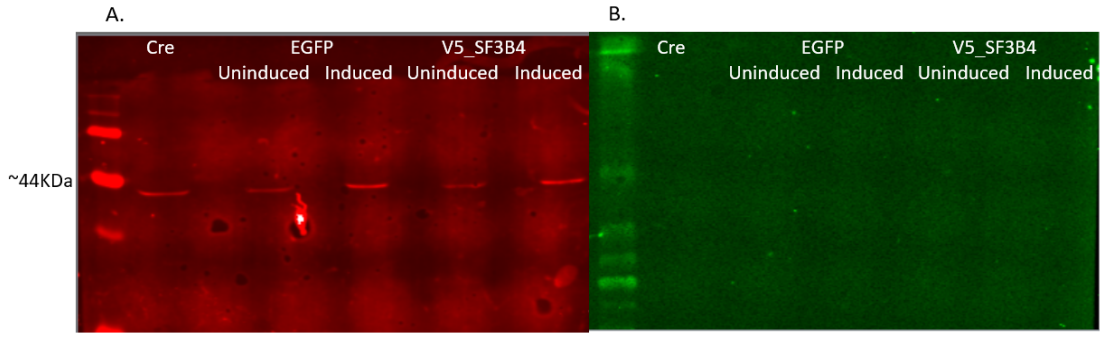

Fig. 2.8 Western blots for V5_SF3B4. Western analysis of whole-cell extract using A) SAP49 as the primary antibody and B) antiV5 as the primary antibody. Cre is the negative control which is the HeLa Crelox cells with an empty Crelox site.

2.17 Genomic DNA PCR

I wanted to investigate why the cells were not expressing V5 tagged SF3B4. The first place to look was the genomic DNA. I needed to test if the locus containing the inducible trans gene was incorporated into the genome of the cells. I did this through pcr. EGFP showed a band at ~1.2 Kb, which confirms integration of the trans gene locus. V5_SF3B4 results were not definitive. There was an amplified band at ~1.8Kb, which was the appropriate size. There was also a band at the same size for the negative control (Cre). Cre should not have a band as this cell line had an empty Cre locus. This band was much fainter than V5_SF3B4.

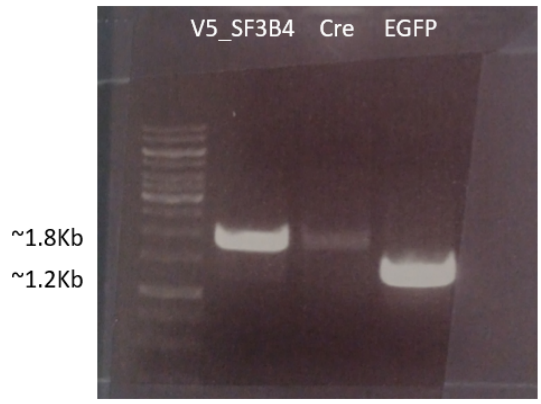

Fig. 2.9 Genomic DNA PCR of V5_SF3B4. Genomic DNA PCR done to test for incorporation of tagged protein genes into the genome of adherent HeLa cells. Cre acted as a negative control and V5_EGFP acted as a positive control.

2.18 Sequencing of the gDNA PCR Product

To determine the identity of the V5_SF3B4 gDNA pcr product, I got it sequenced. The alignment of the sequencing data to the expected sequence shown in Figure 2.10 confirms that the product was the V5_SF3B4 DNA. This confirms that the transgene locus was incorporated into the genome of the cells. The DNA sequencing data also highly suggests that there are no mutations within these genes. The two highlighted marks within the SF3B4 gene were shown to be the correct nucleotides based on the DNA sequencing readout file that was sent with the DNA sequencing file.

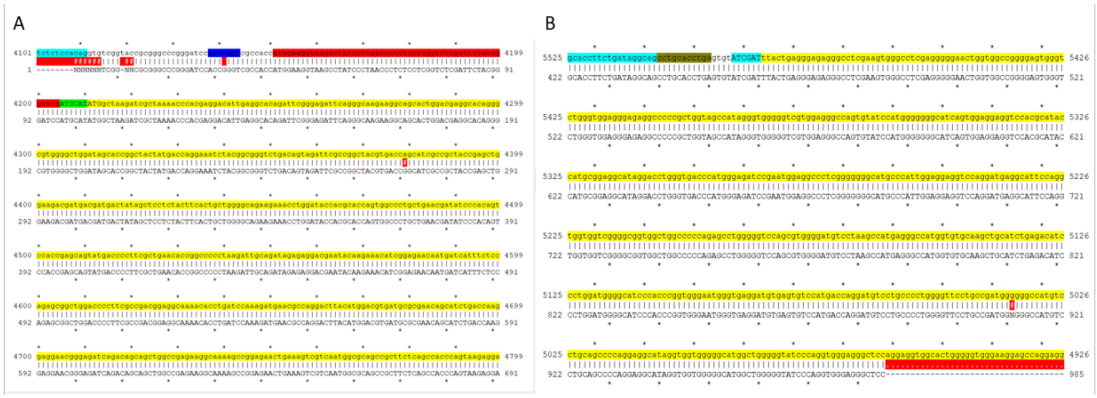

Fig 2.10 Alignment of sequenced DNA with DNA sequence. Image A shows the alignment of the 5' side of the V5_SF3B4 DNA sequence. The nucleotides highlighted in red on the top strand are the V5 gene. The nucleotides highlighted in yellow are the SF3B4 gene. Image B shows the alignment of the 3' side of the SF3B4 DNA sequence.

2.19 Reverse Transcription PCR

The next thing to check was the RNA. I did this through reverse transcription pcr. I had a few different conditions. An experimental condition which looked to amplify the RNA of the trans gene. A negative control that was a no RT pcr reaction. This would test for DNA contamination. A positive control for cyclin A2 RNA, which should be present in all cell lines. Also, a positive control that used the plasmid as the template to test for the pcr parameters.

The EGFP and V5_SF3B4 plasmid lanes both showed amplified bands at ~800bp and ~1.8Kb, showing that the pcr parameters were good. All the cell lines

showed a band for cyclin A2 RNA at ~350bp. This confirms that the RNA is each sample is intact. The band for EGFP was much less amplified than for the other cell lines. There was no amplification of the trans gene RNA for any of the cell lines. Cre was the negative control, so lack of amplification was expected. The lack of amplification of EGFP was unexpected as expression of this protein was confirmed visually (Fig. 2.4 B). I connected this to the lack of amplification of its positive control. Lastly, V5_SF3B4 RNA was not amplified, but this cell line had clear amplification of its positive control. The lack of amplification of the V5_SF3B4 RNA leads me to believe that the cells may be degrading the RNA.

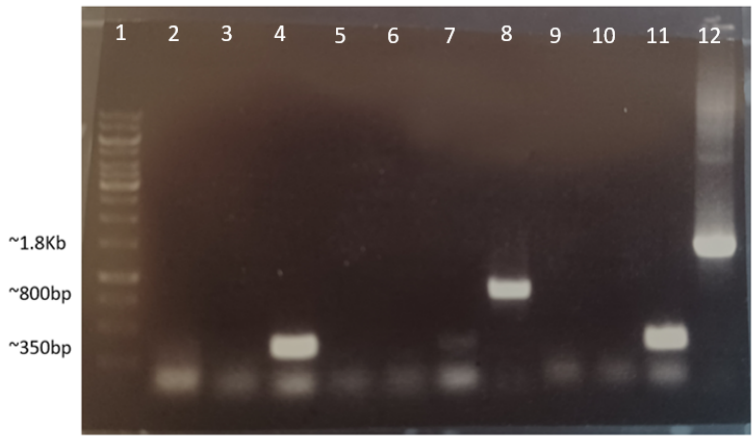

Fig. 2.11 RT PCR. Lanes are 1- 1Kb ladder, 2- Cre experimental, 3- Cre no RT, 4- Cre (+) control, 5- EGFP experimental, 6- EGFP no RT, 7- EGFP (+) control, 8- EGFP plasmid, 9- V5_SF3B4 experimental, 10- V5_SF3B4 no RT, 11- V5_SF3B4 (+) control, 12- V5_SF3B4 plasmid.

2.20 Second Attempt at Tagging SF3B4

I needed to take a new approach in tagging SF3B4. The first thing I considered was is the V5 tag interrupting spliceosome assembly or SF3B4 folding. To address this, I could try a different tag such as streptavidin binding protein (SBP) and move the tag from the amino terminus to the carboxyl terminus. Another thing I considered that the passage number on the cells got fairly high. This could affect the ability of the cells to express the trans gene. To address this, I could try to transfect again with V5_SF3B4 and be sure to keep the passage number as low as possible. So, I transfected and grew up cell lines SF3B4_SBP, V5_SF3B4, and V5_EGFP. Then I ran western blots (Fig. 2.12).

Image A shows blotting for SF3B4 and once again all cell lines are expressing SF3B4. Image B shows blotting for V5. There is no band in the V5_SF3B4 lanes showing that once again the cells are not expressing V5_SF3B4. V5_EGFP did show a band in un-induced and induced lanes. That band is significantly amplified in the induced lane. Image C shows blotting for SBP. All lanes showed non-specific bands at the top of the image, but these non-specific bands are consistent across all lanes. There are many bands in lane 9 which is a positive control, SBP ladder. This shows that the antibody is working. There are no bands at ~44KDa for SF3B4_SBP lanes, indicating no expression of SBP tagged SF3B4.

These results show that I was unable to tag SF3B4. I hypothesize that it may be difficult to tag this protein because it plays an important role in splicing. SF3B4 is part of the complex which recognizes the branch point sequence of the intron (Wilkinson 2020).

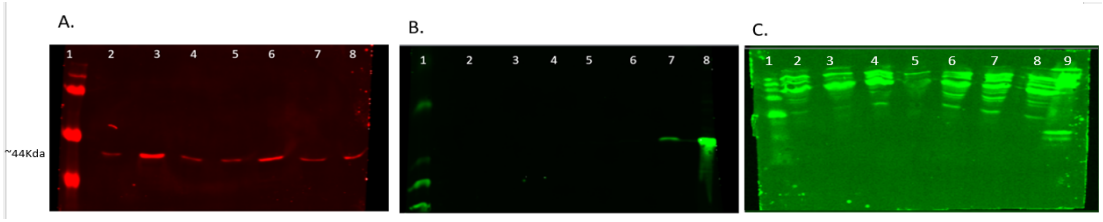

Fig. 2.12 Tagged SF3B4 western blot. Image A shows a western blot that used SAP49, which recognizes SF3B4. Image B shows a western blot that used anti-V5, which recognizes the V5 tag. Image C shows a western blot that used streptavidin, which recognizes the SBP tag. Lane 1 is the protein PAGE plus ruler, lane 2 is SF3B4_SBP uninduced, lane 3 is SF3B4 induced, lane 4 is Cre, lane 5 is V5_SF3B4 uninduced, lane 6 is V5_SF3B4 induced, lane 7 is V5_EGFP uninduced and lane 8 is V5_EGFP induced. For image C lane 9 is a positive control for the SBP antibody.

2.21 Discussion: Creation of Tagged Cell Lines

SF3B4 is part of the A complex of the spliceosome. The A complex is formed during the point in the splicing cycle when the branch point sequence of the intron is recognized (Wilikinson et al. 2020). SF3B4 is a haplo-insufficient gene and a heterozygous loss of function mutation in this gene results in Nager syndrome (Petit

et al. 2014). Over expression of SF3B4 can lead to the most common type of liver disease, hepatocellular carcinoma (Shen and Nam 2018).

RBM17 plays a role in the regulation of splicing. RBM17 dependent splicing changes corelated with the use of cryptic splice sites (Fukumura et al. 2021). Also, a spliceosome protein, DHX15, is regulated by a G-patch protein. A possible candidate for this regulation is RBM17 (Agafonov et al. 2011). These are the reasons I tried to tag these proteins, as it would allow us to further investigate the interactions that these proteins make with other spliceosome proteins.

Use of powerful tools and methods were imperative to the creation of cell lines that express tagged versions of these proteins. The first step in creating these cell lines was the assembly of the donor plasmid that would be used in transfections. I used NEB's HiFi DNA assembly kit, which was a Gibson style reaction. Successful assembly was accomplished multiple times using this method. The transfection method used was recombination mediated cassette exchange. This method involved a co-transfection of the donor plasmid with a wild type Cre recombinase encoding plasmid. The utility of this method was that it incorporated a locus from the donor plasmid into the genome of the HeLa cells, thereby establishing stable cell lines.

I was unable to tag SF3B4. I hypothesized that the reason for this is it has a very important role in splicing. It is part of the complex that recognizes the branch

point sequence (Wilkinson et al. 2020). When investigating where the issue arose, it was found that the cell was either not expressing the trans gene RNA or it was degrading it (Fig. 2.11). It is possible that the protein levels of an integral part of the cell's molecular machinery are tightly regulated.

I was able to establish the V5_RBM17 cell line. The gene regulation previously mentioned is displayed in Figure 2.5 A. The endogenous RBM17 shows up as multiple bands at ~45KDa, which are different isoforms of RBM17. This is seen in lanes Cre, V5_EGFP, and uninduced V5_RBM17. In the induced V5_RBM17 lanes there is a single band at ~45KDa. This shows that upon induction of expression of the tagged protein, there is a reduction in the expression of the endogenous protein. I hypothesize that I was able to tag RBM17 more easily because it is an alternative splicing regulator and is present when it outcompetes U2AF2 for binding to SF3B1 (Fukumura et al. 2021).

SF3B4 is an important and interesting protein worth further pursuit. It has been showed to be difficult to tag, but there could be other ways to learn more about its role in splicing. We could try tagging another protein within the SF3B complex. This would allow us to isolate the complex that contains SF3B4.

Isolation of RBM17 subcomplexes was successful (Fig. 2.7). The next step is to test for the presence of other proteins. Also, we could test for the conditions

necessary for these interactions such as the presence of an intron. This will help to further elucidate the role of RBM17.

Chapter 3

Splicing Inhibiting Drug Assay

3.1 Splicing Inhibiting Herboxidiene Drug

The herboxidiene drug is a splicing inhibitory drug that targets the SF3B complex. It has been shown to be a potent inhibitor of cancer cell growth. I analyzed how different chemical modifications affected the strength of the drug. I also compared my data collected from an in-vivo system to previously published work by Gamboa-Lopez which studied these drugs in-vitro. These results will allow us to see how chemical modifications can modulate the strength of the herboxidiene drug.

3.2 Methods of the In-vivo Drug Assay

A 96 well plate was seeded with adherent HeLa Crelox cells at a concentration of 5000 cells per well in 100µL of media for columns 2 through 11. Columns 1 and 12 received 100µL of media. The 96 well plate was incubated at standard conditions ~24 hours.

After the incubation period, remove 50µL of media from each well. Next a drug plate is created. 190µL of media is added to column 2 and 100µL of media is added to columns 1 and 3 through 12. A stock of 3% DMSO is made. Then a 300µM stock of the two drugs being tested is made as well. 10µL of DMSO is added to 2A

and 2H of the 96 well drug plate. 10µL of drug 1 is added to wells 2B, 2C and 2D. 10µL of drug 2 is added to wells 2E, 2F and 2G. A serial dilution from column 2 to column 10 is done by transferring 50µL of drug media to the adjacent wells. Also transfer 50µL from column 2 to column 12, resuspend and then dispose 50µLs. Column 11 gets no drug. Then transfer 50µL from the drug plate to the corresponding well in the 96 well plate containing cells. Incubate the 96 well plate for ~72 hours at standard conditions. After the incubation period add 5µL of Reliablue reagent to each well, resuspend and incubate for 1 hour. Then take fluorescence readings using the Thermo Scientific Varioskan Lux Plate Reader using 560nm for excitation and 590nm for emission.

The analysis of data was done using Microsoft Excel. The average reading from the wells with media only was averaged as background. This value was subtracted from the reading from the other wells. Each column for DMSO, drug 1 and drug 2 were then averaged with the outlier values removed. The results were plotted using PRISM with drug concentration vs percent cell viability. The percent viability was determined by normalizing the fluorescence data against the DMSO.

3.3 Development of Parameters for the Drug Assay

The concept of the drug assay I developed was simple; seed a 96 well plate with adherent HeLa cells, treat with drug, and then add a dye which would change from blue to colorless based on the concentration of viable cells in the well. It was

not that simple though, as there were many parameters that needed to be developed. The first thing I did was create standard curves for both fluorescence and absorption (Fig 3.1). From this curve, I determined that a good starting concentration is 5000 cells per well since HeLa cells double approximately every 24 hours. For drug assays I seed the plate at 5000 cells per well, incubated for ~24 hours, and then added drug. Another other thing I decided based on the standard curve, was to use fluorescence rather than absorption for the Reliablue reagent reading. The data analysis was more straight forward. For absorption, I was dealing with numbers less than one, and it was easier to see trends in the raw data for fluorescence.

I also needed to consider drug concentration and incubation period for the drug. Initially I was starting at 5μM and doing a 1:3 serial dilution with an incubation period of 48 hours. This dose was too low and did not give a good S curve that showed plateauing of the drug strength. I changed to a starting drug concentration of 15μM and an incubation period of 72 hours, which gave a nice S curve (Figure 3.2). The last thing I needed to consider was the amount of time I would incubate after adding the Reliablue reagent. I initially tried 30 minutes, 1 hour, and 2 hours. I found that there was not much of a change in color after 30 minutes. I decided that a 1-hour incubation period gave the best results, as it became visually apparent which wells had viable cells.

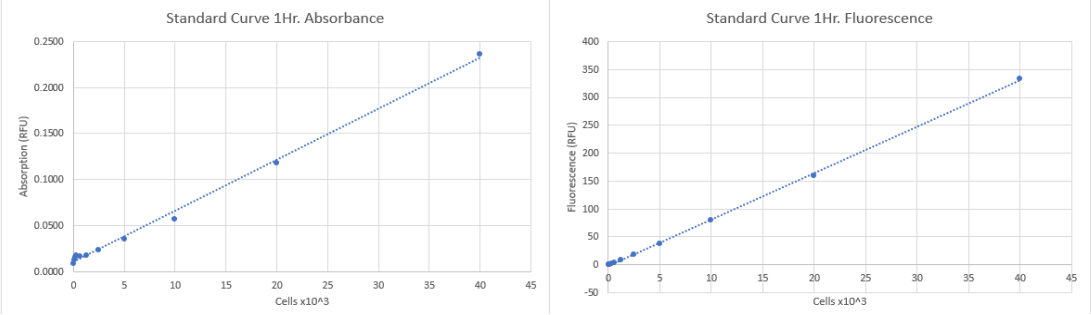

Fig. 3.1 Standard curve for absorbance and fluorescence readings. Standard curves were created to compare the readings to the number of viable cells per well.

3.4 Results for In-vivo Drug Assay

Figure 3.3A shows the basic chemical structure of the splicing inhibitory drugs. Three different classes of drugs were tested, all with this chemical structure but with different modifications. The individual chemical structures are given in Figure 3.4. The amide substitution changes the charge on a site that influences inhibition. The C5 modification is a site that could potentially serve as a location for the addition of a cross-linker. The C6 modification adds a chemical group to a site that influences binding.

The data shows mixed results for the addition of an amide at carbon 1. In the cell viability assay, HB23-19 was stronger than the parent drug, but it was classified as a weak inhibitor in the in-vitro study. HB23-19 had an IC50 of 0.0062µM which is 10 times stronger than the parent drug, IC50 of 0.071µM. HB24-19

(IC50=0.025µM) had approximately the same strength as the parent drug, HB3-20 lost some strength (IC50=0.17µM), while HB25-19 was the weakest (IC50=0.31µM).

The group of drugs with an addition of a hydroxyl at carbon 5 mostly resulted in weak drugs. HB2-20 (IC50=0.096µM) maintained a similar strength as the parent drug (IC50=0.071µM), which contradicted the in-vitro data. HB25-19 (IC50=0.31µM) and HB1-14 (IC50=12.3µM) were weak drugs, which matched the in-vitro data. Lastly, HB4-13 (IC50=30.9µM) and HB3-13 (IC50 too high to calculate from data) were inactive in the in-vivo assay but was classified as a strong inhibitor based on the in-vitro data. These two drugs were enantiomers.

The last group was the carbon 6 modification; HB19-19 (IC50=0.0074µM) was strong, HB8-19 (IC50=4.48µM) and HB10-18 (IC50 too high to calculate from data) were weak. This agreed with the in-vitro data.

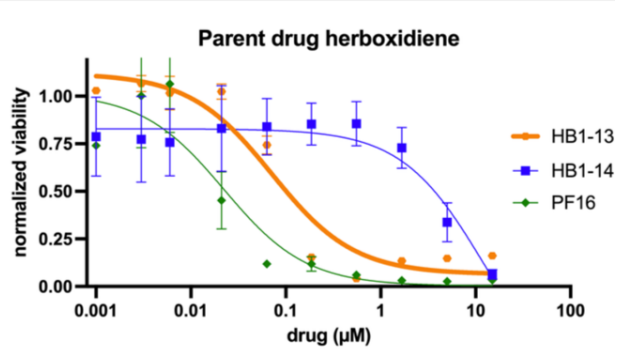

Fig. 3.2 Drug control. This is an example of strong and weak drugs. PF16 is a known strong drug produced by Pfizer, HB1-13 is the parent drug and HB1-14 is an inactive drug invitro.

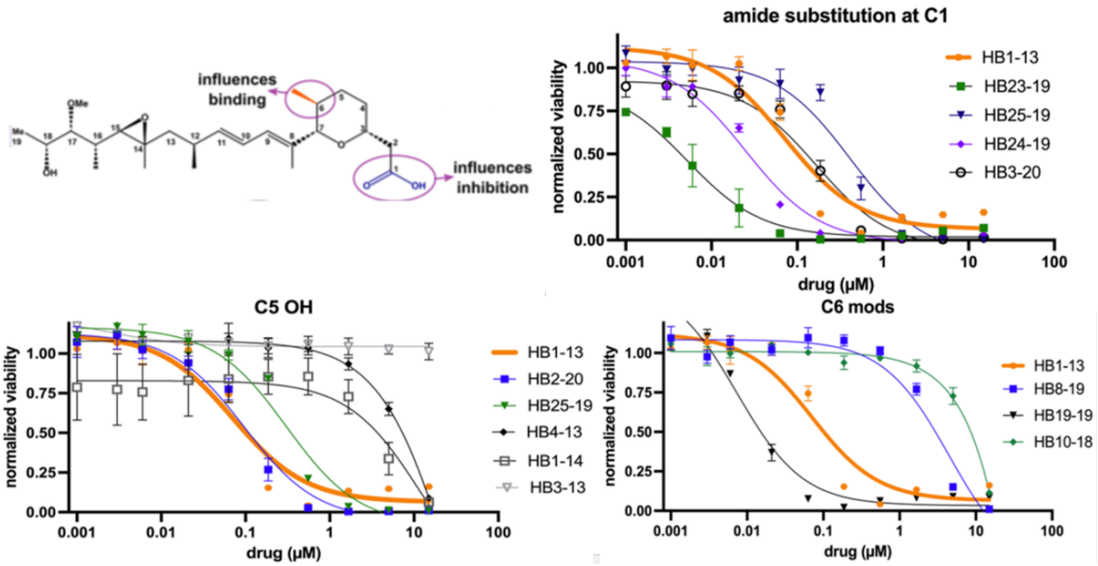

Fig. 3.3 Splicing inhibitory drug's effect on HeLa cell viability. Image D is the basic chemical structure of the splicing inhibitory drug. The title of each plot represents where the modifications take place.

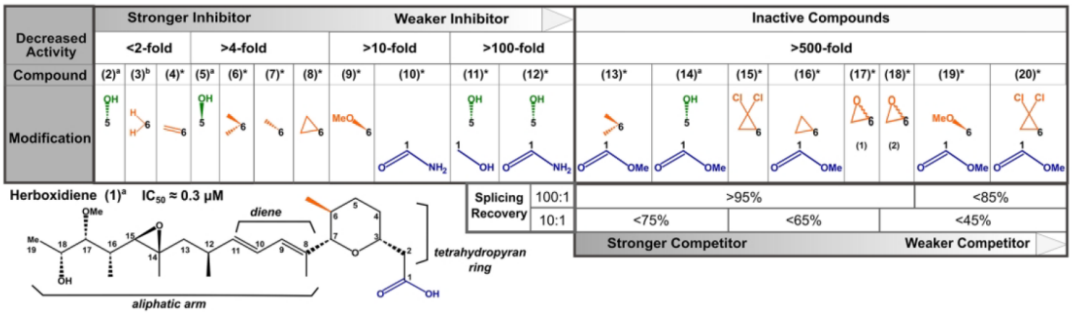

Fig. 3.4 In-vitro drug assay data. This is the in-vitro drug assay data that will be compared to the in-vivo data. This image was adapted from Gamboa Lopez et al. 2021.

Drug	in-vitro data chemical modification number	IC50 invivo (uM)	Classification in-vitro assay	Interpretation
HB1-13	Herboxidiene (parent drug)	0.071	strong inhibitor	n/a
PF16	control strong drug	0.0086	n/a	n/a
HB23-19	10	0.0062	weak inhibitor	results contradict
HB24-19	amide group on C1, large cross-linker	0.025	inactive	results contradict
HB3-20	secondary amide	0.17	weak inhibitor	results agree
HB25-19	12	0.31	weak inhibitor	results agree
HB2-20	11	0.096	weak inhibitor	results contradict
HB25-19	12	0.31	weak inhibitor	results agree
HB1-14	14	12.3	inactive	results agree
HB4-13	5	30.9	strong inhibitor	results contradict
HB3-13	2	inactive	strong inhibitor	results contradict
HB19-19	6	0.0074	strong inhibitor	results agree
HB8-19	15	4.48	inactive	results agree
HB10-18	13	inactive	inactive	results agree

Table 4. Comparing in-vitro to in-vivo data. Yellow represents controls, blue represents carbon 1 amide substitution, tan represents carbon 5 hydroxyl addition and green represents carbon 6 modification.

3.5 Discussion of Drug Assay

I developed an in-vivo drug assay, which analyzes the effects of drugs that target the SF3B complex. I specifically measured the viability of adherent HeLa cells after being treated with drug. The data collected was a follow up on research done by Adriana Gamboa Lopez, formerly of the Jurica Lab. She looked at the effects of these drugs in an in-vitro splicing assay.

The parent drug of the assay is a herboxidiene (Figure 3.3 A). I gathered data on how different modifications to the parent drug affected the strength of the drug. I separated the modified drugs into three groups, carbon 1 amide substitution, carbon 5 addition of a hydroxyl group and carbon 6 modifications. Analyzing these modifications will help us better characterize the drug as well as see how the drugs

strength can be modulated, which is important as one day this drug could be used as a cancer therapeutic.

First, I will consider the drugs with a substitution of an amide for a carboxylic acid on carbon 1. HB23-19 was ten times stronger than the parent drug. The only modification to this drug was the substitution of an amide group for the carboxylic acid. HB24-19 has a similar strength to the parent drug and the only difference between it and HB23-19 was the addition of a large cross-linking molecule at the end of the carbon chain. HB3-20 was weaker than the parent drug with the addition of a secondary amide group rather than a primary amide group. This causes a loss of polarity on carbon 1. HB25-19 was weak in both in-vitro and in-vivo assays, but it had the addition of a hydroxyl group to carbon 5, as well as the addition of an amide group to carbon 1. To conclude, in the in-vitro system substituting an amide group for the carboxylic acid on carbon 1 diminished the strength of the drug and in the in-vivo assay it had the opposite effect most of the time. I hypothesize that maintaining polarity on carbon 1 is important for the strength of the drug.

The next group of modifications that I will address is the addition of a hydroxyl group to carbon 5, which could serve as a site for the addition of a cross-linker. In all but one drug in the in-vivo assay the strength of the drug was diminished. HB2-20 was nearly as strong as the parent drug and had the addition of a hydroxyl on carbon 5 as well as the substitution of the carboxylic acid with a

hydroxyl. The modification to carbon 1 may help to maintain the strength of this drug. Something that was surprising was that HB4-13 and HB3-13 were very strong in-vitro but were inactive in-vivo. I hypothesize that the addition of this hydroxyl group may be inhibiting the drug's ability to get into the cells. Another possibility is that for at least HB4-13 and HB3-13, these drugs could have lost activity while in storage. They were synthesized in 2013. These two drugs need to be resynthesized and tested again before a conclusion can be made.

The last group is carbon 6 modifications. The in-vitro and in-vivo data agreed. The main thing that I would highlight is for HB19-19, the addition of an additional methyl group to carbon 6 resulted in a drug ten times stronger than the parent drug, but when you take away the polar group on carbon 1 while maintaining that extra methyl group, for instance in HB10-18 there is a total loss of activity. This reinforces my hypothesis that polarity on carbon 1 is important. The inactivity of HB8-19 was not surprising as it had the addition of a large bulky group with a couple of chlorines on it and it was inactive in-vitro.

The main takeaways from this study are that having a charged group on carbon 1 seems to be important, no matter what the charge is. The site influences inhibition but there seems to be some flexibility in the drug/target interaction. The addition of a hydroxyl to carbon 5 nearly always leads to a reduction in drug

strength. Lastly, the addition of an extra methyl group to carbon 6 leads to an increase in drug strength.

The parent drug is a powerful splicing inhibitor, but this data shows that this strength can be modulated based on chemical modifications. Studies such as this one is important as they give us further insight into a drug that one day could be used as a cancer therapeutic.

References

Agafonov, D. E., Deckert, J., Wolf, E., Odenwalder, P., Bessonov, S., Will, C. L., Urlaub, H., and Luhrmann, R. (2011). Semiquantitative proteomic analysis of the human spliceosome via a novel two-dimensional gel electrophoresis method. Mol Cell Biol *31*, 2667-2682.

Black, D. L. (2003). Mechanisms of alternative pre-messenger RNA splicing. Annu Rev Biochem *72*, 291-336.

Bohnsack, K. E., Ficner, R., Bohnsack, M. T., and Jonas, S. (2021). Regulation of DEAH-box RNA helicases by G-patch proteins. Biol Chem *402*, 561-579.

Champion-Arnaud, P., and Reed, R. (1994). The prespliceosome components SAP 49 and SAP 145 interact in a complex implicated in tethering U2 snRNP to the branch site. Genes Dev *8*, 1974-1983.

Charenton, C., Wilkinson, M. E., and Nagai, K. (2019). Mechanism of 5' splice site transfer for human spliceosome activation. Science *364*, 362-367.

Company, M., Arenas, J., and Abelson, J. (1991). Requirement of the RNA helicase-like protein PRP22 for release of messenger RNA from spliceosomes. Nature *349*, 487-493.

Cretu, C., Schmitzova, J., Ponce-Salvatierra, A., Dybkov, O., De Laurentiis, E. I., Sharma, K., Will, C. L., Urlaub, H., Luhrmann, R., and Pena, V. (2016). Molecular architecture of SF3b and structural consequences of its cancer-related mutations. Mol Cell *64*, 307-319.

Fica, S. M., Oubridge, C., Galej, W. P., Wilkinson, M. E., Bai, X. C., Newman, A. J., and Nagai, K. (2017). Structure of a spliceosome remodelled for exon ligation. Nature *542*, 377-380.

Fica, S. M., Mefford, M. A., Piccirilli, J. A., and Staley, J. P. (2014). Evidence for a group II intron-like catalytic triplex in the spliceosome. Nat Struct Mol Biol *21*, 464-471.

Fourmann, J. B., Dybkov, O., Agafonov, D. E., Tauchert, M. J., Urlaub, H., Ficner, R., Fabrizio, P., and Lührmann, R. (2016). The target of the DEAH-box NTP triphosphatase Prp43 in Saccharomyces cerevisiae spliceosomes is the U2 snRNP-intron interaction. Elife *5*, e15564.

Fukumura, K., Venables, J. P., and Mayeda, A. (2021). SPF45/RBM17-dependent splicing and multidrug resistance to cancer chemotherapy. Mol Cell Oncol *8*, 1996318.

Gamboa Lopez, A., Allu, S. R., Mendez, P., Chandrashekar Reddy, G., Maul-Newby, H. M., Ghosh, A. K., and Jurica, M. S. (2021). Herboxidiene Features That Mediate

Conformation-Dependent SF3B1 Interactions to Inhibit Splicing. ACS Chem Biol *16*, 520-528.

Hubert, C. G., Bradley, R. K., Ding, Y., Toledo, C. M., Herman, J., Skutt-Kakaria, K., Girard, E. J., Davison, J., Berndt, J., Corrin, P., Hardcastle, J., Basom, R., Delrow, J. J., Webb, T., Pollard, S. M., Lee, J., Olson, J. M., and Paddison, P. J. (2013). Genome-wide RNAi screens in human brain tumor isolates reveal a novel viability requirement for PHF5A. Genes Dev *27*, 1032-1045.

Igel, H., Wells, S., Perriman, R., and Ares, M. (1998). Conservation of structure and subunit interactions in yeast homologues of splicing factor 3b (SF3b) subunits. RNA *4*, 1-10.

Johnson, J. M., Castle, J., Garrett-Engele, P., Kan, Z., Loerch, P. M., Armour, C. D., Santos, R., Schadt, E. E., Stoughton, R., and Shoemaker, D. D. (2003). Genome-wide survey of human alternative pre-mRNA splicing with exon junction microarrays. Science *302*, 2141-2144.

Kelemen, O., Convertini, P., Zhang, Z., Wen, Y., Shen, M., Falaleeva, M., and Stamm, S. (2013). Function of alternative splicing. Gene *514*, 1-30.

Khandelia, P., Yap, K., and Makeyev, E. V. (2011). Streamlined platform for short hairpin RNA interference and transgenesis in cultured mammalian cells. Proc Natl Acad Sci U S A *108*, 12799-12804.

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

Kramer, A., Gruter, P., Groning, K., and Kastner, B. (1999). Combined biochemical and electron microscopic analyses reveal the architecture of the mammalian U2 snRNP. J Cell Biol *145*, 1355-1368.

Matera, A. G., and Wang, Z. (2014). A day in the life of the spliceosome. Nat Rev Mol Cell Biol *15*, 108-121.

Mizui, Y., Sakai, T., Iwata, M., Uenaka, T., Okamoto, K., Shimizu, H., Yamori, T., Yoshimatsu, K., and Asada, M. (2004). Pladienolides, new substances from culture of Streptomyces platensis Mer-11107. III. In vitro and in vivo antitumor activities. J Antibiot (Tokyo) *57*, 188-196.

Nilsen, T. W. (2003). The spliceosome: the most complex macromolecular machine in the cell? Bioessays *25*, 1147-1149.

Parker, R., Siliciano, P. G., and Guthrie, C. (1987). Recognition of the TACTAAC box during mRNA splicing in yeast involves base pairing to the U2-like snRNA. Cell *49*, 229-239.

Petit, F., Escande, F., Jourdain, A. S., Porchet, N., Amiel, J., Doray, B., Delrue, M. A., Flori, E., Kim, C. A., Marlin, S., Robertson, S. P., Manouvrier-Hanu, S., and Holder-Espinasse, M. (2014). Nager syndrome: confirmation of SF3B4 haploinsufficiency as the major cause. Clin Genet *86*, 246-251.

Plaschka, C., Lin, P. C., Charenton, C., and Nagai, K. (2018). Prespliceosome structure provides insights into spliceosome assembly and regulation. Nature *559*, 419-422.

Plaschka, C., Lin, P. C., and Nagai, K. (2017). Structure of a pre-catalytic spliceosome. Nature *546*, 617-621.

Rauhut, R., Fabrizio, P., Dybkov, O., Hartmuth, K., Pena, V., Chari, A., Kumar, V., Lee, C. T., Urlaub, H., Kastner, B., Stark, H., and Luhrmann, R. (2016). Molecular architecture of the Saccharomyces cerevisiae activated spliceosome. Science *353*, 1399-1405.

Semlow, D. R., Blanco, M. R., Walter, N. G., and Staley, J. P. (2016). Spliceosomal DEAH-Box ATPases Remodel Pre-mRNA to Activate Alternative Splice Sites. Cell *164*, 985-998.

Shen, Q., and Nam, S. W. (2018). SF3B4 as an early-stage diagnostic marker and driver of hepatocellular carcinoma. BMB Rep *51*, 57-58.

Spingola, M., Grate, L., Haussler, D., and Ares, M. (1999). Genome-wide bioinformatic and molecular analysis of introns in Saccharomyces cerevisiae. RNA *5*, 221-234.

Stamm, S., Ben-Ari, S., Rafalska, I., Tang, Y., Zhang, Z., Toiber, D., Thanaraj, T. A., and Soreq, H. (2005). Function of alternative splicing. Gene *344*, 1-20.

Steitz, T. A., and Steitz, J. A. (1993). A general two-metal-ion mechanism for catalytic RNA. Proc Natl Acad Sci U S A *90*, 6498-6502.

Tan, Q., Yalamanchili, H. K., Park, J., De Maio, A., Lu, H. C., Wan, Y. W., White, J. J., Bondar, V. V., Sayegh, L. S., Liu, X., Gao, Y., Sillitoe, R. V., Orr, H. T., Liu, Z., and Zoghbi, H. Y. (2016). Extensive cryptic splicing upon loss of RBM17 and TDP43 in neurodegeneration models. Hum Mol Genet *25*, 5083-5093.

Wen, X., Tannukit, S., and Paine, M. L. (2008). TFIP11 interacts with mDEAH9, an RNA helicase involved in spliceosome disassembly. Int J Mol Sci *9*, 2105-2113.

Wilkinson, M. E., Charenton, C., and Nagai, K. (2020). RNA Splicing by the Spliceosome. Annu Rev Biochem *89*, 359-388.

Wirth, D., Gama-Norton, L., Riemer, P., Sandhu, U., Schucht, R., and Hauser, H. (2007). Road to precision: recombinase-based targeting technologies for genome engineering. Curr Opin Biotechnol *18*, 411-419.

Xiong, F., and Li, S. (2020). SF3b4: A Versatile Player in Eukaryotic Cells. Front Cell Dev Biol *8*, 14.

Yu, Y., Maroney, P. A., Denker, J. A., Zhang, X. H., Dybkov, O., Lührmann, R., Jankowsky, E., Chasin, L. A., and Nilsen, T. W. (2008). Dynamic regulation of alternative splicing by silencers that modulate 5’ splice site competition. Cell *135*, 1224-1236.

Zamore, P. D., Patton, J. G., and Green, M. R. (1992). Cloning and domain structure of the mammalian splicing factor U2AF. Nature *355*, 609-614.

Zhang, Y., Weiner, A.M. (1986). A Compensatory Base Change in U1 snRNA Suppresses a 5' Splice Site Mutation. Cell 46(6), 827-835.