UC Santa Cruz

UC Santa Cruz Electronic Theses and Dissertations

Title

Quality Control in Early Spliceosome Assembly — A New Role for DHX15

Permalink

https://escholarship.org/uc/item/9s1757t9

Author

Maul-Newby, Hannah Marie

Publication Date

2023

Copyright Information

This work is made available under the terms of a Creative Commons Attribution-NonCommercial-NoDerivatives License, availalbe at https://creativecommons.org/licenses/by-nc-nd/4.0/

Peer reviewed|Thesis/dissertation

eScholarship.org Powered by the California Digital Library
University of California

UNIVERSITY OF CALIFORNIA
SANTA CRUZ

Quality Control in Early Spliceosome Assembly — A New Role for DHX15

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

DOCTOR OF PHILOSOPHY

in

MOLECULAR, CELL AND DEVELOPMENTAL BIOLOGY

by

Hannah M. Maul-Newby

March 2023

The dissertation of Hannah M. Maul-Newby is approved:

Professor Melissa S. Jurica, Chair

Professor Manuel Ares, Jr.

Professor Douglas Black, UCLA

Peter Biehl
Dean of Graduate Studies

Copyright © by

Hannah M. Maul-Newby

2023

Table of Contents

List of Figures

Chapter 1: Introduction

Chapter 2: A model for DHX15 mediated disassembly of A-complex spliceosomes

Chapter 3: Creation of splicing nuclear extracts from non-HeLa cell lines and HeLa derived lines

Chapter 4: : Isolation of spliceosomal complexes via tripartite purification Introduction

List of Tables

Appendix I:

Abstract

Quality Control in Early Spliceosome Assembly — A New Role for DHX15

Hannah M. Maul-Newby

Pre-mRNA splicing is critical for controlling normal gene expression within the cell. Several splicing associated RNA helicases and their binding partners are misregulated in a variety of cancers including leukemia, prostate, glioma, and non-small-cell lung cancer. Some of the RNA helicases from the DHX-family are involved in multiple cellular pathways, and to define their specific activity they rely on binding partners that contain G-patch domains. For instance, DHX15 has been characterized for its role in spliceosome disassembly after splicing and ribosome biogenesis directed by G-patch proteins TFIP11 and NKRF, respectively. My dissertation focuses on a new model for DHX15 in which it ensures fidelity of early spliceosome assembly. I hypothesize that DHX15 acts in a quality-control (QC) mechanism to disassemble aberrant spliceosomes. Due to the dynamic nature of the spliceosome, in that intronic landmarks are divergent as well as the multitude of rearrangements required to remove an intron, mistakes are bound to happen and the lack of a QC step during early assembly has long been a mystery. To address the role that DHX15 may be playing during early spliceosome assembly, I utilized a combination of RNA substrates and energy sources (ATP, GTP and AMP-PNP) to parse apart the different energy-dependent steps of A-complex assembly. Then through immunodepletion asked how removal of DHX15 affects the system. In attempting to determine where DHX15 binds within our complexes, I utilized a system of truncated RNA substrates and endogenous RNase H digestion of the U2

snRNA. Through this work, I found that depletion of DHX15 results in an accumulation of A- and B-complexes yet results in a decrease in splicing efficiency. Further, certain regions of the U2 snRNA are prone to NTP-dependent rearrangements, whereas others are protected or unaffected. In addition, the work in this thesis solely relied on HeLa cell nuclear extract and protein levels including DHX15 vary across cell types. Therefore, I worked to develop protocols and understand why other cell lines have yet to yield robust activity in nuclear extract in addition to why we observe inconsistent activity between HeLa nuclear extract preps. Finally, to understand how DHX15 may be impacting early complex formation, I devised a purification scheme to allow for early complexes to be purified for further study. Taken together, the data produced during my graduate career and described in this dissertation has yielded another path of study within the spliceosome and may provide a scaffold for future therapeutics in targeting diseases such as cancers and myelodysplastic syndromes.

Dedication

To Mumsie. Without you, none of this would have been possible.

Acknowledgements

While there are a many people who have been vital to my success, I would like to start off my acknowledgements by first thanking my PI and mentor, Dr. Melissa Jurica. Melissa allowed me to pursue a project that sparked my curiosities and listened when I was so very stuck even though my project was not in her wheelhouse, and for that I am so grateful. Thank you for allowing me to try so many different experiments and for not saying no (within reason). Thank you for taking the countless hours to edit my writing… we both know how terrible it was when I first started and for instilling your hatred of white space into my brain. ☺ In addition, thank you for keeping me on track and for always knowing the right thing to say.

To the Jurica Lab members, old and new: Thank you. Thank you for being my lab mates and for putting up with my daily ramblings. Thank you for challenging my science. I would especially like to thank Dr. Veronica Urabe, Beth Prichard, Angela Amorello, and Meredith Stevers for all the laughs in lab. In addition, I would like to thank my wonderful undergrad, Turvi Sharma, who I was lucky enough to mentor for three years. You directly contributed to my successes and without your daily help, this would not have been possible.

To all my friends. Thank you for the countless walks, laughs, many dinners, and hours of complaining. It was necessary. It was perfect. To Tyler, thank you for being my best friend. To Tommy, thank you for getting Ham, my dogphew, and for keeping my head on straight. To Megan, thank you for always being happy and ready to do whatever is needed. To Sadaf, thank you for loving the things I love and being as excited as I am about the things our other friends could care less about. To Ruhi, thank you for always sharing your peach rings and for always being there.

To my family. Thank you for always supporting me and helping me to achieve my dreams. To Kim, thank you for loving me no matter what and making sure Penelope and I are fed. I am sorry I was so bad at keeping up with the laundry.

The experiments I performed in my PhD heavily relied on HeLa cell nuclear extract. HeLa cells are an immortal cell line that were derived from a cervical tumor of then 31-year-old Henrietta Lacks at Johns Hopkins Hospital. Mrs. Lacks was an African American woman who in 1951 was diagnosed with a very aggressive cervical tumor. When her tumor was biopsied, some of her cells were shared with scientists without her permission or knowledge. The cells were found to thrive in culture, and while Mrs. Lacks life ended in 1951 because of her disease, her cells lived on without her or her family's knowledge. HeLa cells were widely distributed around the world and have been key to many successes within the scientific community as a result. My success during my PhD is a direct result from the actions of scientists 70 years ago, and while I do not condone what was done nor can I fix the injustices of previous generations, what I can do is personally acknowledge Henrietta for her contributions to my science and the work of countless others. And strive moving forward to ensure no one forgets the woman behind the cells.

This dissertation was supported by a Chancellor's Dissertation Year Fellowship and the Robert Ludwig Memorial Fellowship.

Chapter 2 of this dissertation is a modified version of material previously published in the journal, RNA:

Maul-Newby HM, Amorello AN, Sharma T, Kim JH, Modena MS, Prichard BE, Jurica MS. A model for DHX15 mediated disassembly of A-complex spliceosomes. RNA. 2022 Apr;28(4):583-595. doi: 10.1261/rna.078977.121. Epub 2022 Jan 19. PMID: 35046126; PMCID: PMC8925973.

Chapter 1: Introduction

Central Dogma and Eukaryotic Gene Expression

Nucleic acids have a foundational role in life. Deoxyribonucleic acid (DNA) provides the genetic blueprint that our cells need to generate new enzymes, proteins, and cellular organelles. For this to happen, the cell follows the central dogma of biology (Fig. 1.1) which states that DNA is transcribed by RNA polymerase II (Pol II) into ribonucleic acid (RNA) which is then translated by the ribosome into protein.

Due to excess genetic information within an RNA transcript, it requires processing prior to translation. Immediately post-transcription, the precursor messenger RNA (pre-mRNA) contains both coding and non-coding regions and as a result, the non-coding regions must be removed prior to translation to ensure only the coding regions are incorporated into the final protein product. This coding region only RNA is referred to as the messenger RNA (mRNA) and will serve as the template for translation.

The process by which introns are removed, and the adjacent exons are ligated together is called pre-mRNA splicing and is carried out by a macromolecular complex called the spliceosome. The removal of introns by the spliceosome is a highly regulated process that has been studied extensively; however there are many questions that remain unanswered such as: how does the spliceosome define or assemble upon a given intron? In addition, it has yet to be determined how many spliceosomal complexes or rearrangements occur during intron removal (especially early in the splicing cycle) nor do we know which enzymes promote or catalyze those changes. However, understanding the

rearrangements required to engage an intron as well as the enzymes involved will be explored within the body of this dissertation.

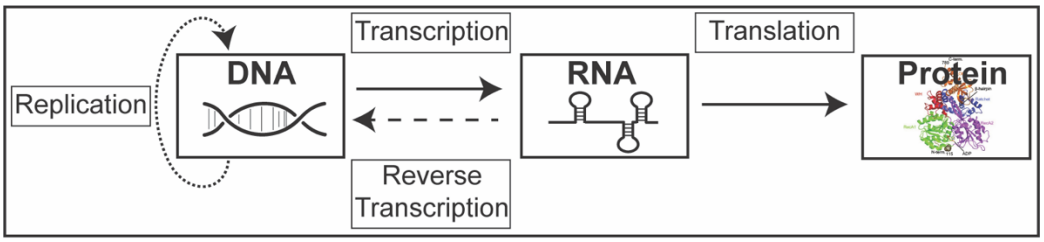


Figure 1.1: Central Dogma of Biology

Pictorial schematic of DNA being transcribed into RNA where the pre-mRNA is capped and then spliced prior to the addition of a polyadenylated tail and exported out of the nucleus. After exportation, the mRNA is translated by the ribosome into protein. The protein structure of DHX15, PDB: 5XDR.

Intronic Landmarks

The genome contains the genetic information necessary to create all mRNAs and non-coding RNAs. However, immediately post release from Pol II the pre-mRNA contains both coding/untranslated regions (exons) and non-coding (introns) regions that must be removed prior to translation by the ribosome (Fig. 1.2) (Berget et al., 1977; Chow et al., 1977). These transcripts possess a 5' methyl guanosine cap and will be polyadenylated on the 3' end after intron removal (Jurado et al., 2014). The removal of introns and the ligation of adjacent exons occurs through the process of pre-mRNA splicing.

Within a pre-mRNA are several sequence landmarks that define the boundaries of exons and introns. The ends of an intron are denoted by the 5' and 3' splice sites (SS), which are determined by GU and AG nucleotides, respectively (Darman et al., 2015). Between the splice sites, is a branch point sequence (BPS) that contains the branch point adenosine, a critical residue that will participate in splicing chemistry. Unlike Saccharomyces cerevisiae (yeast),

which has a conserved branch point sequence (UACUAAC), the human branch point is highly divergent however possesses a consistent YNYURAY nucleotide sequence, with N being any nucleotide, Y's representing pyrimidines and R's representing purines (Fig. 1.2). Immediately downstream of the branch point sequence is the polypyrimidine tract (PYT), a series of pyrimidine bases immediately prior to the 3' splice site. Together, these landmarks define the boundaries and are targets for spliceosome components during the mechanism of splicing (Fig. 1.2).

These landmarks are well defined across transcripts yet are common hotspots for mutations that contribute to disease (Hautin et al., 2020), yet how the spliceosome recognizes these vital landmarks has not been fully determined. This is especially true regarding the branch point sequence, as many human introns contain multiple potential branches, yet only one will yield the correct mRNA transcript. This is further impacted by the observation that the 3' splice site is the first AG pair that is seven nucleotides downstream of the chosen branch point sequence (Wahl et al., 2009) Therefore, not only does the branch point sequence contribute to splicing chemistry, but it also determines the 3' splice site. Taken together, the assembly of the spliceosome onto the correct branch sequence is vital in ensuring the correct intronic sequence is removed and the adjacent exons are ligated together to form the proper mRNA

It has long been known that assembly onto the intron landmarks contributes to cellular health, yet it is largely unknown how many spliceosomal components assist in choosing the correct splice site, branch or PYT. We have lots of evidence as to what happens during later spliceosomal complexes, yet the

early decisions have long eluded discovery at the same rate. Much of our understanding relies on yeast or *in vitro* studies of select human cell lines that have provided robust activity in a test tube. However, this has limited our goals unlocking questions regarding early spliceosome assembly and intron landmark choice due to the limited variability and difficulty in managing scale. To that end, developing new techniques that will allow study of human spliceosomal choices will be vital to exploring the unknown aspects of splicing and the work in later chapters of this dissertation directly address this issue.

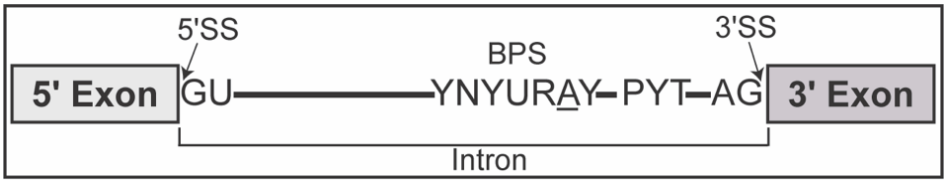


Figure 1.2: Intronic Landmarks
Schematic of intronic landmarks. The 5' and 3' splice sites (SS) are denoted with arrows immediately adjacent. The divergent human branch point sequence (BPS) is located near the 3' end of the intron with the branch point adenosine (A) underlined. Upon formation of the branch helix the branch point adenosine will bulge out into the active site allowing it to facilitate the first nucleophilic attack. Finally, immediately downstream of the branch is the polypyrimidine tract (PYT).

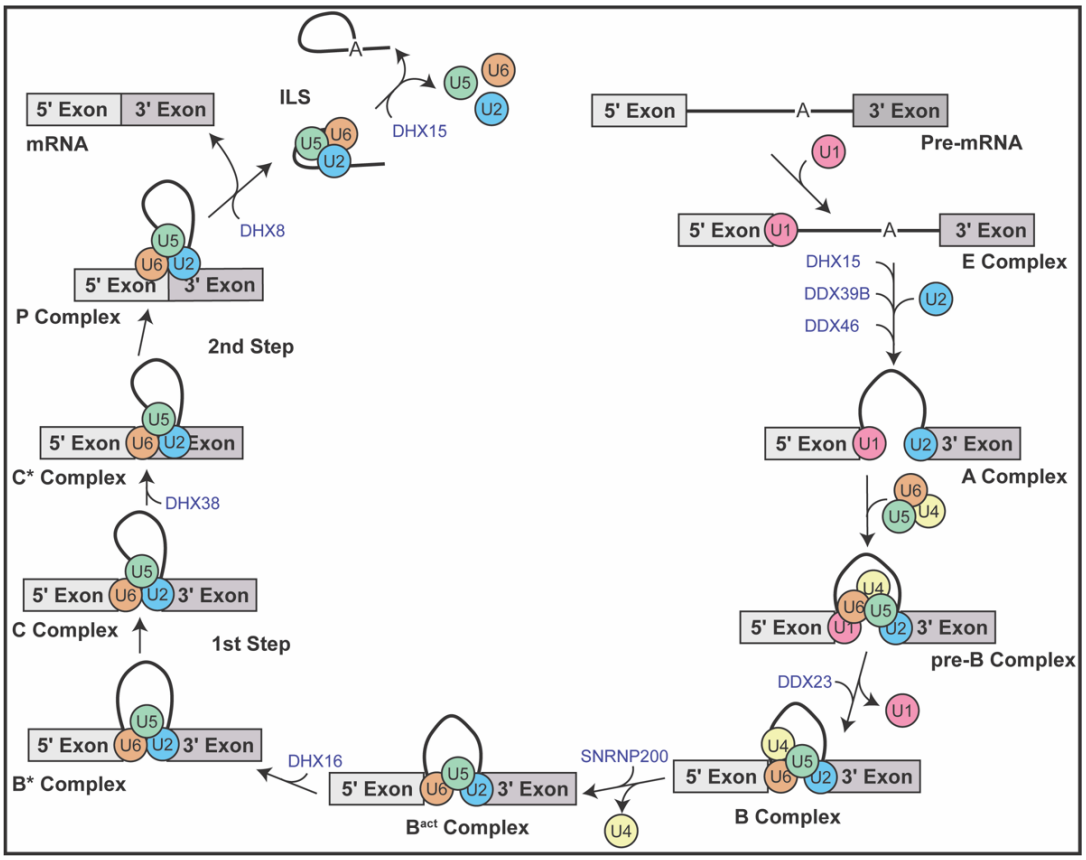
The Spliceosome and Pre-mRNA Splicing

The splicing process is carried out by a molecular machine called the spliceosome. The spliceosome is a macromolecular complex of many proteins and five snRNAs that combine and recombine to remove every intron within a pre-mRNA transcript. The proteins and snRNAs unite to form five small nuclear ribonucleoproteins (snRNPs) (U1, U2, U4, U5 and U6 snRNPs) which facilitate the necessary rearrangements required to remove an intron. This coordinated dance results in the formation of several complexes throughout the splicing cycle, each with their own important function.

The splicing cycle has been categorized by a series of complexes that rearrange as the process progresses (Fig. 1.3). The first complex formed is E-complex, which forms after the binding of the U1 snRNP to the 5' SS of a given intron. The second step of splicing occurs when the U2 snRNP binds the branch point sequence, rearranging to form a stable branch helix and competent A-complex. The formation of A-complex is essential as the bulging of the branch adenosine in the branch helix acts as the nucleophile in the first step of splicing and will contribute to 3' SS choice. After formation of a competent A-complex, the tri-snRNP will be recruited to the transcript and an initial rearrangement will result in B-complex, followed by an internal rearrangement which ejects the U1 and U4 snRNPs resulting in B^{act}. A final internal rearrangement will lead to B* which is the final spliceosome complex prior to the first step of splicing chemistry.

Splicing chemistry occurs through two well-characterized S_N2 transesterification reactions (Fig. 1.4), in which the 2'-hydroxyl group of the exposed branch adenosine, acts as a nucleophile and attacks the phosphodiester bond of the 5' splice site, UG, resulting in the formation of the intron lariat and C-complex. Upon formation of C-complex, the second step of splicing occurs in which a second S_N2 attack occurs in which the 3' hydroxyl group of the free 5' exon attacks the phosphodiester bond of the AG at the 3' splice site at the 3' boundary of the intron. This secondary attack results in the ligation of the 2 exons and the free-intron lariat and P-complex. The intron lariat spliceosome (ILS) then releases the mRNA from the lariat and allows for the snRNP components to be recycled to splice again. The two reactions that occur within each splice cycle are energy neutral and as a result, the spliceosome

relies on helicases and ATP to drive the dynamic rearrangements that are required to successfully remove an intron from a pre-mRNA transcript



(Koodathingal et al., 2010; Koodathingal and Staley, 2013; Mayas et al., 2006; Toroney, 2019; Villa and Guthrie, 2005; Xu and Query, 2007; Yang et al., 2013).

Figure 1.3: Splicing Cycle

To remove an intron, a given pre-mRNA (top right-hand side of cycle) must undergo various rearrangements that are driven by the engagement of many proteins and 5 snRNAs to allow for the lariat to form and the expulsion of an mRNA. In this cycle, the intron is shown in a bold black line with the branch adenosine present as an A and is flanked by 2 exons. The 5 snRNPs are depicted as colored circles: U1 in red; U2 in blue; U4 in yellow; U5 in green; and U6 in orange. The known human intermediate spliceosomal complexes are depicted as well as the helicases (in blue) that are known to drive essential rearrangements during splicing.

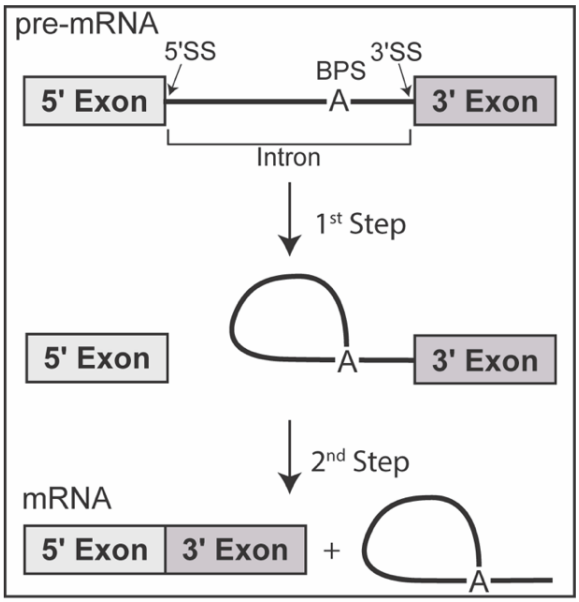


Figure 1.4: Splicing Chemistry
Splicing of a model pre-mRNA through 2 well-characterized S_N2 transesterification reactions (depicted by black arrows) via the catalytic branch adenosine within the BPS.

The U2 snRNP and snRNA

The snRNPs of the spliceosome consist of a single snRNA and many different proteins, all with various properties and roles. The U2 snRNP is essential as the U2 snRNA engages directly with the intron to form the branch helix and U2 snRNP-associated proteins encase the branch; in essence the U2 snRNP defines the 5' and 3' splice sites or the boundaries of the intron to be spliced (Kesarwani et al., 2017). The U2 snRNP itself contains the U2 snRNA as well as ten 12S core proteins and roughly ten 17S proteins in addition to many U2 snRNP-associated proteins. The components change as the splice cycle progresses with for instance the SF3B components leaving the spliceosome after the catalysis (Bertram et al., 2017). In addition, mutations in components of U2

snRNP are associated with disease such as SF3B1 K700E which is commonly associated with several cancer types (Hautin et al., 2020).

Not only does the U2 snRNP undergo vast changes throughout the splicing cycle, the U2 snRNA itself is a highly dynamic structure that requires rearrangements to fulfill its function. Consisting of multiple stem loops and single stranded regions that are constantly opening and closing, the 188-nucleotide snRNA is continually restructured during the splicing process, specifically at the 5' end (Ast et al., 2001) (Fig. 1.5). For the spliceosome to engage with an intron, the snRNAs must make vital connections. Specifically, the U2 snRNA is responsible for directly interacting with the branch point sequence of the intron to form the branch helix and current spliceosome structures have shown the vast rearrangements the U2 snRNA undergoes to accommodate its incorporation onto a given intron.

Much of our current understanding of the structure of the U2 snRNA has come from studying snRNA dynamics in yeast. Previously, a structure called the branch interacting stem loop (BSL) was found to be present within the snRNA prior to interaction with the intron (Perriman and Ares, 2010). In addition, mutations that stabilize or de-stabilize the BSL were shown to change the efficacy of splicing thereby showing not only the importance of this structure but also the required ability for easy rearrangement (Hilliker et al., 2007; Perriman and Ares, 2007). Until recently, the BSL was predicted to exist in humans however had not been confirmed as mutational analyses in human cells are very difficult unlike in yeast. However, a recent cryo-EM structure from the Luhrmann and Stark labs confirmed the existence of this U2 snRNA structure in humans

(Zhang et al., 2020), therefore confirming that like in yeast will need to be unwound to allow the U2 snRNA to interact with an intron and form the branch helix. In addition to interacting directly with the intron, the U2 snRNA will also engage with other snRNAs such as the 3' end of U6, throughout the splice cycle which will require further rearrangements of the 5' end (Ast et al., 2001; Wu and Manley, 1989).

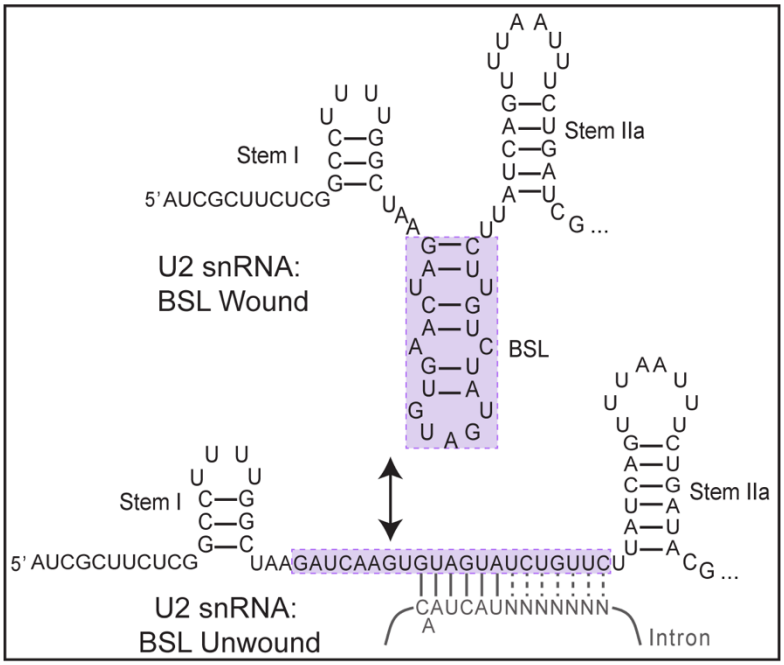


Figure 1.5: U2 snRNA Dynamics

Two known U2 snRNA conformations at the 5' end with the BSL wound and unwound (highlighted in purple). The unwinding of the BSL allows for engagement of the U2 snRNA with the intron and the eventual formation of the branch helix.

Early Spliceosome Assembly: The Transition from E- to A-Complex

The spliceosome assembles in a stepwise fashion, incorporating more snRNPs over time to allow for formation of more complicated complexes and eventually splicing chemistry to occur. E-complex is the first spliceosomal complex to assemble in which the U1 snRNP binds to the 5' splice site in addition

to some spliceosome-associated proteins binding the presumptive branch point sequence and polypyrimidine tract (SF1 and U2AF, respectively) (Berglund et al., 1998). A-complex is the second spliceosomal complex to form and includes both the U1 snRNP and U2 snRNP bound to the intron. However, the competent incorporation of the U2 snRNP onto the intron includes the formation of the branch helix, which ensures that the branch adenosine is placed near the active site for downstream splicing chemistry. Yet the steps that allow the U2 snRNP to engage and incorporate onto an intron and the overall progression from E- to A-complex has long eluded scientific clarity.

Previous work has shown that the U2 snRNP must first undergo internal rearrangements prior to intron engagement. The U2 snRNP will initially interact with the branch point sequence prior to fully engaging with the intron or the formation of the branch helix (Plaschka et al., 2017; Rauhut et al., 2016; Yan et al., 2016). To facilitate this interaction, spliceosome-associated proteins such as SF1 and the U2AF complex (U2AF65 and U2AF35) are bound to the 3' end of the intron and are thought to help recruit the U2 snRNP to the correct branch sequence. The U2 snRNP will make some initial interactions with these proteins which will result in the eventual ejection of SF1 and the U2AF complex and the initial engagement of the U2 snRNP with the intron, however the exact steps remain unclear. Additional internal rearrangements within the U2 snRNP, such as the unwinding of the BSL, occur thereby allowing competent interaction with the branch point sequence and formation of the branch helix (Fig. 1.5) (Fica and Nagai, 2017; Nguyen et al., 2016).

NTP-Dependent Steps

For the spliceosome to progress to the next complex and conformation, energy (nucleoside triphosphate, NTPs) must be utilized to drive the necessary rearrangements (Fig. 1.3). Specifically, both ATP and GTP have been found to promote complex formation and rearrangements of the spliceosome at different steps throughout the splicing cycle. For instance, for the U2 snRNP to incorporate and competently engage an intron, it must undergo several internal rearrangements that are driven by enzymes that utilize NTPs for their modality of action.

Currently, it is known that there are several NTP-dependent rearrangements that occur during the transition from E- to A-complex, such as the unwinding of the branch stem loop (BSL), a potential movement of the upstream region of the intron to allow for an initial engagement with the snRNP and the ejection of certain spliceosome-associated proteins that are known to engage essential intronic landmarks, thereby assisting in the placement of downstream factors.

In addition, prior work to this dissertation has suggested there may be more than one NTP-dependent step during the transition from E- to A-complex. For years, data from various labs in the splicing community has pointed at this fact, however unlocking how many steps are present during this transition and the roles that they play, remain unknown. Yet at the beginning of this dissertation, we knew several key facts that were found utilizing several transcribed model introns (Fig. 1.6) that when assembled upon indicated various ATP-dependent steps as well as hinted at critical moments in the transition from

E- to A-complex in addition to the usage of spliceosome inhibitors (Das and Reed, 1999; Das et al., 2000).

One piece of evidence for more than one ATP-dependent step during the E- to A-complex transition can be observed if we assemble splicing complexes in the presence of the splicing inhibitor, Spliceostatin A (SSA). Roybal and others showed that these complexes assembled in the presence of drug still form in an ATP-dependent manner (Fig. 1.7A) (Corrionero et al., 2011; Roybal and Jurica, 2010). In addition, the Valcarcel group showed that this drug stalled complex does not have a formed branch helix, and a recent structure from the Pena group uncovered that the extended helix upstream of branch is partially formed in the presence of drug (Corrionero et al., 2011; Cretu et al., 2021). Further, the Jurica Lab has shown that even with this partial formation of the extended helix, if we heparin challenge the drug stalled complexes, they are highly unstable when compared to a non-drug stalled A-complex.

Another piece of evidence we have for multiple ATP-dependent steps during E- to A-complex transition was uncovered by Charles Query when building early spliceosome complexes on small intronic substrates (Fig. 1.6) (Newnham, 2001; Query, 1997). When we assemble upon an $A^{3'}$ substrate which contains a branch point sequence, polypyrimidine tract and 15 nucleotides upstream of the branch, an A-like complex called A3' forms in an ATP-dependent manner (Fig.1.7B). In addition to the $A^{3'}$ substrate, the A^{min} substrate provided very interesting evidence for multiple ATP-dependent steps. Unlike $A^{3'}$, the A^{min} substrate lacks the 15-nucleotide upstream sequence and only contains a branch point sequence and a polypyrimidine tract (Query, 1997). In addition, it also

assembles in an ATP-independent manner, forming the A^{min} complex in the absence of ATP. Due to A^{min}'s ability to assemble in an ATP-independent manner, it suggests that it can bypass an ATP-dependent step that other substrates cannot. Further, upon addition of ATP, the A^{min} complex dissipates over time and the lower band that Charles named * (star) increases (Fig. 1.7C. A-like), which suggests that these complexes may be the target of some sort of disassembly pathway (Newnham, 2001).

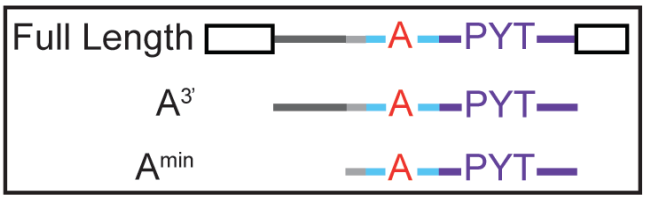


Figure 1.6: RNA Substrates Utilized in this Dissertation

Body labeled RNA substrates utilized throughout this dissertation. The full-length RNA substrate contains two exons, both 5' and 3' SS as well as a BPS and PYT and assembles and splices in an ATP-dependent manner. The $A^{3'}$ substrate is a truncated substrate that contains a BPS and PYT as well as 15 nucleotides upstream of the branch that confers assembly in an ATP-dependent manner. This substrate cannot assemble past an A-like or $A^{3'}$-complex. The A^{min} substrate only contains a BPS and a PYT, lacking the 15 upstream nucleotides and as a result, assembles in an ATP-independent manner and just like the $A^{3'}$ substrate cannot assemble past an A^{min} complex.

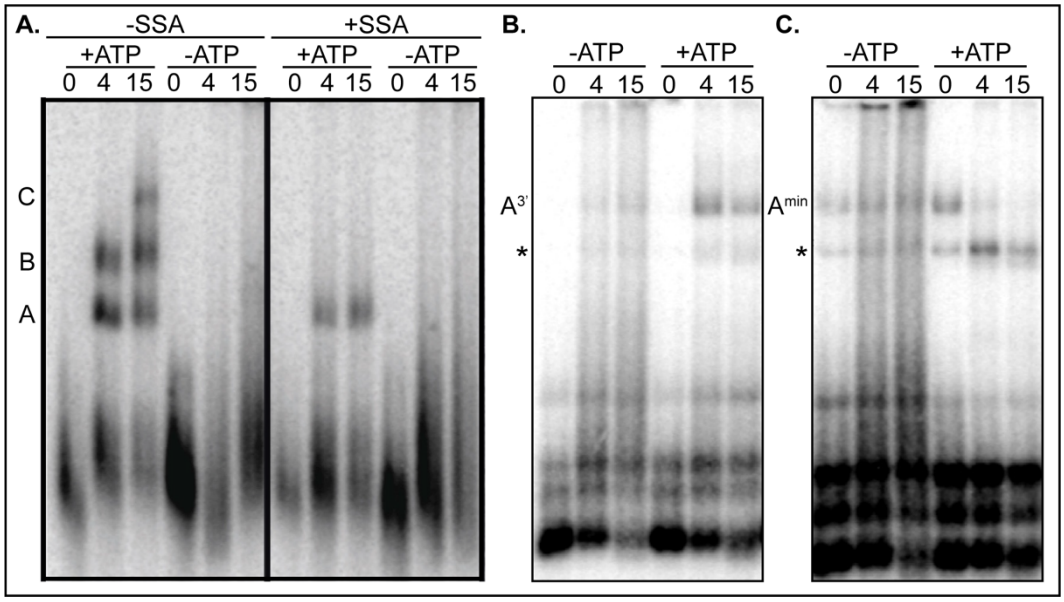


Figure 1.7: **SSA-stalled complexes assemble in an ATP-dependent manner whereas assembly on truncated substrates vary in ATP-dependence**

A. Native gel analysis of a full-length body labeled pre-mRNA substrate assembled in +/- ATP conditions in addition to +/- conditions of splicing inhibitor, Spliceostatin A (SSA). B. Native gel analysis of an $A^{3'}$ substrate over time at the indicated timepoints (minutes). C. Native gel analysis of an A^{min} substrate over time at the indicated timepoints (minutes). Gel in Panel A by Yewande Alabi.

RNA Helicases

RNA helicases are enzymes that are responsible for many important processes such as the unwinding of RNA, and the removal of RNA-binding proteins (Fig. 1.8). Within the human genome, there are over 70 known RNA helicases that are involved across a variety of cellular processes such as ribosome biogenesis, transcription, translation, and RNA splicing. These enzymes can fit into a variety of classes with most RNA helicases falling into the SF2 super family and can be split further into five subclasses: DExH-box, DExD-box, Ski2-like, RIG-I-like and NS3/NPH-II (only found in viruses) (Leitão et al., 2015).

Initially these enzymes were characterized as RNA duplex unwinders, these essential motor proteins tend to move in a single direction, usually 3' to 5' (if DExH-box) along the phosphodiester backbone of a single RNA strand and unwind double stranded RNA-RNA and RNA-DNA interactions as well as remove RNA binding proteins from strands of RNA, all of which requires the energy derived from NTP hydrolysis (Hamann et al., 2019; Mallam et al., 2012; Sengoku et al., 2006; Tauchert et al., 2017; Yang et al., 2007; Yang and Jankowsky, 2006). Further, many helicases are known to be multi-faceted in their roles and have been attributed to functioning in a variety of cellular processes yet only a few have been studied extensively.

For helicases to function they rely on NTP hydrolysis. ATP (adenosine triphosphate) is an essential energy source for many processes as well as GTP (guanosine triphosphate). The cell and its components utilize ATP binding and hydrolysis to ADP (adenosine diphosphate) to generate free phosphate which drives cellular processes. This is true for the other nucleotides as well. We can utilize nucleotide analogs to study whether certain processes require only binding versus hydrolysis of the nucleotide to perform a certain function. One such analog that will be utilized in this dissertation is AMP-PNP (adenylyl imidodiphosphate), which can recapitulate ATP-binding however cannot perform hydrolysis (Fig. 1.9). When studying helicase activity, the use of NTP-analogs becomes very key as they can help to determine whether binding is sufficient for a given transition or if hydrolysis is required. To that end, I will utilize ATP analogs in understanding the helicases of the spliceosome.

Currently, there are eight spliceosome associated helicases that are responsible for driving the essential rearrangements required for the spliceosome to successfully engage and remove an intron. The associated enzymes occur in two flavors: DExD-box helicases and DExH-box helicases, which are named for the presence of a DExD (Asp-Glu-Ala-Asp) or DExH (Asp-Glu-Ala-His) sequence in the second motif of the protein (Fig. 1.10) (Gilman et al. 2017; Jankowsky 2011). In addition, the modality of these two helicase subclasses is dictated by the presence or absence of a Q-motif in the N-terminus of the RecA1 domain (Tanner et al., 2003; Wei et al., 2021). DExD-box helicases possess this motif and as a result directly coordinate the adenosine base of ATP, resulting in the forced usage of ATP binding and hydrolysis as a mode of action (Tanner et al., 2003). DExH-box helicases lack this motif and as a result can bind and hydrolyze any NTP with a preference for ATP or GTP (Fig. 1.10) (He et al. 2010; Walbott et al. 2010).

In addition to the lack of Q-motif, unlike DExD-box helicases, DExH-box helicases are believed to only move in the 3’ to 5’ direction and are processive in activity (Jankowsky, 2011; Pyle, 2008) These helicases bind a particular NTP which allows for stacking of several conserved Arginine and Phenylalanine residues between the two RecA domains which keeps the helicase in an open conformation ready for RNA-loading (Tauchert et al., 2017). For the helicase to close, four nucleotides of the RNA can be accommodated within the RNA binding groove that is formed between the hook turn of the RecA1 domain and the hook loop of the RecA2 domain thereby allowing 3’ RNA overhangs to be bound and

thereby triggering NTP hydrolysis to occur (Bohnsack et al., 2022; Hamann et al., 2019).

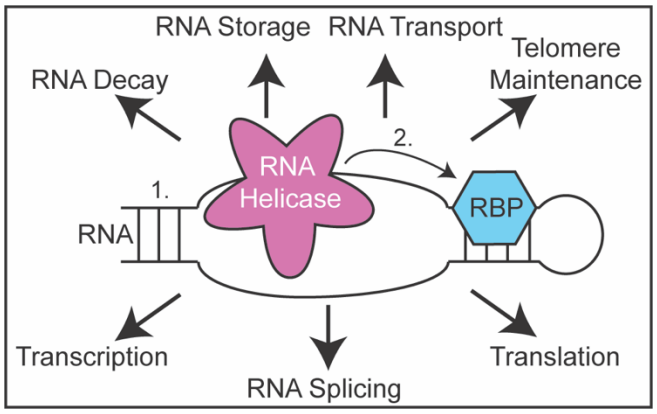


Figure 1.8: Roles of RNA Helicases

RNA helicases are responsible for a variety of cellular processes including those listed and more as well as other processes such as: 1. Unwinding of RNA-RNA duplexes and 2. complex regulation by removal of RNA binding proteins (RBP) in addition to being involved in a multitude of cellular processes that require NTP-driven rearrangements.

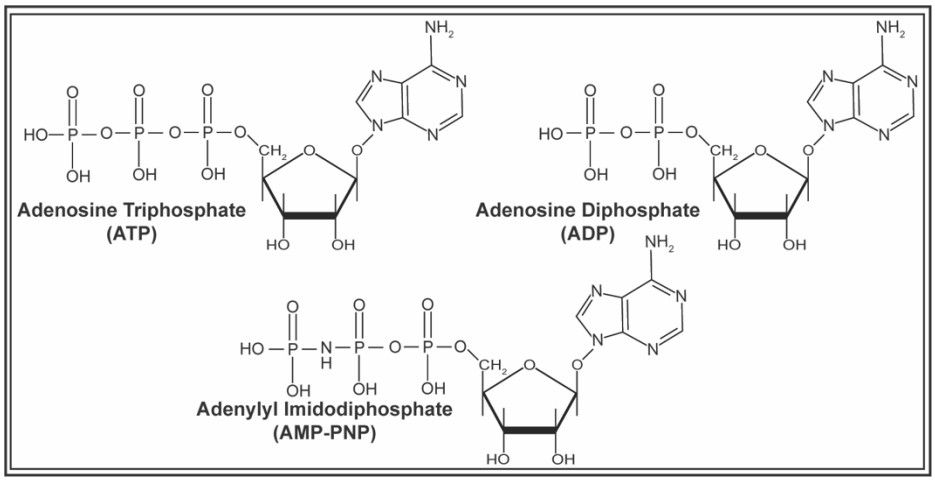


Figure 1.9: Structure of ATP and ATP Analogs

Structures of the nitrogenous bases (top row) followed by the structures of the two possible sugar molecules as well as a phosphate backbone (middle row) and finally the nucleic acid structure of ATP, ADP, and ATP-analog, AMP-PNP.

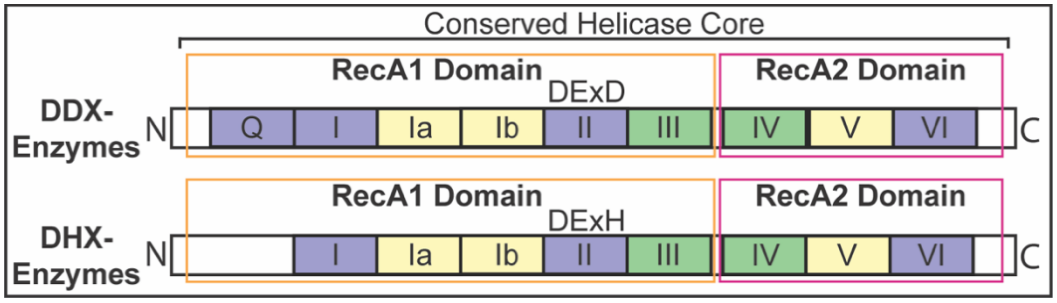


Figure 1.10: Conserved Motifs in DDX- and DHX-Helicases

The conserved helicase core contains several motifs that allow for essential interactions of the helicase: the purple bind and hydrolyze NTPs, the yellow binds double stranded RNA, specifically the substrate of interaction (i.e., whatever is being unwound) and green regulate the binding of RNA to a given NTP.

Helicases of the Early Spliceosome

In early spliceosome assembly, specifically the transition from E- to A-complex, there are three spliceosome-associated helicases: DHX15, DDX39B and DDX46 (Prp43, Uap56/Sub2, and Prp5 respectively, in yeast) (Fig. 1.3). Prior to the work presented in this dissertation, both DDX39B and DDX46 had been shown to play essential roles in A-complex formation, with DDX46 known to unwind the BSL in yeast (Perriman and Ares, 2010) and with the recent observation of this structure in humans, it is be postulated to play a similar role (Zhang et al., 2020). In addition, DDX39B has been shown to also be essential for incorporation of the U2 snRNP and is known to interact with U2AF which binds to the PYT prior to engagement of the U2 snRNP onto an intron (Fleckner et al., 1997). As expected, it has been shown that both DDX39B and DDX46 utilize ATP (O'Day et al., 1996; Shen et al., 2007) whereas DHX15 can utilize either GTP or ATP (Robert-Paganin et al., 2017). DHX15 is most well characterized for its role in the disassembly of the intron lariat spliceosome (ILS) at the end of each splice cycle (Arenas and Abelson, 1997; Burke et al., 2018;

Martin et al., 2002; Toroney, 2019; Tsai et al., 2005). However, previous studies have associated DHX15 with early spliceosomal complexes, however its role has yet to be defined and, in some cases, written off as a contaminant source (Agafonov et al., 2011; Bessonov et al., 2010; Bessonov et al., 2008).

The Known Roles of DHX15

DHX15 is a highly abundant RNA helicase involved in a multitude of cellular processes. It has been associated with ribosome biogenesis, specifically biogenesis of the 40S subunit (Klinge and Woolford, 2019).in addition to the breakdown of the ILS after each splice cycle, however studies disagree whether it interacts directly with the U6 snRNA or the intron (Fourmann et al., 2016; Toroney, 2019).The work in this dissertation as well as recent publications from Qing Feng et al. and Irene Beusch et al. has suggested another role for DHX15 in the early spliceosome in quality control (Beusch et al., 2022; Feng et al., 2022; Maul-Newby et al., 2022)

DHX15 has been implicated in being involved in the process of adding the 5’ cap to pre-mRNAs. DHX15 has also been found in Cajal bodies where it is speculated to have a role in regulating tri-sRNP assembly (Chen et al., 2017). And finally, DHX15 has been associated with having a role in viral infections where it has been identified as signaler for an immune response in the NF-kB/MAPK pathways (Memet et al., 2017; Mosallanejad et al., 2014; Pattabhi et al., 2019).

The vast difference in roles that DHX15 plays within the cell indicates what a vital protein it is (Fig. 1.11) yet its regulation and systematic shuffling to a given process is known to be tightly controlled yet still under heavy investigation.

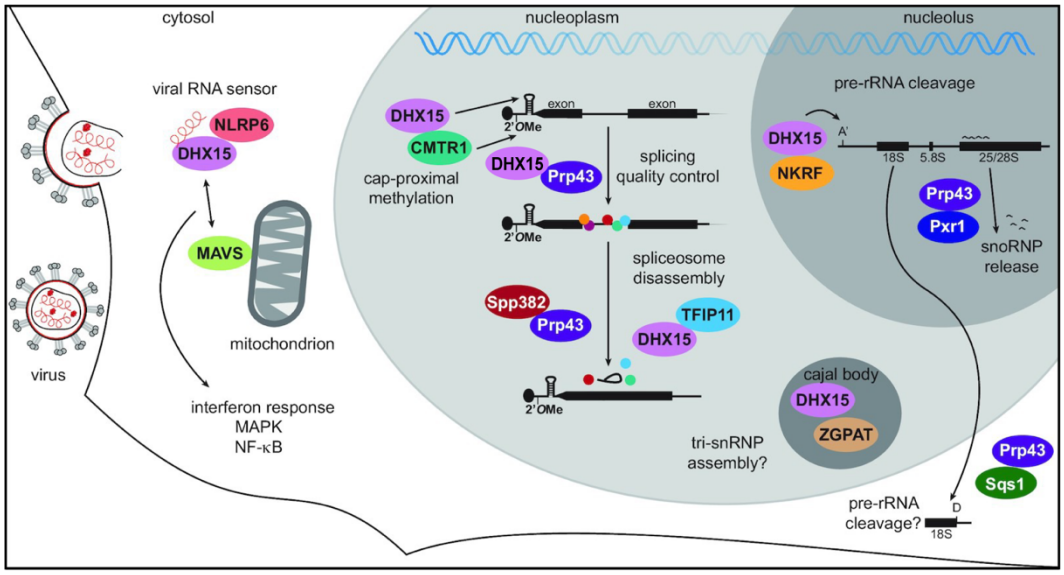


Figure 1.11 – Cellular Pathways involving DHX15/Prp43

Cellular schematic depicting the confirmed and/or predicted cellular pathways in which DHX15/Prp43 play a role. The focus of this dissertation will be the role of DHX15 in splicing quality control. Figure adapted from Bohnsack et al., 2022, © Oxford University Press

The role of G-Patch Proteins in DHX15 Regulation

For DHX15 to perform a specific function, it binds with a G-patch binding protein. The binding of a specific G-patch is what dictates DHX15's function within the cell. For instance, in the case of the ILS, DHX15 directly interacts with TFIP11, a G-patch containing protein and the human homolog of yeasts, NTR1 (Wen et al., 2008). The binding of TFIP11 to DHX15 directs DHX15's helical activity to the disassembly of the ILS and the binding of the G-patch itself, provides rigidity of the two RecA domains within DHX15 allowing for competent binding to the RNA (Bohnsack et al., 2021).

During early spliceosome assembly there are five G-patch associated proteins: CHERP, SUGP1, RBM5, RBM10 and RBM17 (Agafonov et al., 2011). Of the five, it is known that DHX15 interacts with RBM17 during alternative

splicing (De Maio et al., 2018). However, it has remained elusive as to which G-patch interacts with DHX15 during early spliceosome assembly.

Recently, several groups have stated that SUGP1 is the G-patch binding co-factor for DHX15 mediated disassembly of incompetent complexes during early spliceosome assembly (Beusch et al., 2022; Feng et al., 2022; Zhang et al., 2022). However, it has yet to be determined if this G-patch is responsible across all cell types or is cell type specific as its overall expression is highly variable from cell type to cell type. For instance, SUGP1 is lowly expressed in HeLa cells and highly expressed in HEK293T cells and therefore further work is still required to confirm these findings.

Contributions of this Dissertation

Prior to the work in this dissertation, the idea that there are multiple ATP-dependent steps during A-complex formation had been alluded to, yet it had not been confirmed in terms of how many and what they resulted in (Corrionero et al., 2011; Roybal and Jurica, 2010). In addition, a role for DHX15 had not been found in early complex formation nor had the idea of a disassembly step been associated with spliceosomes besides the ILS. However, due to DHX15 association with early complexes and the spliceosome relies on helicases for remodeling, finding a new possible role was intriguing.

I first asked which enzymes are involved during the E- to A-complex transition to parse apart how many ATP-dependent rearrangements may be present. Due to differences in helicase type, I utilized different NTPs in conjunction with assembly and splicing assays to differentiate and determine which may be at play. The results of those assays are described in **Chapter 2**

and resulted in the finding that one of the transitional steps from E- to A-complex is NTP-driven instead of ATP-driven, but also uncovered a new role for DHX15 (Maul-Newby et al., 2022).

In addition to my work on DHX15, I worked to understand why certain cell lines yield nuclear extracts suitable for *in vitro* assays, yet others do not. In addition to seeing differences between different cell lines, differences can occur from parental to derived cell lines as well as inconsistency between preps. As time has progressed science has relied on a multitude of cell lines to answer questions, yet the scientific community relies solely on HeLa cells for human *in vitro* splicing assays. And due to the lack of variety, we may be missing key details in answering many of the unanswered splicing questions. To that end, in **Chapter 3** I describe protocols I adapted, developed, and utilized to make nuclear extracts from cell lines other than HeLa. However, I found that in HEK293T cells, a small molecule inhibitor is most likely co-purified and concentrated during the process of making nuclear extract thereby preventing activity. In addition, I observed that some active HeLa parental lines yield inactive derived lines. To understand what is leading to the inactivity, I investigated snRNA distribution across a gradient to determine if there were differences. Initial results suggest mis-migration of snRNAs and specifically the U2 snRNA thereby preventing activity in -/+ ATP conditions. Together, I have uncovered that there may be a multitude of reasons that lead to inactivity but more than likely the inconsistencies we see across cell lines and from parental to derived may be due to loss of snRNAs during purification.

Finally, structural studies of the spliceosome have traditionally been done in one of two ways: purification of over-expression constructs or the growth of large amounts of cells (hundreds of plates) to yield enough material for structural studies. However, a recent master's student, John Kim, V5-tagged a 12S spliceosome component, SNRPB2 allowing for us to biochemically pull on the spliceosome in ways that we previously were unable to (Kim, 2019; Maul-Newby et al., 2022). In **Chapter 4**, I describe a purification scheme I developed that combines three separate purifications including a final IP against V5-SNRPB2 which results in purified complexes for analysis. Unlike previous purification schemes, this one is unique in that we crosslink, separate by size, RNA substrate as well as pulling directly on the snRNP. Taken together, these steps allow for small scale initial purifications yet yield material in high enough quantities for downstream analyses.

In summary, my work during my graduate studies has uncovered a previously unknown NTP-required assembly step that may be driven by DHX15 or a currently unknown DHX-helicase. In working to decipher the ATP-dependent steps associated with early spliceosome assembly, I also discovered a previously unknown DHX15-mediated disassembly step during assembly that is a potential method of quality control. In addition, I have also provided new protocols to the splicing community such as purification of early spliceosomal complexes and potential reasonings as to the inconsistencies in making nuclear extracts from HeLa cells and how we may broaden our choice of cell lines for *in vitro* assays. Together, both sets of protocols can be utilized by the scientific community to further their studies into the mechanism of splicing.

References

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

Arenas, J.E., Abelson, J.N., 1997. Prp43: An RNA helicase-like factor involved in spliceosome disassembly. Proc Natl Acad Sci U S A 94, 11798-11802.

Ast, G., Pavelitz, T., Weiner, A.M., 2001. Sequences upstream of the branch site are required to form helix II between U2 and U6 snRNA in a trans-splicing reaction. Nucleic Acids Res 29, 1741-1749.

Berget, S.M., Moore, C., Sharp, P.A., 1977. Spliced segments at the 5' terminus of adenovirus 2 late mRNA. Proc Natl Acad Sci U S A 74, 3171-3175.

Berglund, J.A., Abovich, N., Rosbash, M., 1998. A cooperative interaction between U2AF65 and mBBP/SF1 facilitates branchpoint region recognition. Genes Dev 12, 858-867.

Bertram, K., Agafonov, D.E., Liu, W.T., Dybkov, O., Will, C.L., Hartmuth, K., Urlaub, H., Kastner, B., Stark, H., Luhrmann, R., 2017. Cryo-EM structure of a human spliceosome activated for step 2 of splicing. Nature 542, 318-323.

Bessonov, S., Anokhina, M., Krasauskas, A., Golas, M.M., Sander, B., Will, C.L., Urlaub, H., Stark, H., Luhrmann, R., 2010. Characterization of purified human Bact spliceosomal complexes reveals compositional and morphological changes during spliceosome activation and first step catalysis. RNA 16, 2384-2403.

Bessonov, S., Anokhina, M., Will, C.L., Urlaub, H., Luhrmann, R., 2008. Isolation of an active step I spliceosome and composition of its RNP core. Nature 452, 846-850.

Beusch, I., Rao, B., Studer, M., Luhovska, T., Šukytė, V., Lei, S., Oses-Prieto, J., SeGraves, E., Burlingame, A., Jonas, S., Madhani, H.D., 2022. Targeted high throughput mutagenesis of the human spliceosome reveals its in vivo operating principles. bioRxiv, 2022.2011.2013.516350.

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

Bohnsack, K.E., Kanwal, N., Bohnsack, M.T., 2022. Prp43/DHX15 exemplify RNA helicase multifunctionality in the gene expression network. Nucleic Acids Research 50, 9012-9022.

Burke, J.E., Longhurst, A.D., Merkurjev, D., Sales-Lee, J., Rao, B., Moresco, J.J., Yates, J.R., 3rd, Li, J.J., Madhani, H.D., 2018. Spliceosome Profiling Visualizes

Operations of a Dynamic RNP at Nucleotide Resolution. Cell 173, 1014-1030.e1017.

Chen, Z., Gui, B., Zhang, Y., Xie, G., Li, W., Liu, S., Xu, B., Wu, C., He, L., Yang, J., Yi, X., Yang, X., Sun, L., Liang, J., Shang, Y., 2017. Identification of a 35S U4/U6.U5 tri-small nuclear ribonucleoprotein (tri-snRNP) complex intermediate in spliceosome assembly. J Biol Chem 292, 18113-18128.

Chow, L.T., Gelinas, R.E., Broker, T.R., Roberts, R.J., 1977. An Amazing Sequence Arrangement at the 5' Ends of Adenovirus 2 Messenger RNA. Cell 12, 1-8.

Corrionero, A., Minana, B., Valcarcel, J., 2011. Reduced fidelity of branch point recognition and alternative splicing induced by the anti-tumor drug spliceostatin A. Genes Dev 25, 445-459.

Cretu, C., Gee, P., Liu, X., Agrawal, A., Nguyen, T., Ghosh, A., Cook, A., Jurica, M., Larsen, N., Pena, V., 2021. Structural basis of intron selection by U2 snRNP in the presence of covalent inhibitors. Nat Commun 12, 4491.

Darman, R.B., Seiler, M., Agrawal, A.A., Lim, K.H., Peng, S., Aird, D., Bailey, S.L., Bhavsar, E.B., Chan, B., Colla, S., Corson, L., Feala, J., Fekkes, P., Ichikawa, K., Keaney, G.F., Lee, L., Kumar, P., Kunii, K., MacKenzie, C., Matijevic, M., Mizui, Y., Myint, K., Park, E.S., Puyang, X., Selvaraj, A., Thomas, M.P., Tsai, J., Wang, J.Y., Warmuth, M., Yang, H., Zhu, P., Garcia-Manero, G., Furman, R.R., Yu, L., Smith, P.G., Buonamici, S., 2015. Cancer-Associated SF3B1 Hotspot Mutations Induce Cryptic 3' Splice Site Selection through Use of a Different Branch Point. Cell Rep 13, 1033-1045.

Das, R., Reed, R., 1999. Resolution of the mammalian E complex and the ATP-dependent spliceosomal complexes on native agarose mini-gels. RNA 5, 1504-1508.

Das, R., Zhou, Z., Reed, R., 2000. Functional association of U2 snRNP with the ATP-independent spliceosomal complex E. Mol Cell 5, 779-787.

De Maio, A., Yalamanchili, H.K., Adamski, C.J., Gennarino, V.A., Liu, Z., Qin, J., Jung, S.Y., Richman, R., Orr, H., Zoghbi, H.Y., 2018. RBM17 Interacts with U2SURP and CHERP to Regulate Expression and Splicing of RNA-Processing Proteins. Cell Rep 25, 726-736 e727.

Feng, Q., Krick, K., Chu, J., Burge, C.B., 2022. Splicing quality control mediated by DHX15 and its G-patch activator, SUGP1. bioRxiv, 2022.2011.2014.516533.

Fica, S.M., Nagai, K., 2017. Cryo-electron microscopy snapshots of the spliceosome: structural insights into a dynamic ribonucleoprotein machine. Nat Struct Mol Biol 24, 791-799.

Fleckner, J., Zhang, M., Valcarcel, J., Green, M.R., 1997. U2AF65 recruits a novel human DEAD box protein required for the U2 snRNP-branchpoint interaction. Genes Dev 11, 1864-1872.

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

Hamann, F., Enders, M., Ficner, R., 2019. Structural basis for RNA translocation by DEAH-box ATPases. Nucleic Acids Res 47, 4349-4362.

Hautin, M., Mornet, C., Chauveau, A., Bernard, D.G., Corcos, L., Lippert, E., 2020. Splicing Anomalies in Myeloproliferative Neoplasms: Paving the Way for New Therapeutic Venues. Cancers (Basel) 12.

Hilliker, A.K., Mefford, M.A., Staley, J.P., 2007. U2 toggles iteratively between the stem IIa and stem IIc conformations to promote pre-mRNA splicing. Genes Dev 21, 821-834.

Jankowsky, E., 2011. RNA helicases at work: binding and rearranging. Trends Biochem Sci 36, 19-29.

Jurado, A.R., Tan, D., Jiao, X., Kiledjian, M., Tong, L., 2014. Structure and function of pre-mRNA 5'-end capping quality control and 3'-end processing. Biochemistry 53, 1882-1898.

Kesarwani, A.K., Ramirez, O., Gupta, A.K., Yang, X., Murthy, T., Minella, A.C., Pillai, M.M., 2017. Cancer-associated SF3B1 mutants recognize otherwise inaccessible cryptic 3' splice sites within RNA secondary structures. Oncogene 36, 1123-1133.

Kim, J., 2019. A Pipeline for Tagging snRNP Associated Proteins in HeLa. , Molecular, Cell and Developmental Biology. University of California, Santa Cruz, CA.

Klinge, S., Woolford, J.L., Jr., 2019. Ribosome assembly coming into focus. Nat Rev Mol Cell Biol 20, 116-131.

Koodathingal, P., Novak, T., Piccirilli, J., Staley, J., 2010. The DEAH box ATPases Prp16 and Prp43 cooperate to proofread 5' splice site cleavage during pre-mRNA splicing. Mol Cell 39, 385-395.

Koodathingal, P., Staley, J., 2013. Splicing fidelity: DEAD/H-box ATPases as molecular clocks. RNA Biol 10, 1073-1079.

Leitão, A.L., Costa, M.C., Enguita, F.J., 2015. Unzippers, Resolvers and Sensors: A Structural and Functional Biochemistry Tale of RNA Helicases. International Journal of Molecular Sciences 16, 2269-2293.

Mallam, A., Del Campo, M., Gilman, B., Sidote, D., Lambowitz, A., 2012. Structural basis for RNA-duplex recognition and unwinding by the DEAD-box helicase Mss116p. Nature 490, 121-125.

Martin, A., Schneider, S., Schwer, B., 2002. Prp43 is an essential RNA-dependent ATPase required for release of lariat-intron from the spliceosome. J Biol Chem 277, 17743-17750.

Maul-Newby, H., Amorello, A., Sharma, T., Kim, J., Modena, M., Prichard, B., Jurica, M., 2022. A Model for DHX15 Mediated Disassembly of A-Complex Spliceosomes. RNA, rna.078977.078121.

Mayas, R.M., Maita, H., Staley, J.P., 2006. Exon ligation is proofread by the DExD/H-box ATPase Prp22p. Nat Struct Mol Biol 13, 482-490.

Memet, I., Doebele, C., Sloan, K.E., Bohnsack, M.T., 2017. The G-patch protein NF-kappaB-repressing factor mediates the recruitment of the exonuclease XRN2 and activation of the RNA helicase DHX15 in human ribosome biogenesis. Nucleic Acids Res 45, 5359-5374.

Mosallanejad, K., Sekine, Y., Ishikura-Kinoshita, S., Kumagai, K., Nagano, T., Matsuzawa, A., Takeda, K., Naguro, I., Ichijo, H., 2014. The DEAH-Box RNA Helicase DHX15 Activates NF-κB and MAPK Signaling Downstream of MAVS During Antiviral Responses. Science Signaling 7, ra40-ra40.

Newnham, C.M., Query, C. C., 2001. The ATP requirement for U2 snRNP addition is linked to the pre-mRNA region 5’ to the branch site. RNA 7, 1298-1309.

Nguyen, T.H., Galej, W.P., Fica, S.M., Lin, P.C., Newman, A.J., Nagai, K., 2016. CryoEM structures of two spliceosomal complexes: starter and dessert at the spliceosome feast. Curr Opin Struct Biol 36, 48-57.

O'Day, C., Dalbadie-McFarland, G., Abelson, J., 1996. The Saccharomyces cerevisiae Prp5 protein has RNA-dependent ATPase activity with specificity for U2 small nuclear RNA. J Biol Chem 271, 33261-33267.

Pattabhi, S., Knoll, M.L., Gale M, Jr., Loo, Y.M., 2019. DHX15 Is a Coreceptor for RLR Signaling That Promotes Antiviral Defense Against RNA Virus Infection. J Interferon Cytokine Res 39, 331-346.

Perriman, R., Ares, M., Jr., 2010. Invariant U2 snRNA nucleotides form a stem loop to recognize the intron early in splicing. Mol Cell 38, 416-427.

Perriman, R.J., Ares, M., Jr., 2007. Rearrangement of competing U2 RNA helices within the spliceosome promotes multiple steps in splicing. Genes Dev 21, 811-820.

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

Pyle, A.M., 2008. Translocation and unwinding mechanisms of RNA and DNA helicases. Annu Rev Biophys 37, 317-336.

Query, C.C., McCaw, P. S., and Sharp, P. A., 1997. A Minimal Spliceosomal Complex A Recognizes the Branch Site and Polypyrimidine Tract. Molecular and Cellular Biology 17, 2944-2953.

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

Robert-Paganin, J., Halladjian, M., Blaud, M., Lebaron, S., Delbos, L., Chardon, F., Capeyrou, R., Humbert, O., Henry, Y., Henras, A., Réty, S., Leulliot, N., 2017. Functional link between DEAH/RHA helicase Prp43 activation and ATP base binding. Nucleic Acids Res 45, 1539-1552.

Roybal, G.A., Jurica, M.S., 2010. Spliceostatin A inhibits spliceosome assembly subsequent to prespliceosome formation. Nucleic Acids Res 38, 6664-6672.

Sengoku, T., Nureki, O., Nakamura, A., Kobayashi, S., Yokoyama, S., 2006. Structural basis for RNA unwinding by the DEAD-box protein Drosophila Vasa. Cell 125, 287-300.

Shen, J., Zhang, L., Zhao, R., 2007. Biochemical characterization of the ATPase and helicase activity of UAP56, an essential pre-mRNA splicing and mRNA export factor. J Biol Chem 282, 22544-22550.

Tanner, N., Cordin, O., Banroques, J., Doère, M., Linder, P., 2003. The Q motif: a newly identified motif in DEAD box helicases may regulate ATP binding and hydrolysis. Mol Cell 11, 127-138.

Tauchert, M.J., Fourmann, J.B., Luhrmann, R., Ficner, R., 2017. Structural insights into the mechanism of the DEAH-box RNA helicase Prp43. eLife 6.

Toroney, R., Nielsen, K. H., & Staley, J. P., 2019. Termination of pre-mRNA splicing requires that the ATPase and RNA unwindase Prp43p acts on the catalytic snRNA U6. Genes and Development 33, 1-20.

Tsai, R.T., Fu, R.H., Yeh, F.L., Tseng, C.K., Lin, Y.C., Huang, Y.H., Cheng, S.C., 2005. Spliceosome disassembly catalyzed by Prp43 and its associated components Ntr1 and Ntr2. Genes Dev 19, 2991-3003.

Villa, T., Guthrie, C., 2005. The Isy1p component of the NineTeen complex interacts with the ATPase Prp16p to regulate the fidelity of pre-mRNA splicing. Genes Dev 19, 1894-1904.

Wahl, M.C., Will, C.L., Luhrmann, R., 2009. The spliceosome: design principles of a dynamic RNP machine. Cell 136, 701-718.

Wei, Q., Geng, J., Chen, Y., Lin, H., Wang, J., Fang, Z., Wang, F., Zhang, Z., 2021. Structure and function of DEAH-box helicase 32 and its role in cancer. Oncol Lett 21, 382.

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

Wu, J., Manley, J., 1989. Mammalian pre-mRNA branch site selection by U2 snRNP involves base pairing. Genes Dev 3, 1553-1561.

Xu, Y.Z., Query, C.C., 2007. Competition between the ATPase Prp5 and branch region-U2 snRNA pairing modulates the fidelity of spliceosome assembly. Mol Cell 28, 838-849.

Yan, C., Wan, R., Bai, R., Huang, G., Shi, Y., 2016. Structure of a yeast activated spliceosome at 3.5 A resolution. Science 353, 904-911.

Yang, F., Wang, X., Zhang, Z., Pu, J., Fan, Y., Zhou, J., Query, C., Xu, Y., 2013. Splicing proofreading at 5' splice sites by ATPase Prp28p. Nucleic Acids Res 41, 4660-4670.

Yang, Q., Del Campo, M., Lambowitz, A., Jankowsky, E., 2007. DEAD-box proteins unwind duplexes by local strand separation. Mol Cell 28, 253-263.

Yang, Q., Jankowsky, E., 2006. The DEAD-box protein Ded1 unwinds RNA duplexes by a mode distinct from translocating helicases. Nat Struct Mol Biol 13, 981-986.

Zhang, J., Huang, J., Xu, K., Xing, P., Huang, Y., Liu, Z., Tong, L., Manley, J.L., 2022. DHX15 is involved in SUGP1-mediated RNA missplicing by mutant SF3B1 in cancer. Proc Natl Acad Sci U S A 119, e2216712119.

Zhang, Z., Will, C., Bertram, K., Dybkov, O., Hartmuth, K., Agafonov, D., Hofele, R., Urlaub, H., Kastner, B., Lührmann, R., Stark, H., 2020. Molecular architecture of the human 17S U2 snRNP. Nature 583, 310-313.

Chapter 2:

A Model for DHX15 Mediated Disassembly of A-Complex Spliceosomes

Hannah M. Maul-Newby[1,2], Angela N. Amorello[1,2], Turvi Sharma[1,2], John H. Kim[1,2], Matthew S. Modena[1,2], Beth Prichard[1,2] and Melissa S. Jurica[1,2] *

[1]Department of Molecular Cell and Developmental Biology and [2]Center for Molecular Biology of RNA, University of California Santa Cruz, Santa Cruz, California, 95064, USA

Originally Published in RNA 2022

Abstract

A critical step of pre-mRNA splicing is the recruitment of U2 snRNP to the branch point sequence of an intron. U2 snRNP conformation changes extensively during branch helix formation and several RNA-dependent ATPases are implicated in the process. However, the molecular mechanisms involved remain to be fully dissected. We took advantage of the differential nucleotide triphosphates requirements for DExD/H-box enzymes to probe their contributions to *in vitro* spliceosome assembly. Both ATP and GTP hydrolysis support the formation of A-complex, indicating the activity of a DEAH-enzyme because DEAD-enzymes are selective for ATP. We immunodepleted DHX15 to assess its involvement and although splicing efficiency decreases with reduced DHX15, A-complex accumulation incongruently increases. DHX15 depletion also results in the persistence of the atypical ATP-independent interaction between U2 snRNP and a minimal substrate that is otherwise destabilized in the presence of either ATP or GTP. These results lead us to hypothesize that DHX15 plays a quality control function in U2 snRNP's engagement with an intron. In efforts to identify the RNA target of DHX15, we determined that an extended polypyrimidine tract is not necessary for disruption of the atypical interaction between U2 snRNP and the minimal substrate. We also examined U2 snRNA by RNase H digestion and identified nucleotides in the branch binding region that become accessible with both ATP and GTP hydrolysis, again implicating a DEAH-enzyme. Together, our results demonstrate that multiple ATP-dependent rearrangements are likely involved in U2 snRNP addition to the spliceosome and that DHX15 may have an expanded role in maintaining splicing fidelity.

Introduction

Pre-mRNA splicing by the spliceosome is an essential step in eukaryotic gene expression and must be highly accurate to generate functional messenger RNAs. The boundaries of an intron are initially designated by base-pairing interactions between U1 snRNA and the 5' splice site and U2 snRNA and the branch point sequence. U2 snRNA recognition of the branch point sequence takes place in the context of a small ribonucleoprotein particle (snRNP), which also contains ten core proteins, three SF3A proteins, and seven SF3B proteins. These events signal the rest of spliceosome to assemble into a catalytic entity.

A collection of RNA-dependent ATPases that drive rearrangements required during spliceosome assembly have also been linked to enforcing splice site fidelity in *S. cerevisiae* (Koodathingal et al. 2010; Koodathingal and Staley 2013; Mayas et al. 2006; Toroney et al. 2019; Villa and Guthrie 2005; Xu and Query 2007; Yang et al. 2013). Most of these enzymes, classified as either DEAD (DDX) or DEAH (DHX), are considered helicases because they can disrupt RNA base-pairing interactions, however they exhibit some mechanistic differences (Gilman et al. 2017; Jankowsky 2011). DDX-enzymes have a Q-motif that enforces a selectivity for ATP as the source of energy for RNA unwinding (Tanner et al. 2003). In contrast, DHX-enzymes lack this motif and can utilize other nucleotides, with a preference for ATP and GTP (He et al. 2010; Walbott et al. 2010).

During *in vitro* spliceosome assembly, recruitment of U2 snRNP to the branch point sequence is the first ATP-dependent step, and results in formation of a stable entity referred to as A-complex. The mechanistic basis of the

requirement for ATP is not fully understood, but striking differences between recent cryo-EM structures of U2 snRNP and A-complex imply several large-scale molecular rearrangements of interactions within the snRNP and with the intron. Several RNA-dependent ATPases are associated with U2 snRNP: three DDX-enzymes (DDX46, DDX39B, DDX42) and DHX15. The DDX-enzymes display the expected nucleotide preferences for ATP (O'Day et al. 1996; Shen et al. 2007; Uhlmann-Schiffler et al. 2006) and DHX15 is an NTPase with a preference for ATP and GTP (Robert-Paganin et al. 2017). Both DDX46 (aka Prp5) and DDX39B (aka UAP56 or Sub2) have been linked to A-complex assembly, and likely mediate some of the rearrangements. For example, DDX46 is localized near the branch stem loop (BSL) in the 17S U2 snRNP structure (Zhang et al. 2020). In yeast, BSL mutations suppress an N-terminal deletion of the DDX46 ortholog Prp5, which links it to BSL unwinding (Perriman and Ares 2010). DDX46 is also proposed to proofread the branch between U2 snRNA and the branch point sequence of an intron (Liang and Cheng 2015; Xu and Query 2007; Zhang et al. 2021). UAP56 interacts with U2AF, a factor that binds the polypyrimidine tract (PYT) downstream of the branch to help recruit U2 snRNP (Fleckner et al. 1997). DDX42 associates exclusively with a form of U2 snRNP that lacks the SF3A and SF3B complexes and is unlikely to be directly involved in spliceosome assembly (Will et al. 2002). DHX15 (aka Prp43) is best known for its role in disassembly of the intron lariat spliceosome (ILS) at the end of splicing, although studies differ on whether it interacts directly with the intron or U6 snRNA (Fourmann et al. 2016; Toroney et al. 2019). DHX15 also has a role in ribosome biogenesis, and its specificity appears to be regulated by different G-patch

cofactors that direct it to the intron lariat (Wen et al. 2008) or pre-rRNA (Memet et al. 2017). Notably, DHX15 is present in pulldowns of U2 snRNP and A-complex spliceosomes, along with G-patch proteins RBM5, RBM10, RBM17, CHERP and SUGP1 (Agafonov et al. 2011).

For human introns, the sequences that define splice sites are quite variable, especially the branch point sequence. Quality control of branch point selection is likely important to maintain cellular health as alterations in the players involved is evidenced to lead to cancer (Bonnal et al. 2020). Unsurprisingly, these include mutations in several proteins associated with U2 snRNP. For example, specific point mutations in SF3B1, a U2 snRNP protein that directly contacts the branch point sequence in early spliceosome assembly, are enriched in cells from cancers and dysplasia syndromes, particularly in hematological lineages (Dvinge et al. 2016). In addition, SF3B1 mutations result in altered branch point selection for some introns (Alsafadi et al. 2016; Darman et al. 2015; DeBoever et al. 2015; Kesarwani et al. 2017; Wang et al. 2016).

In this study, we use an *in vitro* splicing system to investigate the possibility of a role for DHX15 in mediating ATP-dependent rearrangements involved in U2 snRNP's engagement with an intron. Depletion of DHX15 results in increased A-complex formation, but lower splicing efficiency. The effect is amplified when we assemble spliceosomes on introns incapable of forming complexes competent to complete splicing, leading us to propose that DHX15 may have an additional splicing related function in quality control of branch sequence recognition. Employing a powerful affinity-tagged version of the core component SNRPB2, we also characterize changes in U2 snRNP that show

differential selectivity for ATP and GTP, suggesting involvement of different DDX- and DHX-enzymes in a growing constellation of rearrangements that may impact how the spliceosome finds introns with high fidelity.

Results

Both ATP & GTP can support spliceosome assembly and splicing

To test the hypothesis that molecular rearrangements during A-complex assembly are promoted by one or more of the RNA-dependent ATPases associated with U2 snRNP, we used nucleotide preference to differentiate between the activity of DDX- and DHX-enzymes. DDX-enzymes require ATP to disrupt RNA interactions and are often locked onto their target when bound to a non-hydrolysable ATP analog (Gilman et al. 2017). In contrast, DHX-enzymes are processive RNA helicases that can also utilize GTP (Jankowsky 2011). If a DHX-enzyme is involved in the initial rearrangement, then GTP should support A-complex formation prior to the first rearrangement that requires the activity of a DDX-enzyme. Alternatively, if ATP-binding is sufficient for an initial DDX-mediated rearrangement, then non-hydrolysable AMP-PNP will support assembly to that point.

We carried out *in vitro* splicing assays on a radiolabeled pre-mRNA substrate in the presence or absence of ATP, GTP and AMP-PNP and excluded creatine phosphate to prevent recycling of nucleotides. We used native gels to assess spliceosome assembly on the pre-mRNA and denaturing gels to measure splicing efficiency (Fig. 2.1A and B). As expected, ATP confers robust spliceosome assembly and splicing chemistry, which are both lost with no added nucleotide.

AMP-PNP does not support assembly nor splicing, leading us to conclude that binding of a target RNA by a DDX-enzyme is not sufficient to promote the first rearrangement needed for U2 snRNP to engage an intron, with the caveat that AMP-PNP binding is slower than ATP (Polach and Uhlenbeck 2002). Surprisingly, GTP promotes both a significant amount of spliceosome assembly and splicing, although less efficiently than with ATP. We considered that nucleoside-diphosphate kinase (NME1) could be active in our extracts and transfer a phosphate from GTP to any ADP that remains after extract preparation. To address this possibility, we passed nuclear extract over a G-50 size exclusion column (SEC) to remove endogenous nucleotides (Anderson and Moore 2000). Post SEC, we still observe spliceosome assembly and splicing with the addition of ATP or GTP in the nucleotide-depleted extract (Sup. Fig. 2.1). B-complex and splicing products accumulate more slowly with GTP in the nucleotide-depleted extracts relative to those cycled at 30˚C. This difference indicates that a small amount of ADP recycling over an extended incubation may contribute to a portion of the B-complex formation and splicing activity. However, GTP is clearly sufficient to support robust A-complex assembly.

We also tested nucleotide requirements of early spliceosome assembly in the context of a truncated substrate ($A^{3'}$ substrate) that recapitulates U2 snRNP interactions with an intron (Konarska and Sharp 1986). We find that both ATP and GTP addition results in formation of the $A^{3'}$-complex, but not AMP-PNP (Fig. 2.1C). Because GTP supports A- and $A^{3'}$-complex assembly, we conclude that a DHX-enzyme can promote an initial rearrangement that allows U2 snRNP to engage with an intron. We hypothesize that the higher accumulation of B-

complex with ATP reflects the contribution of a DDX-enzyme, potentially by capturing a transient rearrangement or conformation that allows further spliceosome assembly. Based on their direct association with U2 snRNP, DHX15 and DDX46 are the most-likely candidates for the two activities.

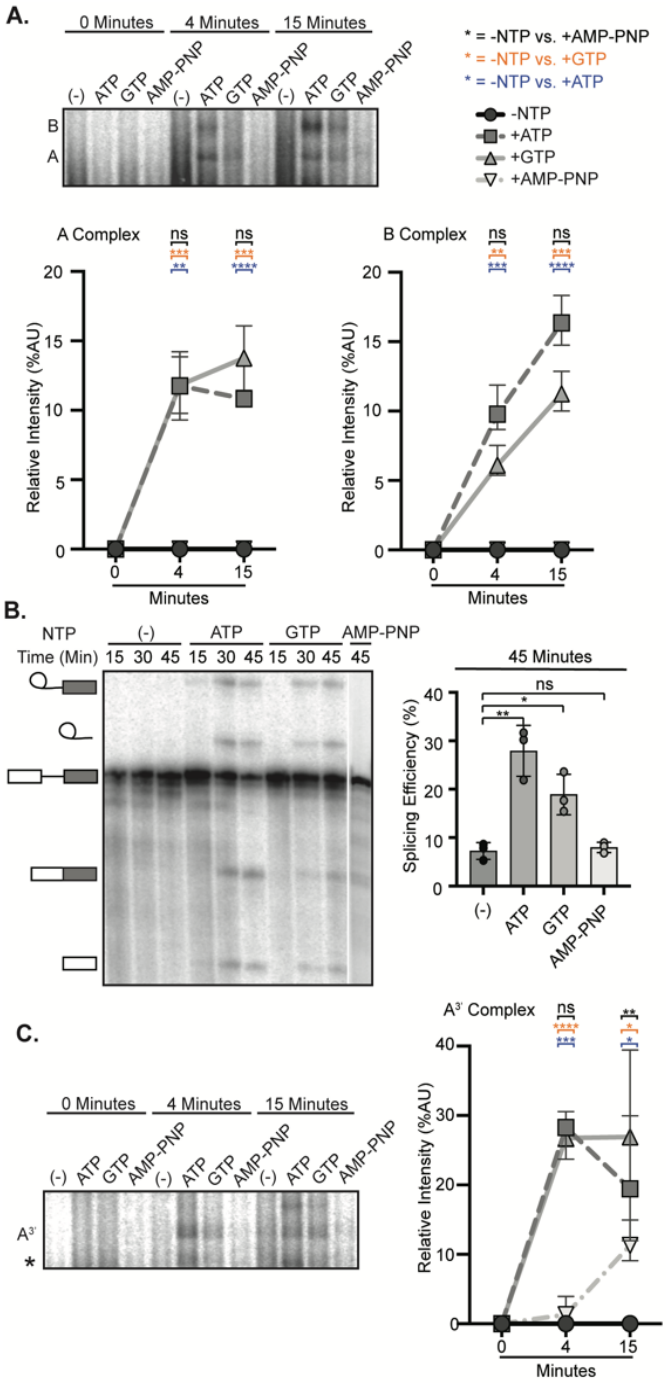


Figure 2.1: Both ATP & GTP can support spliceosome assembly and splicing

Top: Representative native gel analysis of *in vitro* spliceosome assembly with a radiolabeled full-length substrate with different added NTPs at the indicated timepoints. A- and B- complex bands positions are labeled. Bottom: Normalized band intensity relative to the entire lane for the indicated timepoints of three independent trials. Statistical differences were examined by unpaired Student's t-test with *** $p<0.001$, ** $p<0.01$, *$p<0.05$. B. Left: Representative denaturing gel analysis of radiolabeled RNA isolated at the indicated timepoints from *in vitro* splicing reactions using a full-length substrate with different added NTPs. The RNA band identities are illustrated on the left as (top to bottom): lariat intron intermediate, lariat intron, pre-mRNA substrate, mRNA, and 5' exon intermediate. Right: Splicing efficiency is measured as intensity of the mRNA band over total RNA bands and shown for the 45-minute timepoint for three independent trials. Statistical differences were examined as in (A). C. Same as (A) except that a $A^{3'}$-substrate was used for spliceosome assembly.

Depletion of DHX15 reduces splicing efficiency, but not spliceosome assembly

To determine whether DHX15 has a role in U2 snRNP addition to the spliceosome, we immunodepleted the protein from HeLa nuclear extract. Despite the high abundance of DHX15, 55-72% of the protein was removed in three independent immunodepletions relative to mock depletion as determined by western blot analysis (Fig. 2.2A, Sup. Fig. 2.2).

We used DHX15- and mock-depleted extracts for *in vitro* spliceosome assembly on the $A^{3'}$ substrate to focus on U2 snRNP incorporation. Contrary to the prediction that loss of DHX15 would inhibit assembly, the relative band intensity of A-complex in DHX15-depeleted extract is increased compared to the mock-depleted extract (Fig. 2.2B). We repeated the experiment with a full-length substrate to determine if the difference is due to the truncated substrate. A small but consistent increase in the relative band intensities of both A- and B- complex spliceosomes persist with DHX15-depleted extracts (Fig. 2.2C). Surprisingly, overall splicing efficiency is lower in the DHX15-depleted extracts compared to mock-depleted extract (Fig. 2.2D).

In the context of our original hypothesis, these results indicate that DHX15 may not be responsible for the initial rearrangement that allows U2 snRNP to interact with an intron, although we cannot rule out that the remaining DHX15 after immunodepletion is sufficient to carry out that role. However, the opposing increase in spliceosome assembly and decreased splicing efficiency, in tandem with the significant increase of $A^{3'}$-complexes, which cannot complete assembly, lead us to suspect that DHX15 may have a role in quality control of early spliceosome assembly. With reduced DHX15 activity, unproductive complexes accumulate and potentially sequester spliceosome components, which results in overall less splicing efficiency. Alternatively, the decrease in splicing efficiency may indirectly result from the loss of spliceosome disassembly in the ILS and/or another role for DHX15 promoting splicing chemistry after B-complex formation. As we are currently unable to add back purified DHX15, we also cannot rule out the possibility that a co-depleted factor is responsible for the observed effects.

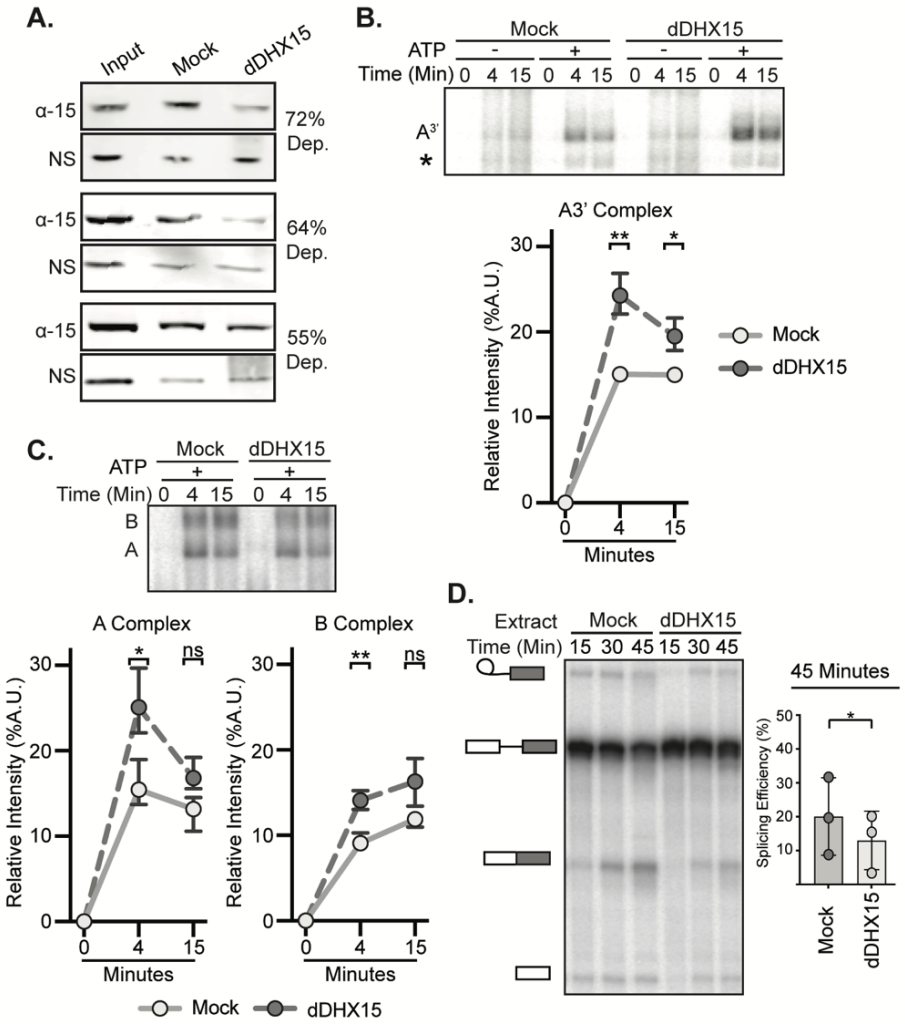


Figure 2.2: Depletion of DHX15 reduces splicing efficiency, but not spliceosome assembly

A. Western blot analysis of three independent immunodepletions of DHX15 from HeLa nuclear extract. Depletions were quantified by comparing the band intensity of DHX15 (α-15) to a non-specific band (NS), both normalized to the intensity of the entire lane. Samples are shown for untreated nuclear extract (Input), control depletion with 1X PBS or IgG (Mock) and depletion with anti-DHX15 (dDHX15). B. Top: Representative native gel analysis of *in vitro* spliceosome assembly with a radiolabeled $A^{3'}$-substrate in Mock and dDHX15 nuclear extracts at the indicated timepoints. Bottom: Normalized band intensity relative to the entire lane from three independent trials. Statistical differences were examined by paired Student's t-test with p-values as in Figure 2.1A. C. Same as (B) but with a full-length splicing substrate. D. Left: Representative denaturing gel analysis of radiolabeled RNA isolated at the indicated timepoints from *in vitro* splicing reactions using Mock or dDHX15 nuclear extracts. The RNA band identities are as described in Fig. 2.1B. Right: Splicing efficiency determined as in Fig. 2.1B.

Reduction of DHX15 stabilizes the ATP-independent interaction between U2 snRNP and minimal intron

Query *et al.* showed a minimal RNA (A^{min} substrate) containing only a branch point sequence followed by a PYT interacts with U2 snRNP in the absence of ATP to form the A^{min}-complex (Query et al. 1997). Notably, in the presence of ATP, the A^{min}-complex is destabilized by an unknown entity (Newnham and Query 2001). To determine whether DHX15 could be responsible, we incubated the A^{min} substrate in both DHX15-depleted and mock-depleted HeLa nuclear extract with and without ATP. In the absence of ATP, the expected A^{min}-complex band forms in both extracts (Fig. 2.3A). Addition of ATP to the mock-depleted extract results in the loss of most A^{min}-complex and the appearance of a faster migrating complex of unknown composition (*) that decreases over time (Query et al. 1997). In DHX15-depleted extracts with ATP, the A^{min}-complex persists, indicating that DHX15 and/or a co-depleted factor is responsible for the ATP-dependent loss. The faster migrating complex (*) is largely unaffected.

If DHX15 is responsible for disrupting the unproductive A^{min}-complex, then GTP should also promote its disassembly in normal nuclear extract. We compared the addition of ATP, GTP and AMP-PNP with the A^{min} substrate, and found that both ATP and GTP result in the loss of A^{min}-complex, while AMP-PNP does not (Fig. 2.3B). Taken together, these results support a model in which DHX15 disrupts an unproductive interaction between U2 snRNP and the intron in A^{min}-complex as a quality control mechanism.

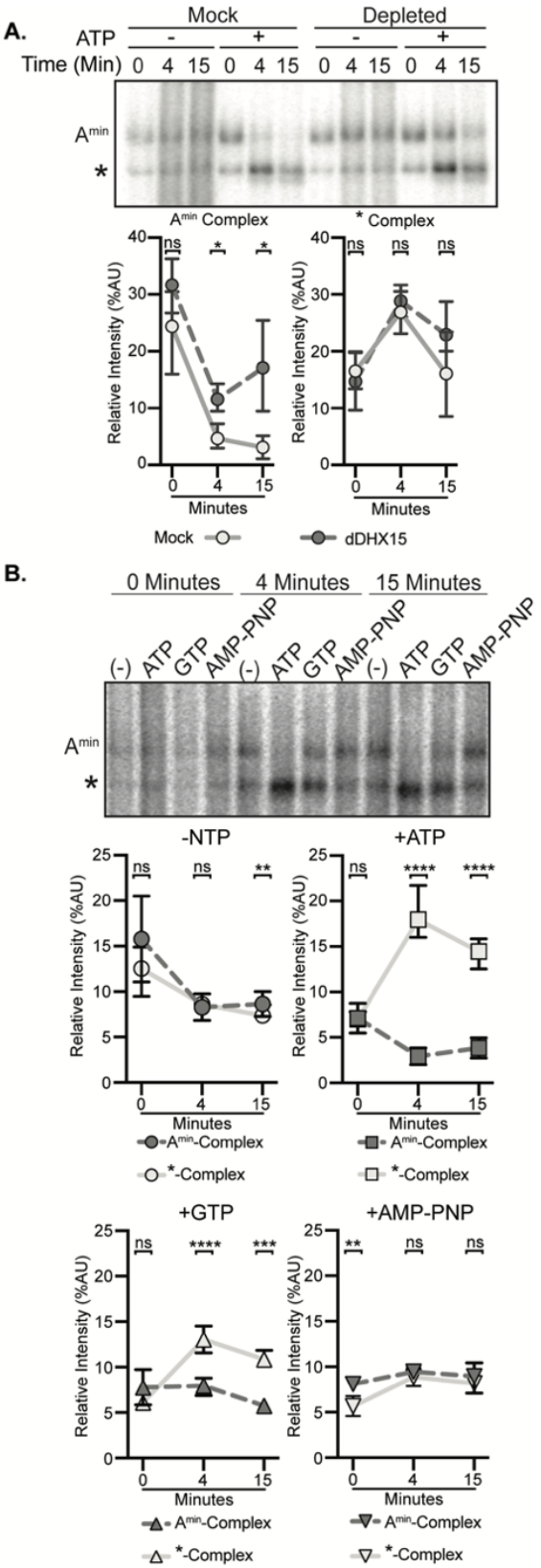

A.
Mock
Depleted
ATP
-
+
-
+
Time (Min) 0 4 15 0 4 15 0 4 15 0 4 15
Amin
*
Amin Complex
* Complex
Relative Intensity (%AU)
Minutes
ns
*
Mock
dDHX15
B.
0 Minutes
4 Minutes
15 Minutes
(-) ATP GTP AMP-PNP
-NTP
+ATP
+GTP
+AMP-PNP
**
****
***
Amin-Complex
*-Complex

Figure 2.3: Reduction of DHX15 stabilizes the ATP-independent interaction between U2 snRNP and minimal intron

A. Top: Representative native gel analysis of *in vitro* spliceosome assembly with a radiolabeled A^{min}-substrate in control (Mock) and DHX15-immunodepleted (dDHX15) nuclear extracts at the indicated timepoints. Bottom: Normalized band intensity relative to the entire lane from three independent trials. Statistical differences were examined by paired Student's t-test. B. Top: Representative native gel analysis of *in vitro* spliceosome assembly with A^{min}-substrate with different added NTPs at the indicated timepoints. Statistical differences were examined by unpaired Student's t-test.

A^{min} destabilization does not occur via an extended PYT

In the A^{min}-complex, there are only two possible RNAs for DHX15 to target: U2 snRNA and the short A^{min} substrate. Given that DHX-enzymes act by translocating from 3' to 5' on a single RNA strand (Hamann et al. 2019; Mallam et al. 2012; Sengoku et al. 2006; Tauchert et al. 2017; Yang et al. 2007; Yang and Jankowsky 2006), the 3' end of the A^{min} substrate is a likely DHX15 target. We tested this possibility by truncating the 3' end of the intron in 4 nucleotide increments to see if loss of the putative target sequence results in A^{min}-complex stabilization despite added ATP (Fig. 2.4A). We stopped at 10 nucleotides downstream of the branch point sequence because too short of a PYT interferes with spliceosome assembly (Bessonov et al. 2010). The truncations did not interfere with A^{min} assembly in the absence of ATP (Fig. 2.4B & 2.4C), and complexes were all destabilized with added ATP. We conclude an extended PYT is not necessary for A^{min} destabilization, meaning that DHX15 either interacts with the remaining 10 nucleotides of the PYT or recognizes another RNA feature within the A^{min}-complex.

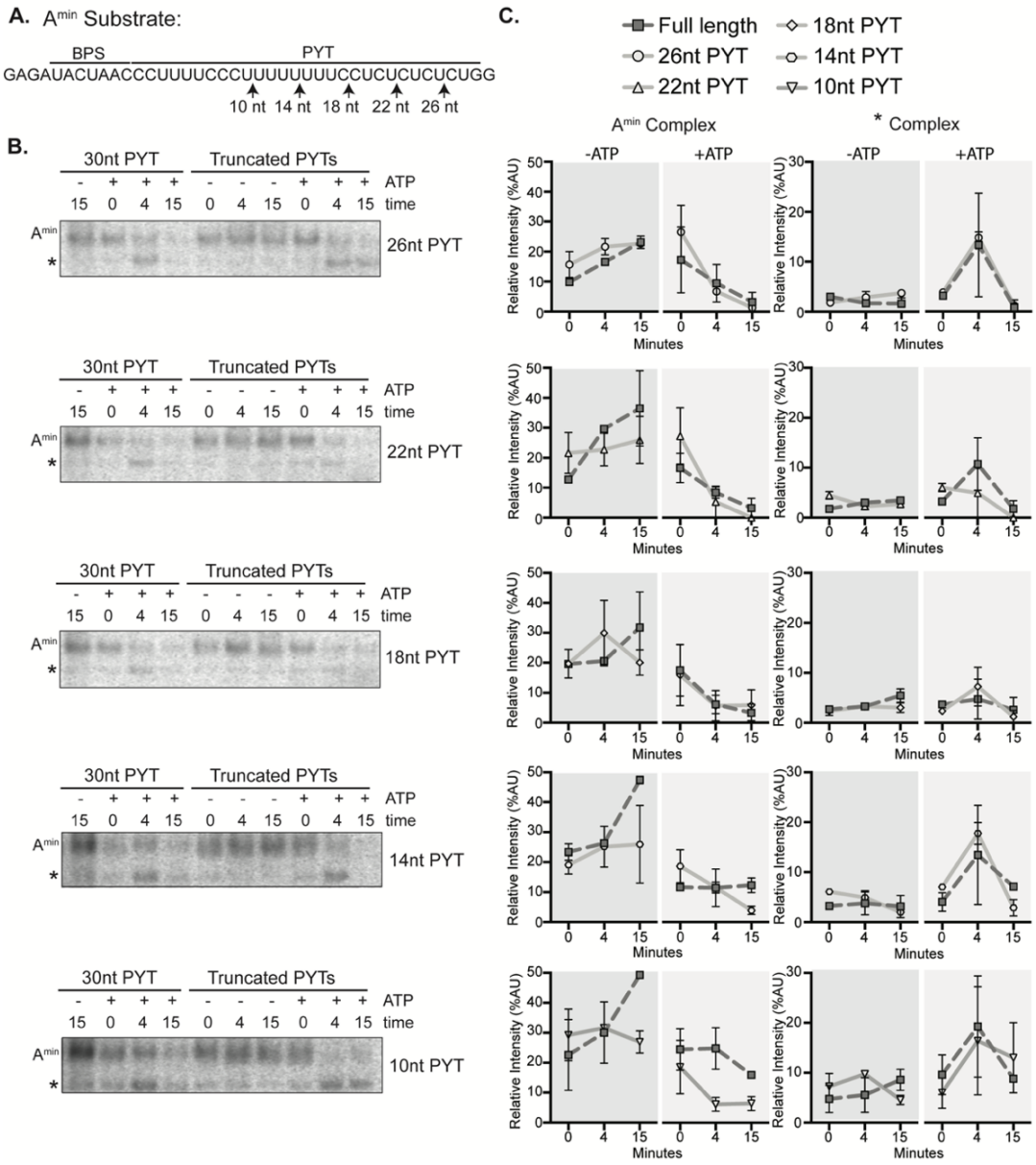


Figure 2.4: A^{min} destabilization does not occur via an extended PYT

A. Schematic of the A^{min} RNA substrate with branch point sequence (BPS) and PYT. The arrows indicate the 3' end of the PYT in each RNA. B. Representative native gel analysis of *in vitro* spliceosome assembly with truncated A^{min} substrate +/- ATP at the indicated timepoints. PYT length is indicated on the right. C. Background-corrected band intensity of A^{min}- and *-complexes from two independent trials. Graphs correspond to the adjacent native gel images in (B).

U2 snRNA accessibility is regulated by NTP hydrolysis

The other RNA in A^{min} is U2 snRNA, of which nucleotides 32-46 are of special interest because they both interact with the intron in an extended branch helix and form the mutually exclusive branch stem loop (BSL) structure (Perriman and Ares 2010; Zhang et al. 2020) (Fig. 2.5A). In the context of the U2 snRNP in nuclear extract, this region was shown to become accessible for base-pairing with a complementary DNA oligonucleotide in an ATP-dependent manner (Black et al. 1985), although the factor responsible for the ATP-dependence was unknown. To determine whether a DHX-enzyme could be involved, we tested whether GTP or AMP-PNP could also enable base-pairing of different regions in the 5' half of U2 snRNA. We created a series of overlapping DNA oligonucleotides complementary to U2 snRNA (nt 1-15, nt 12-26, nt 24-38, nt 32-46, Fig. 2.5A), which we added to HeLa nuclear extract treated with various NTP's. If the targeted region is accessible for base-pairing, endogenous RNase H in the extracts will induce cleavage of the RNA/DNA hybrid. We mapped the specific cleavage sites using primer extension (Fig. 2.5B) and quantified overall digestion of U2 snRNA induced by the oligonucleotides (Fig. 2.5C).

With an oligonucleotide that targets the 5' half of the BSL (nt 24-28), over 60% of U2 snRNA molecules in the extract are cleaved after both ATP and GTP treatment, compared with ~20% cleaved with no added NTP or AMP-PNP. Specifically, most cleavage occurs after nucleotides 32-36 (GUGUA) with some cleavage after nucleotide 37. Using an oligonucleotide targeting the 3' half of the BSL (nt 32-46), around half of U2 snRNA molecules are cleaved, primarily after nucleotides 42-46 independent of added NTP, although there is a small increase

in cleavage after nucleotides 37 and 38 with ATP and GTP. Cleavage in the presence of the other oligonucleotides is not influenced by NTP treatment. With an oligonucleotide targeting the beginning of U2 snRNA (nt 1-15), nearly all the U2 snRNA molecules are cleaved after U9 indicating that while the 5' end of U2 snRNA is available for base-pairing, nucleotides 10-15 are protected, likely because of their participation in Stem I (Fig. 2.5A). An oligonucleotide that targets nucleotides 12-26 also does not induce cleavage, further supporting the stability of Stem I.

Because both ATP and GTP hydrolysis results in increased accessibility, we conclude that a DHX-enzyme mediates the rearrangement that makes U2 snRNA nucleotides 24-38 available for base-pairing. The enzyme may unwind an RNA structure, such as the BSL, or dislodge a protein that protects the RNA.

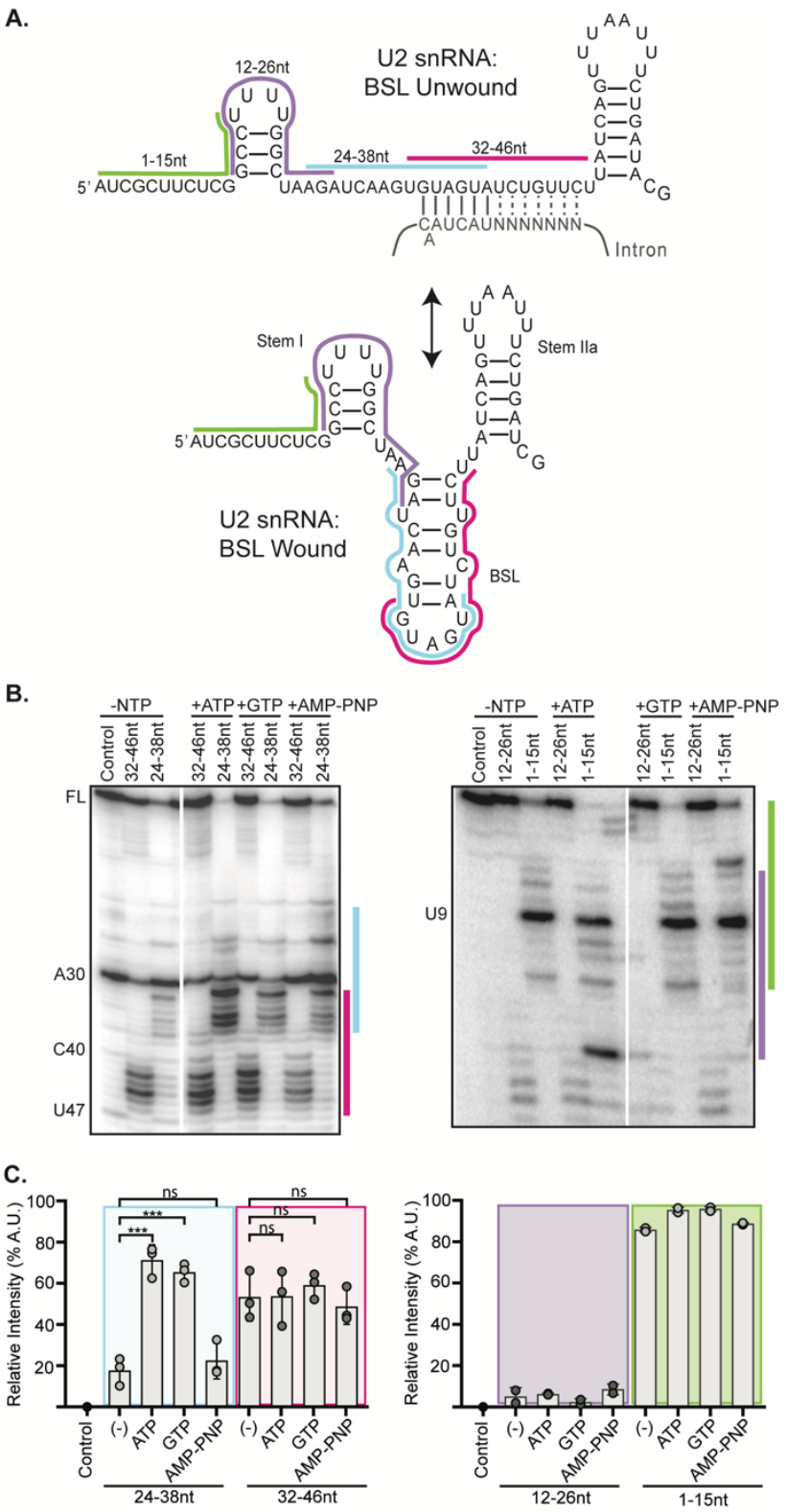

A.
U2 snRNA:
BSL Unwound
12-26nt
1-15nt
24-38nt
32-46nt
5' AUCGCUUCUCG
UAAGAUCAAGUGUAGUAUCUGUUCU
CAUCAUNNNNNNN
A
Intron
Stem I
Stem IIa
5' AUCGCUUCUCG
U2 snRNA:
BSL Wound
BSL
B.
-NTP
+ATP
+GTP
+AMP-PNP
Control
32-46nt
24-38nt
FL
A30
C40
U47
12-26nt
1-15nt
U9
C.
Relative Intensity (% A.U.)
ns
***
(-)
ATP
GTP
AMP-PNP
24-38nt
32-46nt
12-26nt
1-15nt

Figure 2.5: U2 snRNA accessibility is regulated by NTP hydrolysis
Secondary structure of the U2 snRNA with the BSL (bottom) or unwound interaction with the grey intron (top). The position of complementary DNA oligonucleotides used for RNase H digestion are indicated with colored bars. B. Primer extension analysis of U2 snRNA isolated from nuclear extract after RNase H digestion. Extracts were treated with ATP, GTP or AMP-PNP or no NTP. Control digestions were carried out with a non-complementary DNA oligonucleotide. C. Cleavage efficiency was determined as band intensity in the region complementary to the oligonucleotide over the intensity of the entire lane of three independent trials. Statistical differences relative to the -NTP sample were examined by unpaired Student's t-test.

U2 snRNP composition is affected by NTP hydrolysis

To determine if a change in protein interactions is responsible for the NTP-dependent change in U2 snRNA accessibility, we generated a HeLa cell line with a stably integrated transgene encoding a V5-tagged SNRPB2, a core U2 snRNP protein, for pulldown studies (Khandelia et al. 2011; Kim 2019). Western analysis of nuclear extract shows that over half the SNRPB2 expressed in these cells carries the tag (input lane), and that we can specifically immunoprecipitate (IP) the V5-tagged protein (Fig. 2.6A, Sup. Fig. 2.3). We analyzed anti-V5 IP's for RNA and see enrichment of U2 snRNA indicating that the tagged protein incorporates into U2 snRNP (Fig. 2.6B). IP's of *in vitro* splicing reactions using a radiolabeled full-length pre-mRNA reveal higher levels of pre-mRNA co-eluting with the V5-tagged SNRPB2 relative to IgG control, which is further enhanced by addition of ATP to the splicing reactions (Fig. 2.6C). We repeated the IP's with the A^{min} substrate, and in this case more RNA co-elutes from splicing reactions without added ATP (Fig. 2.6D). These results parallel our observations by native gel analysis, indicating that the V5-tagged U2 snRNP incorporates splicing complexes.

To test whether protein composition of U2 snRNP is altered by NTP hydrolysis, we IP'd V5-tagged U2 snRNP from nuclear extract incubated with ATP, GTP or AMP-PNP. Western analysis of eluates showed that the NTP treatments do not alter the immunoprecipitation of V5-tagged SNRPB2, or the association of the 17S U2 snRNP protein SF3A2 (Fig. 2.6E). We detect DHX15 in the eluates under all conditions. In contrast, U2AF1 is reduced after ATP treatment. This change is more likely due to the activity of a DDX-enzyme because GTP treatment does not result in the same decrease. AMP-PNP also has no effect. Relating these results to the increased U2 snRNA accessibility in the snRNP after both ATP and GTP treatment, the ATP-specificity of U2AF1 loss means it is not responsible for blocking access to the branch binding region. Because DHX15 remains with the snRNP, it could be responsible removing a different protein to make the branch binding region of U2 snRNA available for base-pairing interactions. It is tempting to speculate that DHX15 may be disrupting unproductive spliceosomes in a similar manner.

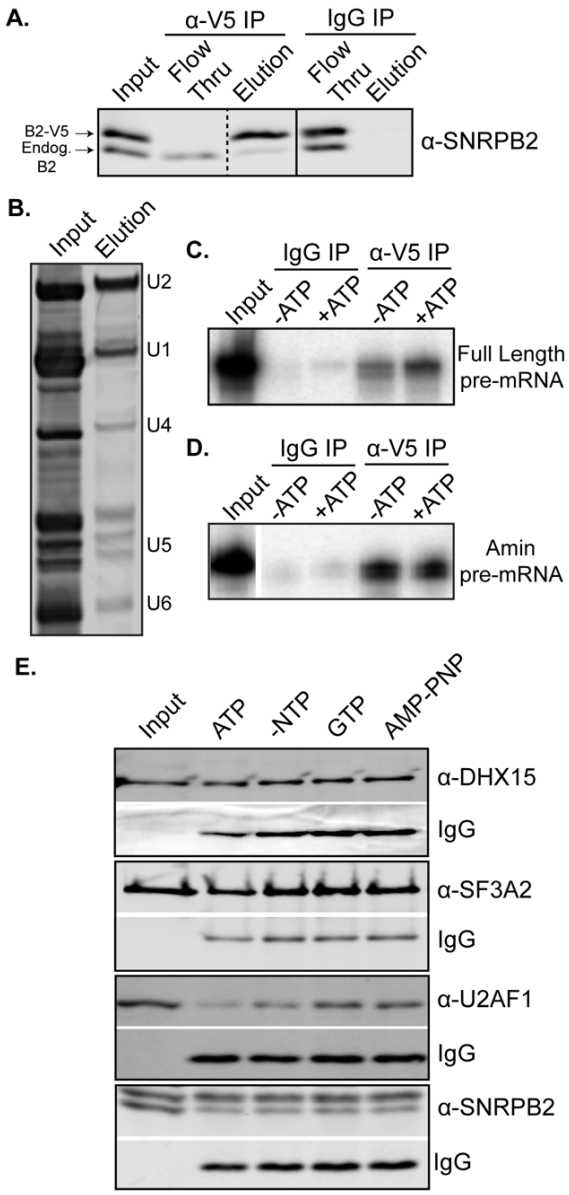


Figure 2.6: ATP hydrolysis alters U2 snRNP protein composition

A. Western blot analysis of IP with anti-V5 and mouse IgG from HeLa nuclear extract containing V5-tagged SNRPB2 and probed with anti-SNRPB2, which detects both endogenous (endog) and V5-tagged SNRPB2 (V5-B2). B. Denaturing gel analysis of RNA isolated from anti-V5 IP eluate stained with SYBR Gold. C. Denaturing gel analysis of RNA isolated anti-V5 IP of *in vitro* splicing reactions +/- ATP using a radiolabeled full-length RNA substrate in the conditions. D. Same a (C) except with an A^{min} substrate E. Western analysis anti-V5 of IP's from nuclear extract incubated under *in vitro* splicing conditions in the presence of ATP, GTP or AMP-PNP. Blots were probed with the indicated antibody. Bands labeled with IgG show the amount of anti-V5 antibody in the elutions.

Discussion

Spliceosome assembly relies on an assortment of DDX- and DHX-enzymes to both drive molecular rearrangements and promote splicing fidelity. Both DDX46 (yeast Prp5) and DDX39B (UAP56, yeast Sub2) play roles in U2 snRNP's addition to the branch point sequence. Recent cryo-EM structures show that DDX46 interacts with U2 snRNP near both the BSL and an unproductive branch-helix, consistent with its proposed roles in BSL rearrangements and fidelity of branch point selection (Liang and Cheng 2015; Perriman and Ares 2010; Zhang et al. 2020; Zhang et al. 2021). However, many questions remain open for both enzymes. For example, does DDX46 directly unwind the BSL, or does it dislodge HTATSF1 to destabilize the BSL? For DDX39B as well, what it targets, what it rearranges, and how it is regulated is still not clear. Using the differential nucleotide triphosphate selectivity of DDX- and DHX-enzymes, our study supports added roles for a DHX-enzyme(s) in regulating U2 snRNP structure and early assembly of human spliceosomes. We show that GTP can substitute for ATP in overcoming whatever blocks U2 snRNP from binding an intron with an anchor sequence. GTP can also mediate remodeling of the U2 snRNP to expose the branch binding sequence. Finally, GTP also promotes the destabilization of the unproductive interaction between U2 snRNP and an anchorless minimal intron, which we linked to the presence of DHX15. Additionally, we identified an ATP-specific loss of U2AF1 from U2 snRNP, which suggests the activity of a DDX-enzyme. DDX39B is a strong candidate because it has been shown to directly interact with U2AF1 (Shen et al. 2007).

DHX15 and its yeast ortholog Prp43 are ubiquitous nuclear RNA-dependent NTPases with disparate cellular functions and are best characterized as promoting ribosomal RNA biogenesis and spliceosome disassembly (Wen et al. 2008; Wild et al. 2010). DHX15's specificity is regulated by G-patch co-factors (Heininger et al. 2016), with NKRF controlling rRNA biogenesis (Memet et al. 2017) and TFIP11 mediating spliceosome disassembly (Studer et al. 2020). In this manuscript, we define a new role for DHX15 in early spliceosome assembly that may provide higher eukaryotes an additional quality control mechanism to ensure splicing fidelity in the face of divergent branch sequences. Our rationale for quality control stems from the increased accumulation of unproductive A-like complexes on three different substrates, in combination with decreased splicing efficiency for a full-length pre-mRNA (Fig. 2.2). In recent years links between recognition of the branch point region and cancer have dramatically increased (Agrawal et al. 2018). For example, hematological malignancies frequently select for specific point mutations in the U2 snRNP-associated proteins SF3B1 and U2AF1 (Hautin et al. 2020). Notably, DHX15 is also often misregulated in blood cancers (Jing et al. 2018; Pan et al. 2017).

Our working model is that DHX15 performs quality control of the initial interaction between U2 snRNP and an intron by disassembling complexes that are incompetent to form a productive branch helix (Fig. 2.7). It could work in concert with DDX46, which is proposed to directly proofread the branch helix. The yeast ortholog Prp5 releases from A-complex less readily with a mutant vs. consensus branch helix, with its release being necessary for continued spliceosome assembly (Liang and Cheng 2015). If we extend this model to

mammalian introns, which have more variable branch point sequences, DDX46 is more likely to stall on nonproductive branch helices. As a result, the more complex intronic landscape necessitates a mechanism for disassembly. We suspect that DHX15 fulfills this role through a disassembly mechanism that parallels disruption of U2 snRNA interactions with the intron in the ILS.

If DHX15 destabilizes stalled A-complex, how it is recognized comes into question. So far, DHX15's RNA targets are imposed by the identity of the complex being remodeled. For example, the enzyme promotes ILS disassembly by binding the 3' end of U6 snRNA (Toroney et al. 2019), although how the branch helix between U2 snRNA and the intron is disrupted remains unclear. A constitutively active DHX15 can also interact with the intron to promote disassembly of B^{act} spliceosomes (Fourmann et al. 2016). In the context the early spliceosome, DHX15's target is not yet defined. One option is the intron. With the limited real estate of the minimal intron substrate, we were able to narrow down the potential RNA handle to 10 nucleotides downstream from the branch point sequence. U2 snRNA is the other possibility, and we find that the branch binding region of U2 snRNA becomes accessible upon addition of both ATP and GTP. Because DHX15 associates with U2 snRNP, it is a good candidate for mediating the change, which could mimic removal of an intron in either the context of the ILS or A-complex.

Our model also requires a G-patch partner to direct DHX15 to stalled spliceosomes and may further dictate its RNA target. Notably, five different G-patch proteins associate with U2 snRNP and/or A-complex: RBM5, RBM10, RBM17, CHERP and SUGP1. Of these, SUGP1 binds to SF3B1 to prevent use

of cryptic branch points, although its DHX-enzyme partner was not identified (Zhang et al. 2019). Mutations in SUGP1 are also associated with aberrant splicing related to branch point selection in cancer cells that also correlate with SF3B1 mutations (Liu et al. 2020).

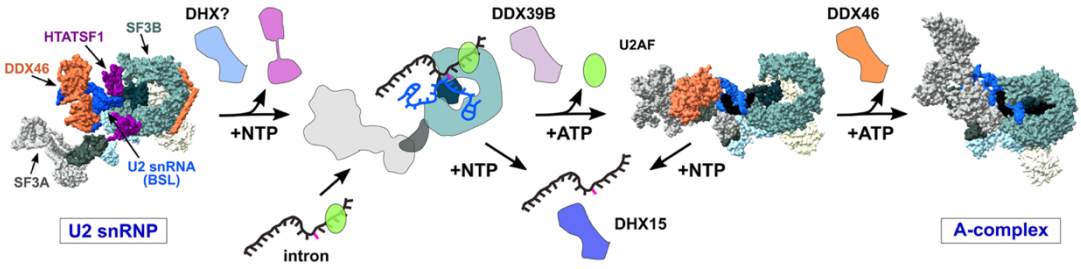


Figure 2.7: Model of DDX- and DHX-enzyme contributions to early spliceosome assembly.
Structural models are adapted from PDB entries: 6Y5Q, 7OQE, and 6G90.

Materials and Methods

HeLa nuclear extract

Nuclear extract was generated as previously described (Dignam et al. 1983) from HeLa S3 cells grown in DMEM/F-12 1:1 and 5% (v/v) newborn calf serum. HeLa cells containing an integrated transgene for V5-tagged SNRPB2 were generated using the HILO-RMCE cassette and RMCE acceptor cells (generously provided by E. Makeyev) as described in Khandelia, *et al.* (Khandelia et al. 2011). Expression of V5-tagged SNRPB2 was induced with doxycycline 48 hours prior to harvest (Kim 2019). Prior to use for the NTP studies, HeLa nuclear extract was either incubated at 30˚C for 15 minutes to cycle residual ATP or passed through a G-50 size exclusion column to deplete residual nucleotides.

T7 run-off RNA transcription

Full-length pre-mRNA sequence is derived from the adenovirus major late (AdML) transcript with uracil substituted at the last four bases of the 5' exon, and a UACUAAC branch point sequence. Sequences for the $A^{3'}$ and A^{min} RNA substrates are provided in Supplemental Fig. 2.4. Double stranded DNA oligos (Eurofins) were generated as templates for PYT truncations. All RNA substrates were generated by T7 run-off transcription reactions supplemented with ^{32}P-UTP. Full-length substrate was capped by G(5')ppp(5')G and gel purified. $A^{3'}$, A^{min} and PYT truncation substrates were GMP-capped and purified by size-exclusion spin column.

Native gel analysis of *in vitro* splicing reactions

Splicing reactions containing 20 nM RNA substrate, 40 mM potassium glutamate, 2 mM magnesium acetate, 0.05 mg/mL tRNA, 40% (v/v) HeLa nuclear extract

and 2 mM ATP, GTP, AMP-PNP or no added NTP were incubated at 30˚C for 0-15 minutes. 2X native loading dye (20mM Tris, 20mM glycine, 25% glycerol, 0.1% bromophenol blue, 0.1% cyan blue, 1mg/ml heparin) was added directly to each reaction, which were separated on native gels with 1.9% low melting temp agarose (Invitrogen, 15517-014) in 20 mM Tris, 20 mM glycine for 3 hours at 72 volts. Gels were dried onto Whatman paper and visualized by phosphorimaging (Amershan Typhoon).

Denaturing gel analysis of *in vitro* splicing reactions

Splicing reactions were generated as described above except that pre-mRNA substrate was at 6 nM and incubation time was extended to 45 minutes. Reactions were stopped with addition of splicing dilution buffer (100 mM Tris pH 7.5, 10 mM EDTA, 1% SDS, 150 mM sodium chloride, and 0.3 M sodium acetate pH 5.2) followed by phenol:chloroform:isoamyl alcohol (25:24:1) extraction and ethanol precipitation. RNA was resuspended in FEB (95% formamide, 20 mM EDTA, 0.01% bromophenol blue and 0.01% cyan blue), separated by 15% urea-PAGE and visualized by phosphorimaging.

DHX15 depletion

10 µg DHX15 antibody (Abcam, ab70454), IgG (Epredia™, NC748P) or 1X PBS was incubated with protein A beads (NEB, S1425S) with agitation for 1 hour at 4˚C. HeLa nuclear extract with 500 mM potassium chloride was added to the antibody bound beads and rotated end over end for 2 hours at 4˚C. The depleted nuclear extract was then dialyzed into Buffer E for 5 hours (100 mM potassium chloride, 20% glycerol, 20 mM Tris, pH 7.9, 1.5 mM magnesium chloride, 0.5 mM DTT) in small dialysis cups (Thermo Scientific™, 69570).

RNase H digestion and primer extension

HeLa nuclear extract was supplemented with 2 mM magnesium acetate and 2 mM ATP, GTP, AMP-PNP or no NTP. DNA oligos (Eurofins) complementary to the U2 snRNA nt 1-15, nt 12-26, nt 24-38 and nt 32-46 or a control oligo (Sup. Fig. 2.5) was added at 5 μM and incubated for 15 minutes at 30˚C to allow for cleavage by endogenous RNase H. The oligonucleotides were degraded by addition of 1μl RQ1 DNase (Promega, M6101) for 10 minutes at 30˚C. RNA was isolated by phenol:chloroform:isoamyl alcohol (25:24:1) and chloroform:isoamyl alcohol (24:1) extraction, followed by ethanol precipitation and resuspended in water. For primer extension, 10 picomoles of DNA oligonucleotides complementary to either U2 snRNA nt 97-117 or nt 28-42 (U2L15 (Black et al. 1985)) were labeled with γ-32P ATP and purified via Sephadex G-25 (MilliporeSigma™ Supelco™ G258010G) column. The isolated RNA and radiolabeled oligonucleotide were annealed by incubation at 95˚C for 2 min, 53˚C for 5 min, and on ice for 5 minutes and then added to reverse transcription reactions containing 50 mM Tris pH 7.9, 75 mM potassium chloride, 7 mM DTT, 3 mM magnesium chloride, 1 mM dNTPs, and 0.5 μg reverse transcriptase (MMLV variant). Reactions were incubated for 30 minutes at 53˚C, and the DNA was isolated by the addition of 0.3 M sodium acetate pH 5.2, 0.5 mM EDTA, and 0.05% SDS followed by ethanol precipitation. The labeled DNA was resuspended in FEB and separated on a 9.6% urea:PAGE. The gel was dried on Whatman paper and visualized by phosphorimaging.

Immunoprecipitations

For IP's in Fig. 6A, C & D, 1.5 µg of anti-V5 antibody (Invitrogen, R960-25) or IgG (Epredia™, NC748P) was added directly to splicing reactions containing 6 nM full-length RNA substrate, 2 mM magnesium acetate, 40 mM potassium glutamate, 0.1 mg/ml tRNA and 40% V5-SNRPB2 HeLa +/- 2 mM ATP incubated at 30˚C for 10 minutes. For other co-IP's, 3 µg of anti-V5 antibody (Genescript, A01724) was added to 2 mM magnesium acetate, 20 mM potassium glutamate, 0.1 mg/ml tRNA and 60% V5-SNRPB2 HeLa nuclear extract +/– 2 mM ATP, GTP or AMP-PNP and incubated at 30˚C for 10 minutes. All samples were incubated with antibody for 13.5 hours at 4˚C with rocking. Samples were then added to protein A magnetic beads (NEB, S1425S) and rotated at 4˚C for 4 hours. The beads were washed three or more times with IP wash buffer (100 mM Tris pH 7.5, 120 mM potassium chloride, 1 mM EGTA, 0.1% NP-40 or IGEPAL). For western blots, samples were eluted with 0.1 M glycine, pH 2.5 and quenched with equal volume 1 M Tris, pH 7.9. For RNA analysis, samples were eluted with splicing dilution buffer followed by precipitation with phenol:chloroform:isoamyl alcohol (25:24:1) and ethanol precipitation. RNA samples were analyzed by 15% urea-PAGE as described for *in vitro* splicing reactions.

Western blot analysis

Samples were prepared in 5X Laemmli buffer (62.5 mM Tris, 25% glycerol, 6.25% SDS, 0.1% bromophenol blue, 5% beta-mercaptoethanol) and heated at 95˚C for 2 minutes prior to separation by 10% SDS-PAGE. Gels were transferred to PVDF membrane (Bio-Rad Mini Trans-Blot® Cell) and blocked in 1% non-fat milk in 1X Tris-buffered saline with Tween 20 (TBST) for 1 hour at room

temperature while rocking. The following antibodies were added directly to the blocking buffer at the indicated concentrations and incubated at 4˚C overnight while rocking. From Proteintech: DHX15 (12265-1-AP, 1:1000), SNRPB2 (13512-1-AP, 1:2500), and U2AF1 (60289-1-Ig, 1:1000). From Santa Cruz Biotech: SF3A2 (sc-390444, 1:1000) and HDAC1 (sc-7872, 1:5000). Blots were washed three times in 1X TBST and corresponding LICOR secondaries (1:15000) were added in blocking buffer and rocked at room temperature for 1 hour. The blots were again washed three times and then imaged on a LICOR Imaging System. Images were quantified and processed utilizing LICOR Image Studio Lite.

Acknowledgements

This work was supported by a National Institutes of Health grant R01GM72649 to M.S.J.

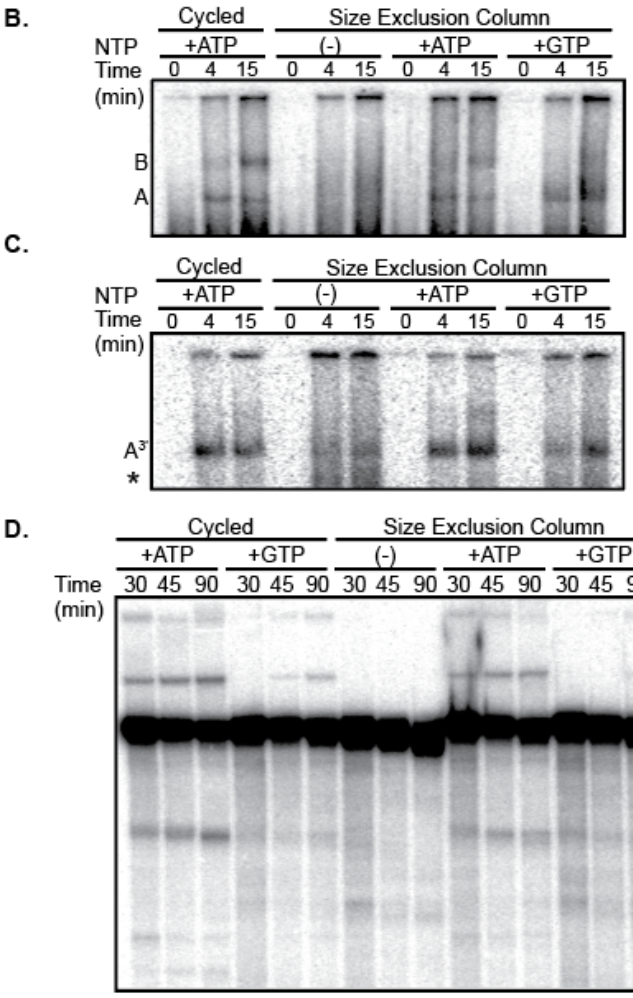
Author Contributions: Conceived and designed the experiments: HMN, AA, JK, MM, MJ. Performed the experiments: HMN, AA, TS, JK, MM, BP. Analyzed the data: HMN, AA, JK, MJ. Wrote the paper: HMN, AA, MJ.

Competing interests: The authors have declared that no competing interests exist.

Supplementary Figures:

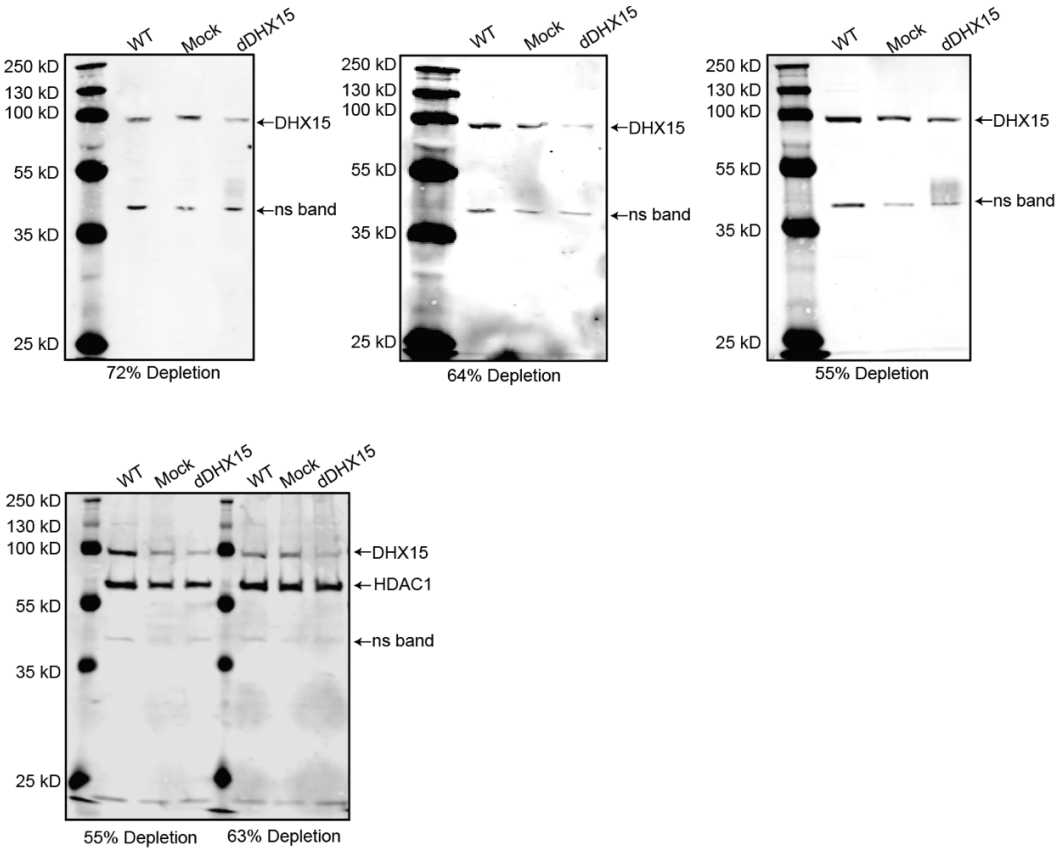
A.

Prior to SEC (cpm)	Post SEC (cpm)	Water Blank (cpm)
1348.33	41.33	20.00
Normalized Prior to SEC (cpm)	Normalized Post SEC (cpm)	Fold Change (%)
1328.33	21.33	63.21



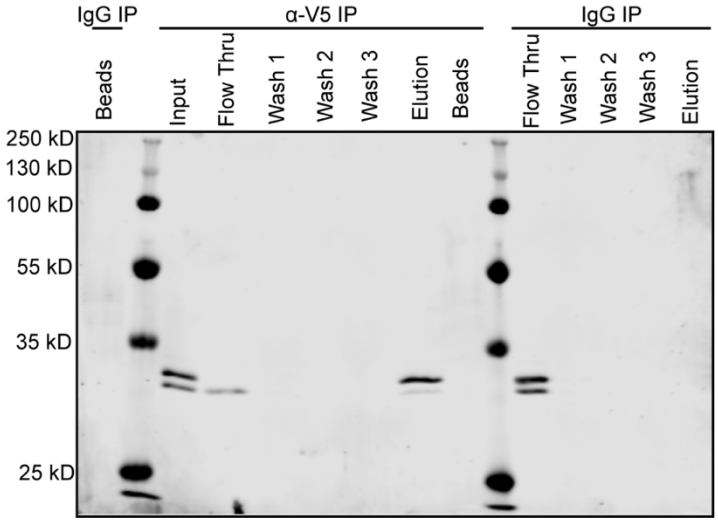
Supplemental Figure 2.1: Size Exclusion Column Treatment of Nuclear Extract

Removal of radiolabeled UTP by size exclusion column. B. Native assembly on a full-length substrate with nuclear extract treated at 30˚C (cycled) or run over a size exclusion column prior to use. C. Same analysis as B. however assembly occurred on a $A^{3'}$ substrate. D. Denaturing gel analysis of splicing reactions in nuclear extract treated at 30˚C (cycled) or run over a size exclusion column prior to use.



Supplemental Figure 2.2: DHX15 Depletion Westerns.

Full western blot images for DHX15 depletion shown in Figure 2.2 as well as DHX15 depletions blotting for anti-HDAC1 as a loading control for depletion.



Supplemental Figure 2.3: Western blot analysis of anti-V5 and IgG IP's

Full western blot IP's probed with anti-SNRPB2 from main Figure 2.6A.

Full-Length

GAGACCGGCAGATCAGCTTGGCCGCGTCCATCTGGTCATCTAGGATCTGAT
ATCATCGATGAATTCGAGCTCGGTACCCCGTTCGTCCTCACTCTCTTCCGCA
TCGCTGTCTGCGAGGGCCAGCGTAAAAGGTGAGTACTCCCTCTCAAAAGCG
GGCATGACTTCTGCCCTCGAGTTATTAACCCTCACTAAAGGCAGTAGTCAAG
GGTTTCCTTGAAGCTTTCGTGCTGACCCTGTCCCTTTTTTTTCCACAGCTGC
AGGTCGACGTTGAGGACAAACTCTTCGCGGTCTTTCCAGTACTCTTG

$A^{3'}$

GTCAAGGGTTTCCTTGAAGCTTTCGTACTAACCCTTTCCCTTTTTTTTCCTCT
CTCTCTCGAGCGTACACCATCAGGGTACGAGCTAGCCCATGGCGTACACCA
TCAGGGTACGACTAGTAGATCTCGTACACCATCAGGGTACGGAATTCT

A^{min}

GTACTAACCCTTTCCCTTTTTTTTCCTCTCTCTCTGCTCGAGCGTACACCATC
AGGGTACGAGCTAGCCCATGGCGTACACCATCAGGGTACGACTAGTAGATC
TCGTACACCATCAGGGTACGGAATTCT

Supplemental Figure 2.4: RNA substrates.

DNA template sequences of the RNA substrates utilized in this study.

Name	DNA Sequence
U2 snRNA nt 32-46	AGAACAGATACTACA
U2 snRNA nt 24-38	TACTACACTTGATCT
U2 snRNA nt 12-26	TCTTAGCCAAAAGGC
U2 snRNA nt 1-15	AGGCCGAGAAGCGAT
Control Oligo	CTGATGTCCGCAGACAATGT

Supplemental Figure 2.5: Sequences of DNA oligonucleotides complementary to the indicated U2 snRNA region for RNase H digestion.

References

Agafonov DE, Deckert J, Wolf E, Odenwalder P, Bessonov S, Will CL, Urlaub H, Luhrmann R. 2011. Semiquantitative proteomic analysis of the human spliceosome via a novel two-dimensional gel electrophoresis method. *Mol Cell Biol* **31:** 2667-2682. doi: 10.1128/MCB.05266-11

Agrawal AA, Yu L, Smith PG, Buonamici S. 2018. Targeting splicing abnormalities in cancer. *Curr Opin Genet Dev* **48:** 67-74. doi: 10.1016/j.gde.2017.10.010

Alsafadi S et al. 2016. Cancer-associated SF3B1 mutations affect alternative splicing by promoting alternative branchpoint usage. *Nat Commun* **7:** 10615. doi: 10.1038/ncomms10615

Anderson K, Moore MJ. 2000. Bimolecular exon ligation by the human spliceosome bypasses early 3' splice site AG recognition and requires NTP hydrolysis. *RNA* **6:** 16-25. doi: 10.1017/s1355838200001862

Bessonov S, Anokhina M, Krasauskas A, Golas MM, Sander B, Will CL, Urlaub H, Stark H, Luhrmann R. 2010. Characterization of purified human Bact spliceosomal complexes reveals compositional and morphological changes during spliceosome activation and first step catalysis. *RNA* **16:** 2384-2403. doi: 10.1261/rna.2456210

Black DL, Chabot B, Steitz JA. 1985. U2 as well as U1 small nuclear ribonucleoproteins are involved in premessenger RNA splicing. *Cell* **42:** 737-750.

Bonnal SC, López-Oreja I, Valcárcel J. 2020. Roles and mechanisms of alternative splicing in cancer - implications for care. *Nat Rev Clin Oncol* **17:** 457-474. doi: 10.1038/s41571-020-0350-x 10.1101/542506

Darman RB et al. 2015. Cancer-Associated SF3B1 Hotspot Mutations Induce Cryptic 3' Splice Site Selection through Use of a Different Branch Point. *Cell Rep* **13:** 1033-1045. doi: 10.1016/j.celrep.2015.09.053

DeBoever C, Ghia EM, Shepard PJ, Rassenti L, Barrett CL, Jepsen K, Jamieson CH, Carson D, Kipps TJ, Frazer KA. 2015. Transcriptome sequencing reveals potential mechanism of cryptic 3' splice site selection in SF3B1-mutated cancers. *PLoS Comput Biol* **11:** e1004105. doi: 10.1371/journal.pcbi.1004105

Dignam JD, Lebovitz RM, Roeder RD. 1983. Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. *Nucleic Acids Res* **11:** 1475-1489.

Dvinge H, Kim E, Abdel-Wahab O, Bradley RK. 2016. RNA splicing factors as oncoproteins and tumour suppressors. *Nat Rev Cancer* **16:** 413-430. doi: 10.1038/nrc.2016.51

Fleckner J, Zhang M, Valcarcel J, Green MR. 1997. U2AF65 recruits a novel human DEAD box protein required for the U2 snRNP-branchpoint interaction. *Genes Dev* **11:** 1864-1872.

Fourmann JB, Dybkov O, Agafonov DE, Tauchert MJ, Urlaub H, Ficner R, Fabrizio P, 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. doi: 10.7554/eLife.15564

Gilman B, Tijerina P, Russell R. 2017. Distinct RNA-unwinding mechanisms of DEAD-box and DEAH-box RNA helicase proteins in remodeling structured RNAs and RNPs. *Biochem Soc Trans* **45:** 1313-1321. doi: 10.1042/BST20170095

Hamann F, Enders M, Ficner R. 2019. Structural basis for RNA translocation by DEAH-box ATPases. *Nucleic Acids Res* **47:** 4349-4362. doi: 10.1093/nar/gkz150

Hautin M, Mornet C, Chauveau A, Bernard D, Corcos L, Lippert E. 2020. Splicing Anomalies in Myeloproliferative Neoplasms: Paving the Way for New Therapeutic Venues. *Cancers (Basel)* **12:** E2216. doi: 10.3390/cancers12082216

He Y, Andersen GR, Nielsen KH. 2010. Structural basis for the function of DEAH helicases. *EMBO Rep* **11:** 180-186. doi: 10.1038/embor.2010.11

Heininger AU et al. 2016. Protein cofactor competition regulates the action of a multifunctional RNA helicase in different pathways. *RNA Biol* **13:** 320-330. doi: 10.1080/15476286.2016.1142038

Jankowsky E. 2011. RNA helicases at work: binding and rearranging. *Trends Biochem Sci* **36:** 19-29. doi: 10.1016/j.tibs.2010.07.008

Jing Y et al. 2018. DHX15 promotes prostate cancer progression by stimulating Siah2-mediated ubiquitination of androgen receptor. *Oncogene* **37:** 638-650. doi: 10.1038/onc.2017.371

Kesarwani AK, Ramirez O, Gupta AK, Yang X, Murthy T, Minella AC, Pillai MM. 2017. Cancer-associated SF3B1 mutants recognize otherwise inaccessible cryptic 3' splice sites within RNA secondary structures. *Oncogene* **36:** 1123-1133. doi: 10.1038/onc.2016.279

Khandelia P, Yap K, Makeyev EV. 2011. Streamlined platform for short hairpin RNA interference and transgenesis in cultured mammalian cells. *Proc Natl Acad Sci U S A* **108:** 12799-12804. doi: 10.1073/pnas.1103532108

Kim J. A Pipeline for Tagging snRNP Associated Proteins in HeLa. Santa Cruz, CA: University of California; 2019. p. Dissertation.

Konarska MM, Sharp PA. 1986. Electrophoretic separation of complexes involved in the splicing of precursors to mRNAs. *Cell* **46:** 845-55.

Koodathingal P, Novak T, Piccirilli JA, Staley JP. 2010. The DEAH box ATPases Prp16 and Prp43 cooperate to proofread 5' splice site cleavage during pre-mRNA splicing. *Mol Cell* **39:** 385-395. doi: 10.1016/j.molcel.2010.07.014

Koodathingal P, Staley JP. 2013. Splicing fidelity: DEAD/H-box ATPases as molecular clocks. *RNA Biol* **10:** 1073-1079. doi: 10.4161/rna.25245

Liang WW, Cheng SC. 2015. A novel mechanism for Prp5 function in prespliceosome formation and proofreading the branch site sequence. *Genes Dev* **29:** 81-93. doi: 10.1101/gad.253708.114

Liu Z, Zhang J, Sun Y, Perea-Chamblee TE, Manley JL, Rabadan R. 2020. Pan-cancer analysis identifies mutations in SUGP1 that recapitulate mutant SF3B1 splicing dysregulation. *Proc Natl Acad Sci U S A* **117:** 10305-10312. doi: 10.1073/pnas.1922622117

Mallam AL, Del Campo M, Gilman B, Sidote DJ, Lambowitz AM. 2012. Structural basis for RNA-duplex recognition and unwinding by the DEAD-box helicase Mss116p. *Nature* **490:** 121-125. doi: 10.1038/nature11402

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

Memet I, Doebele C, Sloan KE, Bohnsack MT. 2017. The G-patch protein NF-κB-repressing factor mediates the recruitment of the exonuclease XRN2 and activation of the RNA helicase DHX15 in human ribosome biogenesis. *Nucleic Acids Res* **45:** 5359-5374. doi: 10.1093/nar/gkx013

Newnham CM, Query CC. 2001. The ATP requirement for U2 snRNP addition is linked to the pre-mRNA region 5' to the branch site. *RNA* **7:** 1298-1309.

O'Day CL, Dalbadie-McFarland G, Abelson J. 1996. The Saccharomyces cerevisiae Prp5 protein has RNA-dependent ATPase activity with specificity for U2 small nuclear RNA. *J Biol Chem* **271:** 33261-33267. doi: 10.1074/jbc.271.52.33261

Pan L et al. 2017. DHX15 is associated with poor prognosis in acute myeloid leukemia (AML) and regulates cell apoptosis via the NF-kB signaling pathway. *Oncotarget* **8:** 89643-89654. doi: 10.18632/oncotarget.20288

Perriman R, Ares M, Jr. 2010. Invariant U2 snRNA nucleotides form a stem loop to recognize the intron early in splicing. *Mol Cell* **38:** 416-427. doi: 10.1016/j.molcel.2010.02.036

Polach KJ, Uhlenbeck OC. 2002. Cooperative binding of ATP and RNA substrates to the DEAD/H protein DbpA. *Biochemistry* **41:** 3693-3702. doi: 10.1021/bi012062n

Query CC, McCaw PS, Sharp PA. 1997. A minimal spliceosomal complex A recognizes the branch site and polypyrimidine tract. *Mol Cell Biol* **17:** 2944-2953.

Robert-Paganin J et al. 2017. Functional link between DEAH/RHA helicase Prp43 activation and ATP base binding. *Nucleic Acids Res* **45:** 1539-1552. doi: 10.1093/nar/gkw1233

Sengoku T, Nureki O, Nakamura A, Kobayashi S, Yokoyama S. 2006. Structural basis for RNA unwinding by the DEAD-box protein Drosophila Vasa. *Cell* **125:** 287-300. doi: 10.1016/j.cell.2006.01.054

Shen J, Zhang L, Zhao R. 2007. Biochemical characterization of the ATPase and helicase activity of UAP56, an essential pre-mRNA splicing and mRNA export factor. *J Biol Chem* **282:** 22544-22550. doi: 10.1074/jbc.M702304200

Studer MK, Ivanović L, Weber ME, Marti S, Jonas S. 2020. Structural basis for DEAH-helicase activation by G-patch proteins. *Proc Natl Acad Sci U S A* **117:** 7159-7170. doi: 10.1073/pnas.1913880117

Tanner NK, Cordin O, Banroques J, Doère M, Linder P. 2003. The Q motif: a newly identified motif in DEAD box helicases may regulate ATP binding and hydrolysis. *Mol Cell* **11:** 127-138. doi: 10.1016/s1097-2765(03)00006-6

Tauchert MJ, Fourmann JB, Lührmann R, Ficner R. 2017. Structural insights into the mechanism of the DEAH-box RNA helicase Prp43. *Elife* **6:** e21510. doi: 10.7554/eLife.21510

Toroney R, Nielsen KH, Staley JP. 2019. Termination of pre-mRNA splicing requires that the ATPase and RNA unwindase Prp43p acts on the catalytic snRNA U6. *Genes Dev* **33:** 1555-1574. doi: 10.1101/gad.328294.119

Uhlmann-Schiffler H, Jalal C, Stahl H. 2006. Ddx42p--a human DEAD box protein with RNA chaperone activities. *Nucleic Acids Res* **34:** 10-22. doi: 10.1093/nar/gkj403

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

Walbott H, Mouffok S, Capeyrou R, Lebaron S, Humbert O, van Tilbeurgh H, Henry Y, Leulliot N. 2010. Prp43p contains a processive helicase structural architecture with a specific regulatory domain. *EMBO J* **29:** 2194-2204. doi: 10.1038/emboj.2010.102

Wang L et al. 2016. Transcriptomic Characterization of SF3B1 Mutation Reveals Its Pleiotropic Effects in Chronic Lymphocytic Leukemia. *Cancer Cell* **30:** 750-763. doi: 10.1016/j.ccell.2016.10.005

Wen X, Tannukit S, Paine ML. 2008. TFIP11 interacts with mDEAH9, an RNA helicase involved in spliceosome disassembly. *Int J Mol Sci* **9:** 2105-2113. doi: 10.3390/ijms9112105

Wild T, Horvath P, Wyler E, Widmann B, Badertscher L, Zemp I, Kozak K, Csucs G, Lund E, Kutay U. 2010. A protein inventory of human ribosome biogenesis reveals an essential function of exportin 5 in 60S subunit export. *PLoS Biol* **8:** e1000522. doi: 10.1371/journal.pbio.1000522

Will CL, Urlaub H, Achsel T, Gentzel M, Wilm M, Luhrmann R. 2002. Characterization of novel SF3b and 17S U2 snRNP proteins, including a human Prp5p homologue and an SF3b DEAD-box protein. *EMBO J* **21:** 4978-4988.

Xu YZ, Query CC. 2007. Competition between the ATPase Prp5 and branch region-U2 snRNA pairing modulates the fidelity of spliceosome assembly. *Mol Cell* **28:** 838-849. doi: 10.1016/j.molcel.2007.09.022

Yang F, Wang XY, Zhang ZM, Pu J, Fan YJ, Zhou J, Query CC, Xu YZ. 2013. Splicing proofreading at 5' splice sites by ATPase Prp28p. *Nucleic Acids Res* **41:** 4660-4670. doi: 10.1093/nar/gkt149

Yang Q, Del Campo M, Lambowitz AM, Jankowsky E. 2007. DEAD-box proteins unwind duplexes by local strand separation. *Mol Cell* **28:** 253-263. doi: 10.1016/j.molcel.2007.08.016

Yang Q, Jankowsky E. 2006. The DEAD-box protein Ded1 unwinds RNA duplexes by a mode distinct from translocating helicases. *Nat Struct Mol Biol* **13:** 981-986. doi: 10.1038/nsmb1165

Zhang J, Ali AM, Lieu YK, Liu Z, Gao J, Rabadan R, Raza A, Mukherjee S, Manley JL. 2019. Disease-Causing Mutations in SF3B1 Alter Splicing by Disrupting Interaction with SUGP1. *Mol Cell* **76:** 82-95.e7. doi: 10.1016/j.molcel.2019.07.017

Zhang Z, Rigo N, Dybkov O, Fourmann JB, Will CL, Kumar V, Urlaub H, Stark H, Lührmann R. 2021. Structural insights into how Prp5 proofreads the pre-mRNA branch site. *Nature* **596:** 296-300. doi: 10.1038/s41586-021-03789-5

Zhang Z et al. 2020. Molecular architecture of the human 17S U2 snRNP. *Nature* **583:** 310-313. doi: 10.1038/s41586-020-2344-3

Chapter 3: Creation of splicing nuclear extracts from non-HeLa cell lines and HeLa derived lines

Introduction

To study the mechanism of splicing in a cell free environment, scientists have relied upon nuclear extracts purified from cell lines, specifically HeLa cells, to provide the spliceosomal components. The findings in Chapter 2 of this dissertation relied on nuclear extract derived from two distinct HeLa lines, one from the Luhrmann lab and one derived from a HeLa line engineered to contain a Lox/P site (Khandelia et al., 2011; Kim, 2019), both of which yield active nuclear extracts. Our lab created a V5-SNRPB2 HeLa line with Lox/P recombination that also yields nuclear extract that is active for *in vitro* splicing (Kim, 2019). My studies in Chapter 2 relied on depletion of DHX15 from nuclear extracts because we did not have a way to genetically remove it. Recently, Qing Feng, a postdoc in Chris Burge's Lab at MIT created a DHX15-degron line in HEK293T (Feng et al., 2022). To allow us to directly ask how removal of DHX15 from HeLa nuclear extract affects *in vitro* spliceosome assembly, she also added a degron tag to DHX15 in our V5-SNRNPB2 HeLa line. In this chapter, I discuss the challenges in using nuclear extracts from both DHX15-degron lines to assemble spliceosomes *in vitro*.

HeLa cells are derived from a cervical cancer tumor and have accumulated several additional anomalies, but revolutionized scientific discovery for the past 80 years (Landry et al., 2013). In 1983, Dignam et al. published a protocol to make an extract from HeLa cell nuclei that is active for *in vitro* transcription (Dignam et al., 1983) and this protocol was adapted to pre-mRNA splicing by the Sharp Lab (Padgett et al., 1983). Since this time, the protocol for

purification of nuclear extract from HeLa cells has not changed despite the establishment of other immortalized cell lines. Even though this line was derived from a tumor and is known to have cellular defects, scientists continue to rely on HeLa nuclear extract for splicing assays despite the inherently variable nature of HeLa from different labs and the inconsistencies in splicing efficiency between nuclear extract extracts from the same cell line. A handful of groups have utilized other immortalized lines for *in vitro* splicing assays with some success, however these systems are not used widely (Chou et al., 2000; Kataoka and Dreyfuss, 2004).

Genomic manipulation of HeLa cells is challenging due to high copy number of chromosomes, therefore the ability to generate nuclear extract in other cell lines more amenable to CRISPR would allow scientists to ask genetic questions regarding spliceosome assembly that have thus far precluded scientific discovery. In addition, expanding the number of cell lines utilized to provide spliceosome components would allow us to address the caveats of the splicing field working with a single cell line and confirm that observations in one cell type expands to others. One challenge we face is that *in vitro* spliceosome studies require a significant amount of nuclear extract, with researchers often growing liters of cells to generate enough extract for a given experiment. While this approach is straightforward in suspension-adapted HeLa, it is not conducive to testing activity of multiple cell lines. Therefore, I adapted a tabletop nuclear extract protocol from Lee et al. to generate nuclear extracts from other cellular backgrounds including HEK293T and genetically engineered HeLa (Lee et al., 1988). I found that generating extracts from HEK293T and genetically altered

HeLa is not straightforward and will require additional inquiry and optimization to ensure robust activity in all extract preps.

Results

HEK293T nuclear and whole cell extracts do not support spliceosome assembly

To determine whether another cell line could be utilized to make nuclear extract active for *in vitro* splicing, I used HEK293T cells, which are a common line used for genetic manipulation and more closely resemble normal human cells. To generate the extract, I grew HEK293T cells in adherent culture and used the Dounce protocol (see **Appendix I – NE HEK293T Dounce Protocol**). The nuclear extract was tested for its ability to support in vitro spliceosome assembly on three distinct RNA substrates- A^{min}, $A^{3'}$ and full-length AdML derivatives (Fig. 1.6) in addition to two non-AdML based introns (Ftz and ß-Glo, data not shown).

Unlike HeLa nuclear extract, the A^{min} complex that in HeLa co-migrates with A-complex does not form in HEK293T nuclear extract in either - or +ATP conditions. In HEK293T nuclear extract, the A^{min} substrate did form a complex that co-migrated at the * band we observe in HeLa with A^{min} (Newnham, 2001). The * complex was present in both - and +ATP conditions, however the A^{min} complex that co-migrates with A-complex in HeLa extract, was not present in either - or +ATP conditions in HEK293T NE (Fig. 3.1A). We refer to this co-migratory species as HEK* and like * in HeLa, have yet to unlock what these bands represent in terms of the components interacting with the radiolabeled A^{min} pre-mRNA. The amount of HEK* band present in each HEK293T prep would vary, however HEK* in HEK293T nuclear extract was consistently higher than HeLa *.

To ensure that the HEK* band is not due to a non-specific RNA/protein interaction unrelated to splicing, I compared the ability of two separate HEK293T nuclear extracts to assemble upon two RNA substrates: A^{min} and AA12, a scrambled pre-mRNA control of A^{min} that does not possess a canonical BPS nor PYT. The HEK* complex did not form on the scrambled intron (AA12) compared to the A^{min} substrate, meaning that the RNA sequence is important (Fig. 3.1B). These results suggest that the HEK293T extracts assemble in an RNA sequence dependent manner, even though they do not assemble in a comparable manner to HeLa.

In addition to the A^{min} substrate, I asked whether the $A^{3'}$ and full-length substrate yielded complexes when assembled upon by HEK293T nuclear extract. In the presence of the $A^{3'}$ substrate, HEK* was observed in +ATP conditions (Fig. 3.1C) as well as in -ATP conditions (data not shown) unlike that of $A^{3'}$ assembly in HeLa nuclear extract. However, the full-length substrate failed to assemble any complex including the newly observed A^{min} and $A^{3'}$-associated HEK* complexes (Fig. 3.1C).

Aiming to increase the activity of HEK293T extract, I adapted a whole cell extract (WCE) prep for HEK293T cells that was originally determined by Kataoka and Dreyfuss (**Appendix I – WCE Protocol**). Like the nuclear extract, I found that HEK293T whole cell extract also yielded a HEK* band when the A^{min} (-/+ATP) or $A^{3'}$ (+ATP) substrates were assembled upon, yet full-length pre-mRNA substrate did not yield a noticeable complex (Fig. 3.1D). Further, the use of the HEK293T whole cell extract did not increase complex assembly for any of the substrates.

Finally, I performed denaturing gel analysis overtime utilizing a full-length substrate to test whether splicing chemistry was possible with HEK293T nuclear and whole cell extracts providing spliceosomal components compared to HeLa nuclear extract. If the HEK293T derived complexes were inherently unstable and falling apart during assembly analysis (unstable in the native agarose gel matrix), but are capable of chemistry, we should observe splicing intermediates in the denaturing gel. When compared to the HeLa nuclear extract, neither the HEK293T nuclear nor whole cell extracts produced splicing intermediates (Fig. 3.1E). These results suggest that the lack of assembly observed on all tested RNA substrates in the HEK293T extracts was either because a critical spliceosome component associated with early assembly is absent or inactive, or because the extract contains a factor that inhibits assembly.

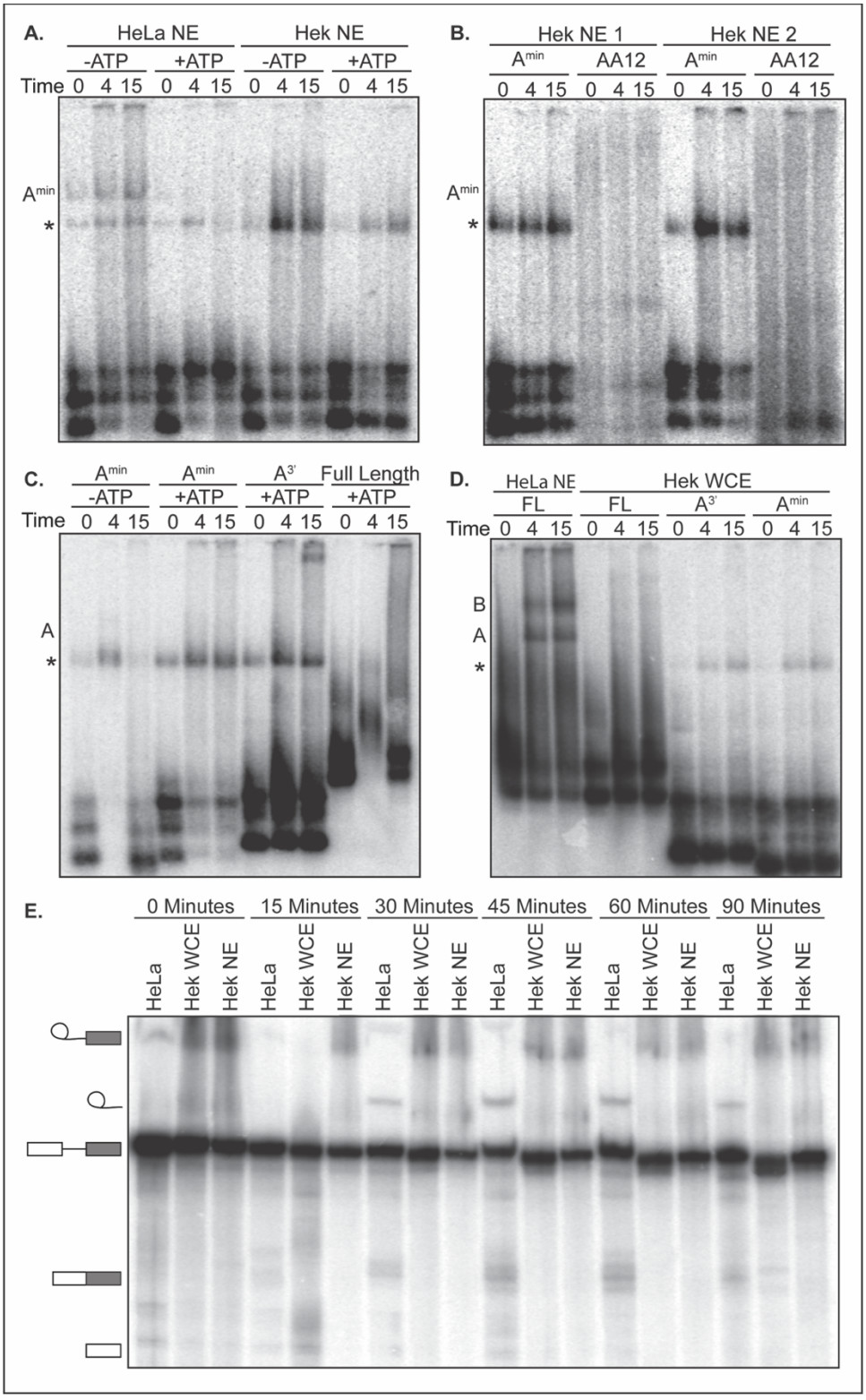

A.
HeLa NE
Hek NE
-ATP
+ATP
-ATP
+ATP
Time 0 4 15 0 4 15 0 4 15 0 4 15
Amin
*
B.
Hek NE 1
Hek NE 2
Amin
AA12
Amin
AA12
0 4 15 0 4 15 0 4 15 0 4 15
Amin
*
C.
Amin
Amin
A3'
Full Length
-ATP
+ATP
+ATP
+ATP
Time 0 4 15 0 4 15 0 4 15 0 4 15
A
*
D.
HeLa NE
Hek WCE
FL
FL
A3'
Amin
Time 0 4 15 0 4 15 0 4 15 0 4 15
B
A
*
E.
0 Minutes
15 Minutes
30 Minutes
45 Minutes
60 Minutes
90 Minutes
HeLa
Hek WCE
Hek NE

Figure 3.1: HEK293T nuclear and whole cell extract are not sufficient for A-complex assembly or splicing chemistry

A. Representative native gel analysis of *in vitro* spliceosome assembly in HeLa and HEK293T nuclear extract (NE) with truncated A^{min} substrate in -/+ATP conditions at the indicated timepoints. B. Representative native gel analysis of *in vitro* spliceosome assembly in two distinct HEK293T nuclear extract preps with truncated A^{min} substrate and AA12 substrate (scrambled sequence) in +ATP conditions at the indicated timepoints. C. Representative native gel analysis of *in vitro* spliceosome assembly in HEK293T nuclear extract with various RNA substrates in -/+ATP conditions at the indicated timepoints. D. Representative native gel analysis of *in vitro* spliceosome assembly in HeLa nuclear extract and HEK293T whole cell extract (WCE) with various RNA substrates in +ATP conditions at the indicated timepoints. E. Denaturing gel analysis of a radiolabeled full-length substrate over time in various extracts prepared from HeLa or HEK293T cells. Splicing intermediates are labeled on the left-hand side of the gel.

Is HEK293T nuclear or whole cell extract missing spliceosome components?

If the HEK293T extracts are missing an essential component, then supplementing with active HeLa nuclear extract at some ratio should provide the missing component and restore assembly. I performed assembly assays in which the nuclear extract providing the spliceosome components was some ratio of HeLa:HEK293T nuclear extract to determine how much active HeLa nuclear extract was necessary to restore activity. In addition, the total reaction volume usually used in assembly assays is 40% HeLa nuclear extract. To accommodate the possibility that dilution of our active components could negatively impact our results, I increased the total amount of combined extract to 60% of the total reaction volume.

Assembly on a full-length substrate at a ratio of 90/10 HeLa/HEK293T nuclear extract was required to restore assembly at 15 minutes. Both 70/30 and 80/20 HeLa/HEK293T nuclear extract were inactive (Fig. 3.2). The 70/30 ratio of HeLa/HEK293T nuclear extract is roughly equivalent to a normal assembly

reaction in which 40% of the total volume is active HeLa nuclear extract and therefore should provide sufficient components to accommodate spliceosome assembly. A similar observation when I used a combination of active HeLa nuclear extract and HEK293T whole cell extract, with a ratio of roughly 90/10 HeLa/HEK293T extracts required for complex assembly (data not shown). Taken together, these results suggest that the HEK293T extracts may contain something that prevents the assembly of spliceosomal complexes.

If HEK293T extracts contain an inhibitory molecule, then it should be separable by fractionation. I used small-scale SEC to fractionate HEK whole cell extract (**Appendix II, SEC Protocol**), and tested each fraction at 50/50 ratio with HeLa nuclear extract in assembly assays with both A^{min} and $A^{3'}$ substrates. I hypothesized that if a molecular inhibitor was present, only a subset of fractions would interfere with assembly and that we would then be able to determine the identity of such an inhibitor by sending that fraction for further testing. However, unlike the unfractionated HEK293T nuclear and whole cell extracts, none of the tested fractions affected $A^{3'}$ and A^{min} complex formation, (Fig. 3.3A and B). This result suggests that the inhibitory factor includes multiple components that were separated by fractionation or retained by the SEC column.

In these experiments, we also noticed with the A^{min} substrate formed a doublet at the HeLa/HEK* band in +ATP. It is possible that the previously observed HEK* band may be a separate complex from * band associated with HeLa nuclear extract. Notably, there is a small amount of this lower band in the HeLa only lane, but it is more intense in the HEK fraction samples (Fig. 3.3B).

Interpreting this finding is further compounded by the fact that we do not know what * band is.

The spliceosomes' ability to assemble on these substrates post-fractionation when supplemented with active HeLa nuclear extract suggests that fractionation of the HEK293T whole cell extract by S300 SEC was sufficient to separate the remaining fractions from whatever was inhibiting assembly with that entity likely in the column or void space. Taken together, these results suggest that HEK293T extract may contain a naturally occurring inhibitor that is preventing assembly of early spliceosomal complexes.

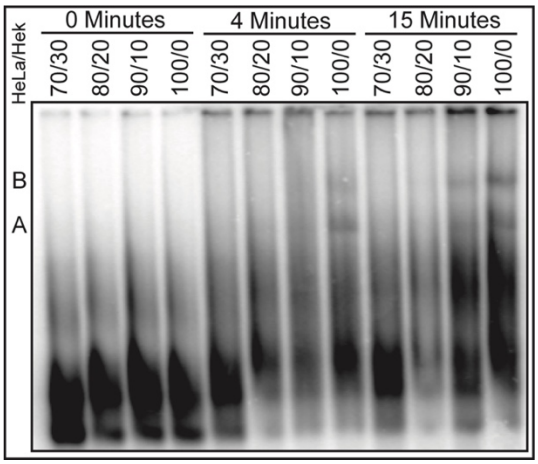


Figure 3.2: Rescue of HEK293T nuclear extract assembly by HeLa nuclear extract addition

Representative native gel analysis of *in vitro* spliceosome assembly in HeLa and HEK293T nuclear extracts mixed at various ratios on a full-length substrate in +ATP conditions at the indicated timepoints.

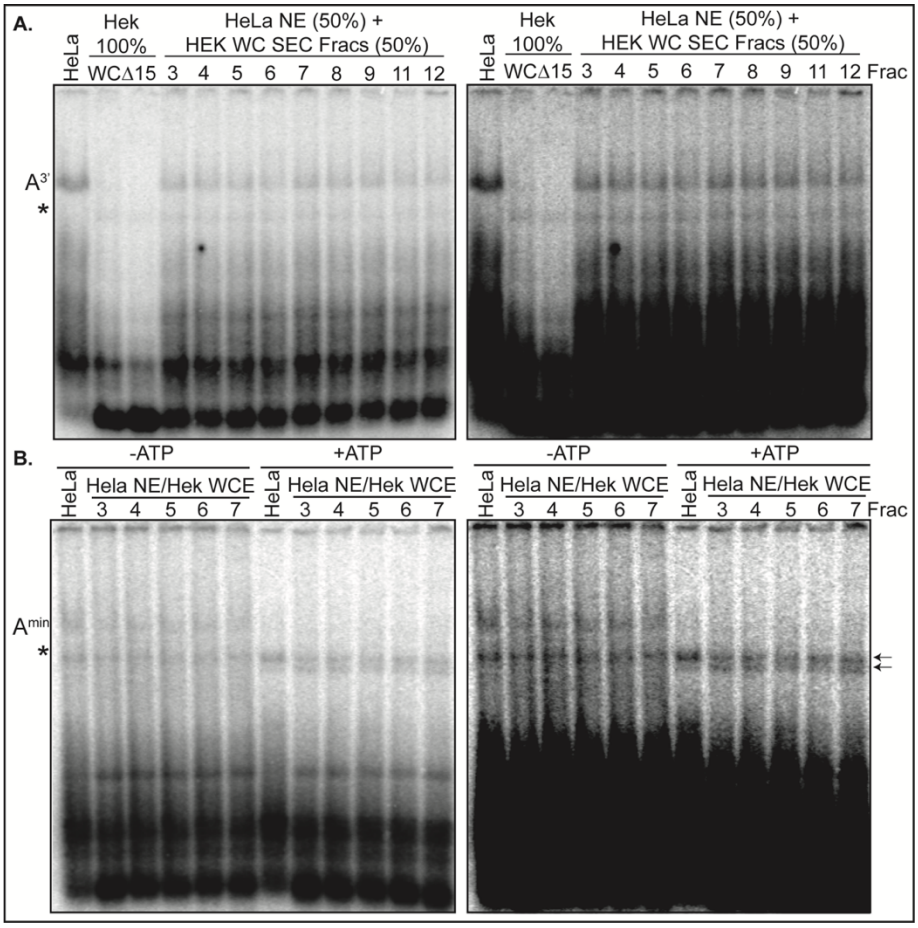


Figure 3.3: SEC fractionated HEK293T whole cell extract assembly

A. Representative native gel analysis of *in vitro* spliceosome assembly in HeLa and fractionated HEK293T whole cell extract with a $A^{3'}$ substrate in +ATP conditions at 15 minutes. The gels are identical with the intensity increased on the right side. B. Representative native gel analysis of *in vitro* spliceosome assembly in HeLa and fractionated HEK293T whole cell extract with an A^{min} substrate in -/+ATP conditions at 15 minutes. The gels are identical with the intensity increased on the right side.

Variability within HeLa cell lines – Do altered levels/migration of snRNAs contribute to nuclear extract inactivity?

In addition to testing HEK293T cells for splicing activity, I tested nuclear extract preps from different HeLa cell lines. Previously, we found that some preps result in nuclear extracts sufficient for assembly yet inefficient for splicing chemistry. In addition, we recently determined that genetic manipulation of an

active parental cell line can affect the activity of the derived line's nuclear extracts. The parental V5-SNRPB2 HeLa cell line was generated after LoxP recombination and puromycin selection against a population of cells (Khandelia et al., 2011; Kim, 2019). This active line was then further genetically altered by CRISPR of the endogenous DHX15 locus followed by expansion of single cell clones (Feng et al., 2022). The resulting cell lines are perfectly healthy in culture yet do not produce nuclear extract active for splicing *in vitro* (Fig. 3.4). To further confirm that there were no differences immediately post-purification, I measured UV absorbance and 280 nm of the parental and derived extracts to estimate overall protein/nucleic acid concentration. I found no significant difference between the protein concentrations eliminating the possibility that the inactive extract was too concentrated or dilute to support activity *in vitro* (data not shown).

I next asked whether the distribution of spliceosome-associated snRNAs is altered within the extracts that failed to assemble spliceosomal complexes. I hypothesized a problem with U2 snRNA specifically due to the lack of A-complex formation. To assess snRNAs, I separated the nuclear extracts on glycerol gradients in -/+ ATP conditions which allowed for separation of the snRNAs as well as distribution of the 12S and 17S U2 snRNPs (**Appendix I, Glycerol Gradients Protocol**). After fractionating the nuclear extract by glycerol gradient, I performed northern blot analysis using northern probes first optimized by Konaska and Sharp (Konarska and Sharp, 1986) (**Appendix** I, **Northern Blot Protocol**) to determine which gradient fractions contained the five snRNAs and how ATP affected the distribution of the snRNAs between the active and inactive HeLa cell lines (Fig. 3.5).

After separating the extract and probing for the snRNAs by northern, I found that the distribution of the snRNAs between the active and inactive HeLa nuclear extracts differed significantly (Fig. 3.6). When comparing the relative amounts of snRNAs between the active and inactive extract, all snRNAs appear decreased to some extent, however the levels of U2 snRNA over the gradient were significantly lower still (Fig. 3.7). This trend was observed across several replicates and different inactive nuclear extract preps (data not shown). The overall decreased snRNA levels could represent lower expression in the DHX15-degron line, or loss during the purification process, potentially at the chromatin release step, and both possibilities could be contributing to the inactivity of the extract. However the large decrease in the U2 snRNA specifically in comparison to the general decrease in the other snRNAs has yet to be determined.

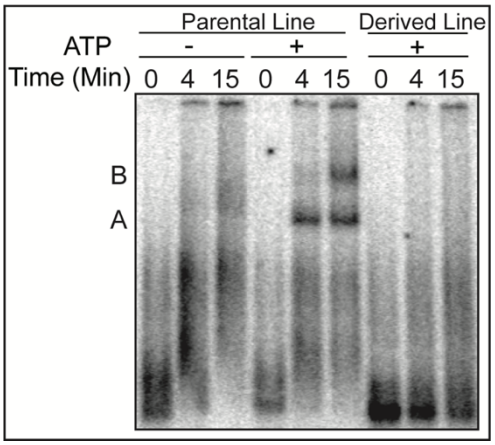


Figure 3.4: Cell lines derived from an active line does not ensure activity

Representative native gel analysis of *in vitro* spliceosome assembly in two HeLa nuclear extracts (parental HeLa extract (V5-snRPB2) and the derived HeLa extract (V5-snRNPB2, DHX15-HA-dTAG) on a full-length substrate in -/+ATP conditions at the indicated timepoints.

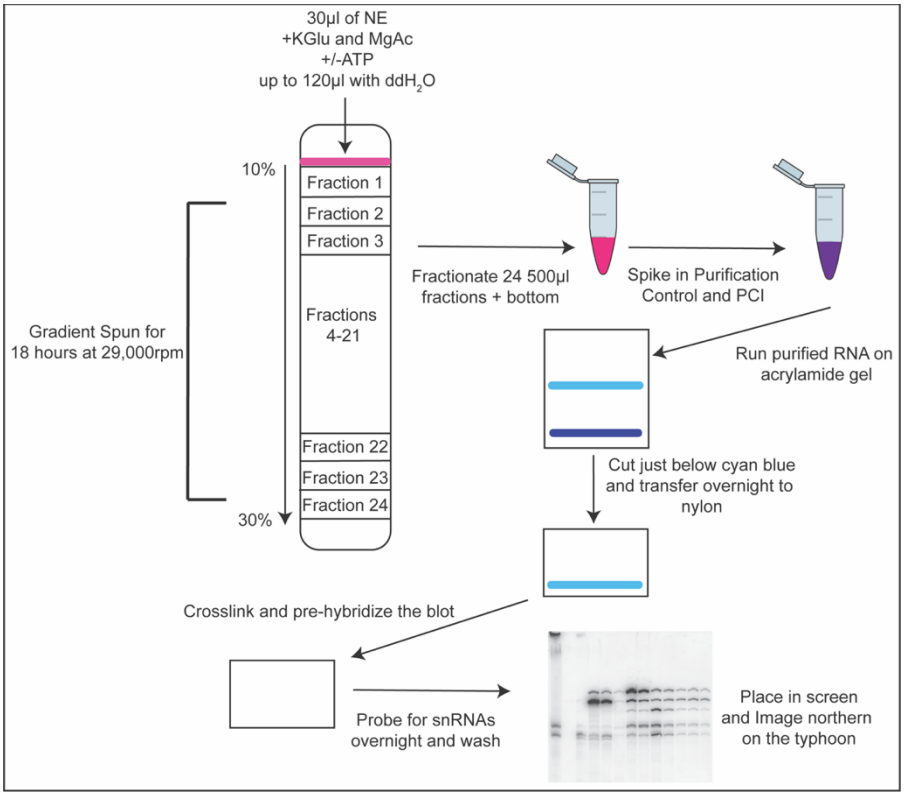


Figure 3.5: Glycerol Gradient and Northern Blot Protocol Overview

Purification schematic of glycerol gradient and Northern Blot analysis. After nuclear extract is loaded onto a glycerol gradient, it is spun at high speed overnight. The gradient is then fractionated, and the RNA is purified and run on a denaturing acrylamide gel. The RNA is then transferred to nylon, crosslinked, and exposed to anti-sense radiolabeled probes against the five snRNAs.

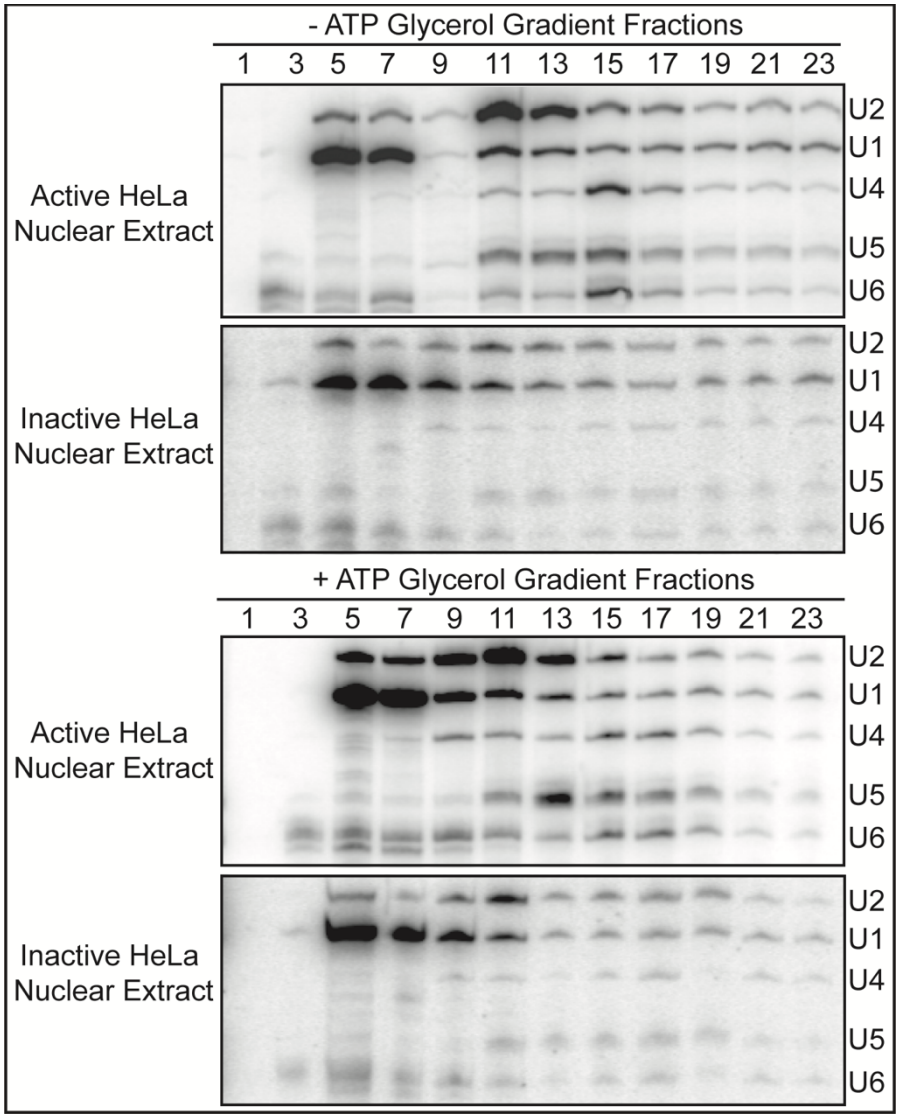


Figure 3.6: Northern Blot analysis of active and inactive HeLa nuclear extract across a glycerol gradient

Comparison of two large scale prep HeLa nuclear extracts: active parental HeLa nuclear extract (V5-snRPB2) and the inactive derived HeLa extract (V5-snRNPB2, DHX15-HA-dTAG). Both extracts were diluted in splicing conditions and purified according to the protocol in **Appendix I.**

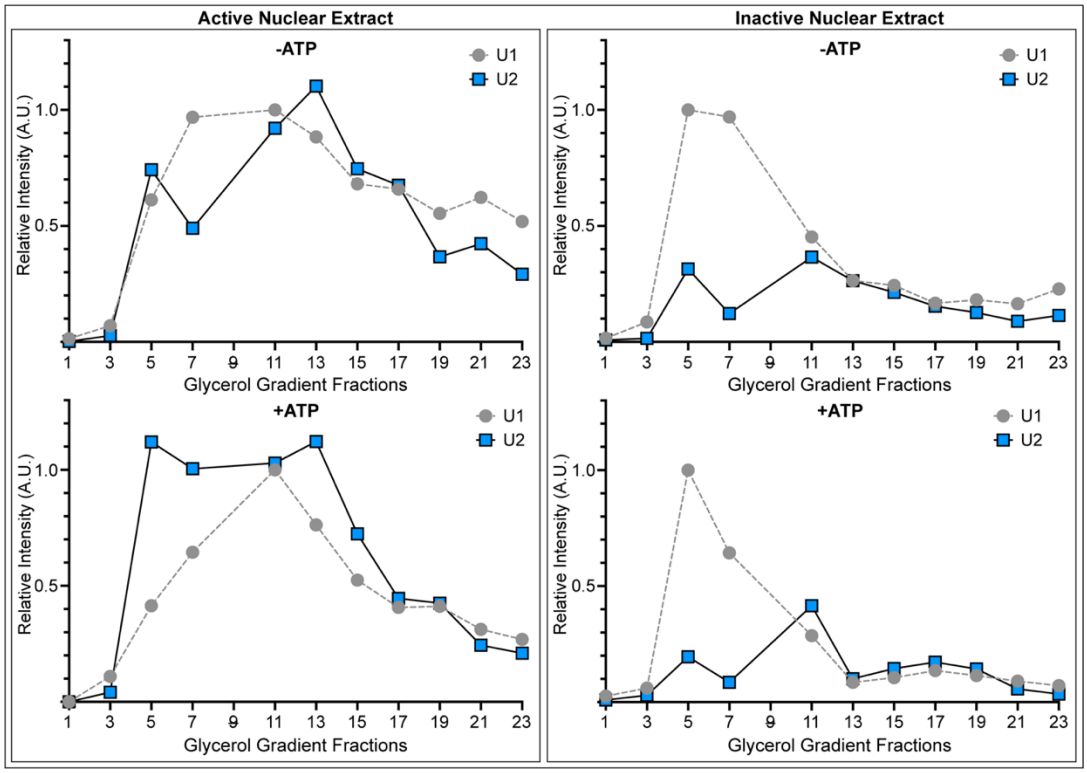


Figure 3.7: Relative Intensity of U1 and U2 Northern Blot bands of active and inactive HeLa nuclear extracts across a glycerol gradient

Quantification of Figure 3.6 Northern blots of two HeLa nuclear extracts: active parental HeLa nuclear extract (V5-snRPB2), normalized to the highest intensity of U1, Fraction 11 and the inactive derived HeLa extract (V5-snRNPB2, DHX15-HA-dTAG), normalized to the highest intensity of U1, Fraction 5. Fraction 9 values were removed from all datasets due to the observed underloading of the lane in Fig. 3.6 -ATP Active nuclear extract blot.

Discussion

Since 1983, the splicing community has heavily relied on HeLa nuclear extracts to provide the spliceosomal components and associated proteins for *in vitro* splicing assays. However, to this day we have yet to fully unlock key steps in early spliceosome assembly, such as branch recognition, and this may be a result of the lack of variability amongst the cells we utilize to provide our spliceosomal components. In addition, it has become increasingly apparent that

different cell types will yield different splicing landscapes and alternative splicing profiles which could potentially help to answer our remaining splicing questions.

The splicing community has long relied on HeLa nuclear extract for *in vitro* splicing assays. Yet, splicing labs have long known of the inherent variability in activity between nuclear extract preps within HeLa itself, no one has published a study regarding the reasoning behind the variability nor has a review of tested protocols for making active extract been collected and published. As a result, we as a collective group have accepted that one HeLa nuclear prep will yield robust activity in both assembly and splicing assays whereas a second will be efficient for assembly yet inefficient for chemistry.

The work I have presented in this chapter represents the first steps in trying to answer why HeLa extracts are variable and why other cell lines have not yet produced extracts for wide use. One potential reason for HeLa inactivity is upon purification, the U2 snRNA is lost, resulting in an inactive final product and contributing to changes in distribution of snRNAs across the glycerol gradient. However, further investigation will be required to determine if optimization of the current protocols at the chromatin release step can restore activity to extracts prepared from HeLa cells that currently generate inactive nuclear extract. It is still unclear how genetic manipulation of some cell lines disturbs activity of the snRNAs, however further analysis will need to be done to determine other factors that could be contributing to this issue.

Taken together, *in vitro* splicing has allowed scientists to answer many fundamental questions about splicing, however, it has relied heavily on a single cell line to provide spliceosome components. The expansion into other cell lines

to provide these components may be necessary to expand the kinds of questions we can ask with the *in vitro* assembly system, yet, as my work has shown, it is not as simple as just growing cells and making a nuclear extract. However, I believe it is important that we put in the effort to do so. RNA sequencing and other robust single cell technologies have shown how different cells are from one another and splicing changes infer that the spliceosome is affected in some way which may be possible to study mechanistically if we expand our system into new cellular backgrounds. The *in vitro* system certainly still has much to teach us, however without the addition of new nuclear extract technologies to support it, unlocking remaining questions in splicing and specifically those associated with early complex assembly may not be possible.

References

Chou, M.Y., Underwood, J.G., Nikolic, J., Luu, M.H., Black, D.L., 2000. Multisite RNA binding and release of polypyrimidine tract binding protein during the regulation of c-src neural-specific splicing. Mol Cell 5, 949-957.

Dignam, J.D., Lebovitz, R.M., Roeder, R.D., 1983. Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. Nucleic Acids Res. 11, 1475-1489.

Feng, Q., Krick, K., Chu, J., Burge, C.B., 2022. Splicing quality control mediated by DHX15 and its G-patch activator, SUGP1. bioRxiv, 2022.2011.2014.516533.

Kataoka, N., Dreyfuss, G., 2004. A simple whole cell lysate system for in vitro splicing reveals a stepwise assembly of the exon-exon junction complex. J Biol Chem 279, 7009-7013.

Khandelia, P., Yap, K., 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.

Kim, J., 2019. A Pipeline for Tagging snRNP Associated Proteins in HeLa. , Molecular, Cell and Developmental Biology. University of California, Santa Cruz, CA.

Konarska, M.M., Sharp, P.A., 1986. Electrophoretic separation of complexes involved in the splicing of precursors to mRNAs. Cell 46, 845-855.

Landry, J.J.M., Pyl, P.T., Rausch, T., Zichner, T., Tekkedil, M.M., Stütz, A.M., Jauch, A., Aiyar, R.S., Pau, G., Delhomme, N., Gagneur, J., Korbel, J.O., Huber, W., Steinmetz, L.M., 2013. The Genomic and Transcriptomic Landscape of a HeLa Cell Line. G3 Genes|Genomes|Genetics 3, 1213-1224.

Lee, K.A., Bindereif, A., Green, M.R., 1988. A small-scale procedure for preparation of nuclear extracts that support efficient transcription and pre-mRNA splicing. Gene Anal Tech 5, 22-31.

Newnham, C.M., Query, C. C., 2001. The ATP requirement for U2 snRNP addition is linked to the pre-mRNA region 5’ to the branch site. RNA 7, 1298-1309.

Padgett, R.A., Hardy, S.F., Sharp, P.A., 1983. Splicing of adenovirus RNA in a cell-free transcription system. Proc Natl Acad Sci U S A 80, 5230-5234.

Chapter 4: Isolation of spliceosomal complexes via tripartite purification

Introduction

Early spliceosome complexes are highly dynamic and short-lived and as a result, make structural studies difficult. Recent advances in cryo-EM technologies have allowed great leaps our understanding of what is happening at a molecular level during early assembly particularly when the U2 snRNP engages an intron. However, differences between these structures provide very tantalizing evidence for large rearrangements that must occur to form a competent A-complex, yet we have not found a way to purify and study transient intermediate complexes.

Current spliceosome structure protocols require scientists to grow enormous quantities of cells to overcome problems of heterogeneity and instability associated with purifying dynamic and short-lived complexes. For reference, when we generate HeLa nuclear extract, a single prep will yield between 10 and 20mL of nuclear extract from 10L suspension culture of HeLa cells and requires roughly three weeks of labor-intensive work from start to end. Other groups have pushed this even further. For instance, the Luhrmann group utilized 12mL splicing reactions with 40% HeLa nuclear extract (4.8mL) to generate enough purified sample to solve a B-complex structure (Deckert et al., 2006) or 1.5 liters of nuclear extract to solve the structure of the 17S U2 snRNP (Zhang et al., 2020) Even with these efforts, the overall resolution of the solved structures is variable leading to ambiguity in the more dynamic regions of the given complex. My thesis work focused on understanding how the presence of ATP drives the rearrangements necessary for early spliceosome assembly as

well as the role of DHX15 during this process. To that end, I wanted to develop a protocol that would allow me to study transient complexes and interactions such as those that would incur DHX15 activity during early spliceosome assembly.

To overcome the limitations, I set out to devise a robust purification protocol for spliceosome complexes in transition from E- to A-complex (Fig. 4.1). Within HeLa nuclear extract, it is widely accepted that only a subpopulation of the spliceosomal components is competent for *in vitro* splicing. For instance, if we utilize a tagged pre-mRNA in a system without any other biochemical purification, most of the pre-mRNA will become trapped in some non-active state. The same issue arises when we rely on only protein tags as many of our tagged proteins get split across a variety of splicing intermediates therefore making it difficult to separate out the complex of interest. Therefore, individual tagging systems result in a large amount of heterogeneity between active and inactive complexes, which greatly complicates structures determination. Therefore, I devised a small scale purification scheme which combines multiple tags and affinity steps to yield early spliceosome complexes that represent active conformations and stable for downstream analyses (Ilagan and Jurica, 2014; Jurica and Moore, 2002; Kim, 2019; Maul-Newby et al., 2022; Nilsen, 2013). The protocol can be manipulated by changing the pre-mRNA substrate to isolate different spliceosome assembly intermediates.

My protocol starts with pre-mRNA substrates engineered to contain MS2-hairpins for affinity purification. To increase the stability of transient interactions, I incorporated the crosslinkable element 4-thiouridine (4sU) into the pre-mRNA transcript at a rate of one crosslinkable element per transcribe. And finally to

ensure we enrich for pre-mRNAs containing U2 snRNP, I utilized an engineered active V5-SNRPB2 HeLa nuclear extract, described in Chapter 3 of this thesis (Kim, 2019). Together, these steps allow us to not only separate complexes of interest but increase the likelihood of capturing rare or transitional interactions during early spliceosome assembly. A step-by-step protocol can be found in **Appendix II** of this dissertation.

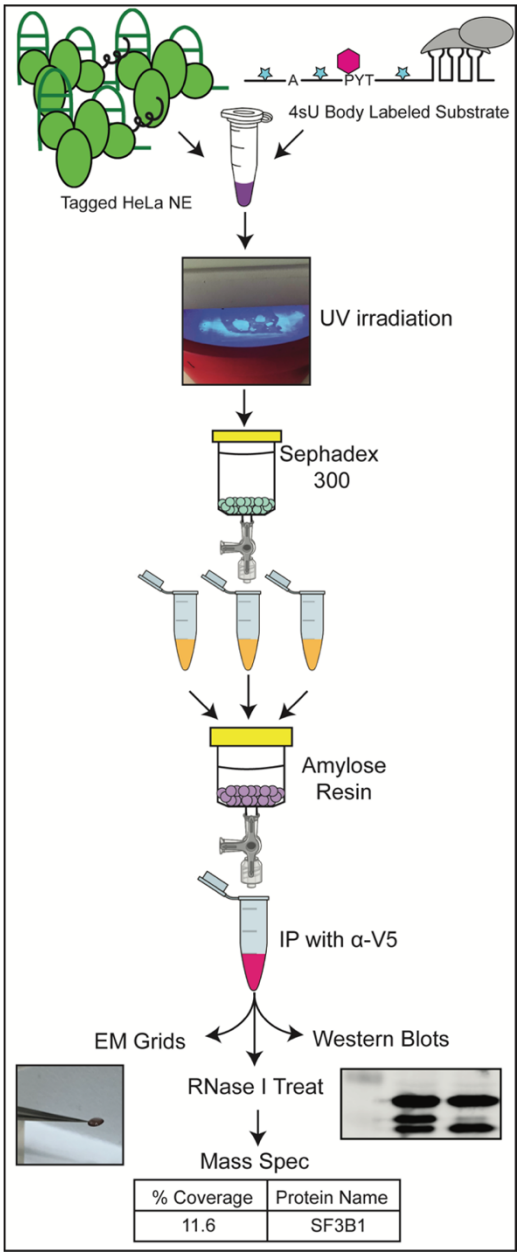


Figure 4.1: Early Spliceosome Complex Purification Schematic

To purify a spliceosomal complex of interest, a V5-SNRPB2 HeLa NE (green) is combined with a 4sU body labeled pre-mRNA substrate bound to MS2-MBP (grey) under normal assembly conditions. Post-assembly these reactions are exposed to UV light (302nm) and separated via a Sephadex 300 size exclusion column. The fractions of interest are pooled and placed over an amylose column and eluted with maltose. This elution is then IP'd against V5 and eluted depending on the final downstream analysis.

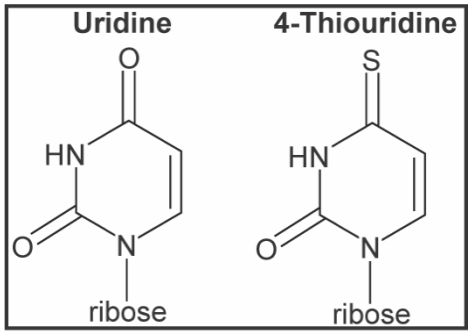


Figure 4.2: Uridine versus 4-thiouridine

Structural difference between Uridine and crosslinkable nucleotide analog, 4-thiouridine in which the carbon 4 oxygen is replaced with a sulfur.

Results

Adding a crosslinkable element to the pre-mRNA substrate does not inhibit assembly and results crosslinks

Previous efforts in the lab to capture early splicing complexes indicated that they are inherently unstable potentially due to short lived interactions and the absence of stabilizing structural elements such as the branch helix((Roybal and Jurica, 2010) unpublished data from former graduate student, Andrew MacRae). I hypothesized that addition of a crosslinkable element will improve stability and therefore incorporated 4sU into truncated pre-mRNAs containing MS2 hairpins and body labeled with ^{32}P. The truncated substrate ensures that assembly will not go past A-complex. Immediately after assembly and crosslinking, the complexes are separated by size to effectively remove inactive nuclear extract

components and unincorporated pre-mRNA (Fig.4.1). We used S300 resin because it provides sufficient separation of early spliceosomal complexes, as observed in the two distinct peaks of radioactivity across the SEC fractions (Fig. 4.3A). In addition, we also observed that complexes assembled and crosslinked on a 4sU-containing pre-mRNA consistently increased the amount of radioactivity recovered in the early SEC fractions, consistent with stabilizing more flexible and labile complexes in comparison to the non-crosslinkable control. To further enrich for early spliceosome, only the first five fractions over background were collected and pooled to be taken into the next step of the purification process (Fig. 4.3, grey boxes).

To determine the efficacy of the crosslinking, SEC fractions were treated with RNase T1, an endoribonuclease that specifically targets guanosine residues in single stranded RNA (Fig. 4.3B, Top RNase T1). If a crosslink is present, it will protect the RNA from digestion by the RNase T1 and evidence of a crosslink will be observed by an upwards shift in the gel of the body labeled substrate. In SEC fractions 9-11 as well as the 4sU positive splicing reaction, we observe a robust band in the radiograph at roughly 55kD (naked RNA substrate runs at roughly 37kD) indicating the presence of at least one crosslink (Fig. 4.3B, Bottom). Taken together, these results indicate that the pre-mRNA is incorporated into a large complex and that the 4sU crosslinked to other entities.

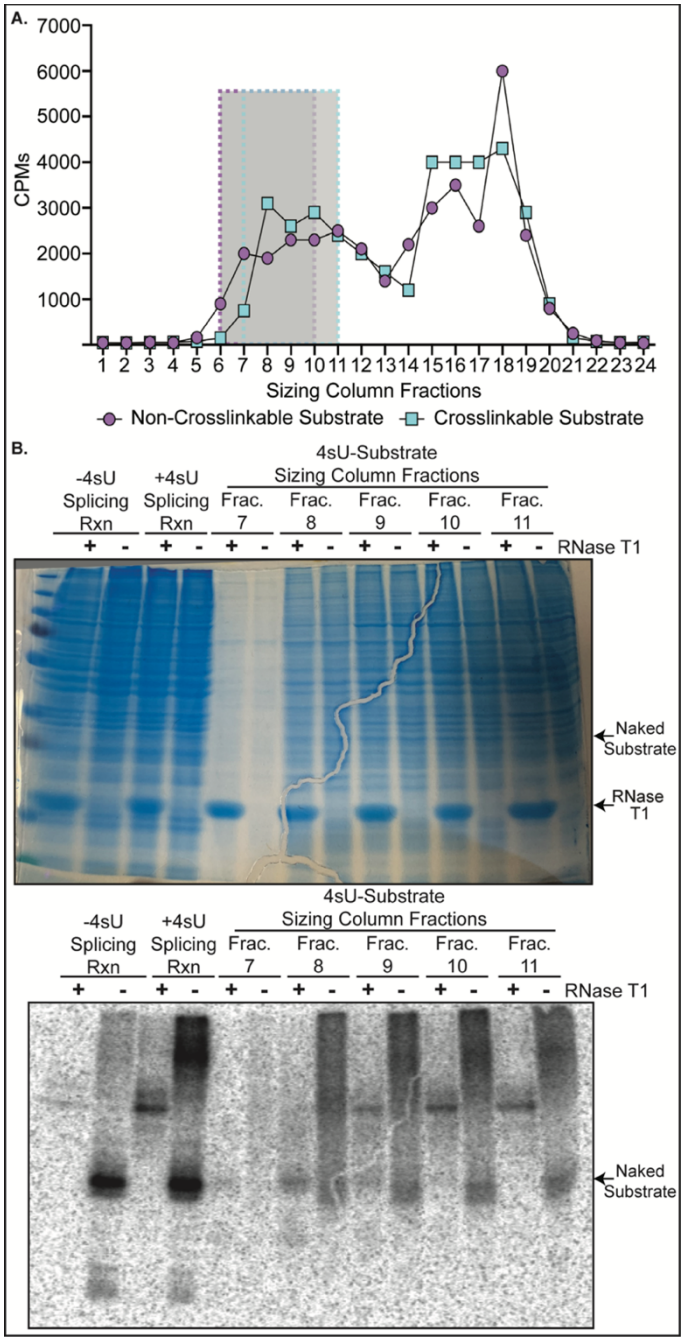


Figure 4.3: SEC Fractions and Crosslink Confirmation by RNase T1

A. CPM vs. SEC fraction number plotted for both body labeled pre-mRNA substrates with (cyan) or without (purple) 4sU. Fractions under the grey boxes were pooled for further purification. B. SEC fractions treated with and without RNase T1. Top: PAGE gel analysis of proteins associated with the SEC fractions compared to the initial splicing reactions. Bottom: Same gel after being placed in a phosphor screen to detect body labeled substrates.

Purification of crosslinked complexes results in higher yield in final immunoprecipitation elutions

To ensure enrichment of complexes bound to pre-mRNA, I passed the first radioactive peak over an amylose to bind the MS2-MBP bound to the MS2 hairpins. The resulting eluent with maltose was incubated with magnetic beads bound to a V5 antibody to isolate complexes that also contain U2 snRNP tagged with V5-SNRBP2. To maintain the complexes in a native state for future structural studies, we released them by TEV-protease cleavage of a TEV linker between V5 and SNRBP2 (**Appendix II)**. In the TEV eluate, the amount of pre-mRNA isolated is higher with the +4sU pre-mRNA when compared to the non-crosslinkable control (-4sU, Fig. 4.4A), indicating that crosslinking increases the yield of complexes containing both pre-mRNA and U2 snRNP.

With a positive crosslinking hit in the RNase T1 digestion (Fig. 4.3), I looked for proteins that current structures would predict to be near the pre-mRNA and hypothesized that if a crosslink had occurred, I would observe a shift in molecular weight by Western blot. PHF5A became a tantalizing target as it has been shown to be located near the pre-mRNA in several spliceosome structures (Cretu et al., 2021; Dybkov et al., 2006; Zhang et al., 2020). To determine if PHF5A had been crosslinked, Western analysis was performed and upon blotting for PHF5A in the Elution 2 samples, we observe the presence of higher molecular weighted bands in the crosslinked samples (Fig. 4.4B), including a band at roughly 55kD which would equate to the size we would expect if the pre-mRNA and PHF5A were crosslinked together.

As a control for crosslinking, I blotted for SF3B3 which has been observed in spliceosome structures to be far away from the intron therefore

would not be expected to crosslink to the pre-mRNA (Cretu et al., 2021; Dybkov et al., 2006; Zhang et al., 2020). Further, SF3B3 is a component of the 17S U2 snRNP and is an excellent readout for the overall purification between the different samples (Fig. 4.4C). Western blot analysis of Elution 2 blotting for SF3B3 resulted in no discernable difference between the crosslinked and non-crosslinked sample with no crosslinks detected.

Notably, SNRPB2-V5, PHF5A and SF3B3 were all detected in both the crosslinked and non-crosslinked samples, which suggests that even without crosslinking, complexes with both pre-mRNA and U2 snRNP components remain stable throughout the purification scheme.

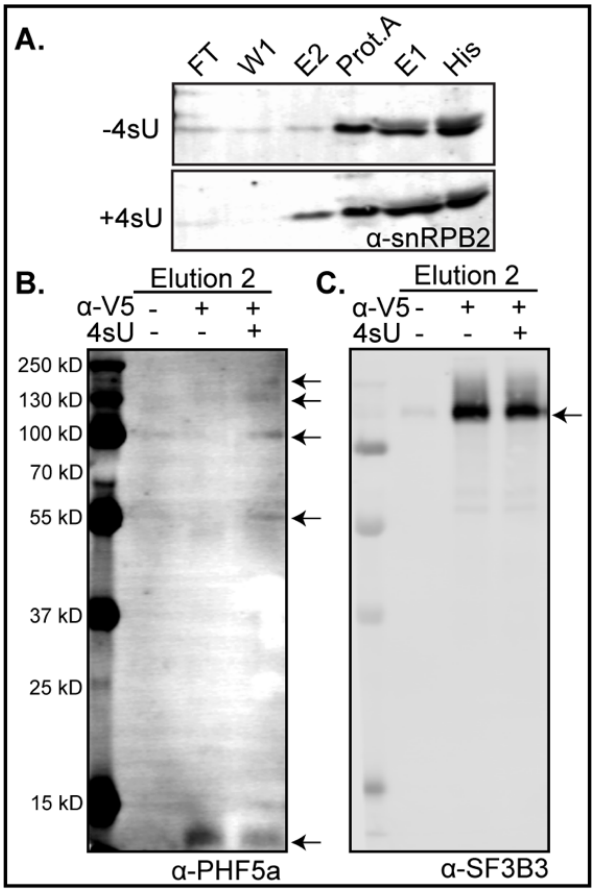


Figure 4.4: 4sU crosslinking increases the concentration of SNRPB2-incorporated complexes post-purification

Western blot analysis of IP samples post-elution. A. Western blot analysis of various timepoints from the IP blotting for SNRPB2 (**Appendix II**). Lanes are as follows: FT – Flow Through, W1 – Wash 1, E2 – Elution 2, Prot.A – Protein A Beads, E1 – Elution 1 and His – His Beads. B. Western blot analysis of Elution 2 blotting for PHF5A. Arrows are indicative of bands in crosslinked lane. C. Western blot analysis of Elution 2 blotting for SF3B3. Arrows are indicative of bands in the crosslinked lane.

Discussion

Purification of early spliceosomal complexes has proved to be difficult for a variety of reasons such as being highly dynamic, short-lived, and hard to slow down/stop. By combining a crosslinkable substrate with HeLa nuclear extract that contains V5-SNRBP2 in addition to several stringent purification steps, we were able to successfully purify intact spliceosome complexes. Next steps are to analyze the complexes by mass spectrometry and image them by electron microscopy. Further, due to the ease of reproducibility in terms of how few cells and overall manpower this protocol requires, we may observe enough complexes to fill in current gaps in knowledge such as where transient proteins such as helicases like DHX15 interact in addition to the types of complexes they may associate with.

The protocol is flexible for purifying other complexes by substituting the pre-mRNA substrate. For instance, both A^{min} or $A^{3'}$ with MS2 hairpins and 4sU could be used to assemble splicing complexes in -/+ ATP conditions to establish how they differ from one another. It also provides a means to determine the composition of * band and to ask whether the HEK* band observed in the native gels from Chapter 3 is a different entity.

References

Cretu, C., Gee, P., Liu, X., Agrawal, A., Nguyen, T., Ghosh, A., Cook, A., Jurica, M., Larsen, N., Pena, V., 2021. Structural basis of intron selection by U2 snRNP in the presence of covalent inhibitors. Nat Commun 12, 4491.

Deckert, J., Hartmuth, K., Boehringer, D., Behzadnia, N., Will, C.L., Kastner, B., Stark, H., Urlaub, H., Luhrmann, R., 2006. Protein composition and electron microscopy structure of affinity-purified human spliceosomal B complexes isolated under physiological conditions. Mol Cell Biol 26, 5528-5543.

Dybkov, O., Will, C.L., Deckert, J., Behzadnia, N., Hartmuth, K., Luhrmann, R., 2006. U2 snRNA-protein contacts in purified human 17S U2 snRNPs and in spliceosomal A and B complexes. Mol Cell Biol 26, 2803-2816.

Ilagan, J., Jurica, M., 2014. Isolation and accumulation of spliceosomal assembly intermediates. Methods Mol Biol 1126, 179-192.

Jurica, M.S., Moore, M.J., 2002. Capturing splicing complexes to study structure and mechanism. Methods 28, 336-345.

Kim, J., 2019. A Pipeline for Tagging snRNP Associated Proteins in HeLa. , Molecular, Cell and Developmental Biology. University of California, Santa Cruz, CA.

Maul-Newby, H., Amorello, A., Sharma, T., Kim, J., Modena, M., Prichard, B., Jurica, M., 2022. A Model for DHX15 Mediated Disassembly of A-Complex Spliceosomes. RNA, rna.078977.078121.

Nilsen, T.W., 2013. Detecting RNA–Protein Interactions by Photocross-Linking Using RNA Molecules Containing Uridine Analogs. Cold Spring Harbor Protocols 2013, pdb.prot073569.

Roybal, G.A., Jurica, M.S., 2010. Spliceostatin A inhibits spliceosome assembly subsequent to prespliceosome formation. Nucleic Acids Res 38, 6664-6672.

Zhang, Z., Will, C.L., Bertram, K., Dybkov, O., Hartmuth, K., Agafonov, D.E., Hofele, R., Urlaub, H., Kastner, B., Luhrmann, R., Stark, H., 2020. Molecular architecture of the human 17S U2 snRNP. Nature 583, 310-313.

Appendix I:

Chapter 3: Creation of splicing nuclear extracts from non-HeLa cell lines and HeLa derived lines

Mini preparation of nuclear extracts from HeLa / Hek293T cells grown in monolayer - adapted from Lee et al, Gene Anal Tech 5: 22-31 (1988)

All buffers used in this protocol should be ice-cold, Table AI.2

1 - Wash the cells grown in at least 3 15-cm dishes at 80% confluence (~20 x 10^6 cells/plate) with ice-cold 1X PBS.

2 - Harvest the cells using a rubber policeman in 1-4 ml of ice-cold 1X PBS / plate, pool together in a 15 ml Flacon tubes and spin for 5 minutes at 2000 rpm in a table-top centrifuge at 4˚C.

3 - Discard the supernatant and note the packed cell volume (PCV). Work on ice from now on.

4 - Resuspend in 1 PCV of Buffer A, transfer to a single Eppendorf tube and allow the cells to swell on ice for 15 minutes.

5 - Homogenize your cells based on cell type: See Table AI.1
*I recommend using the Dounce protocol for HeLa preps of more than 3 15-cm plates

6 - Centrifuge the cell homogenate for 20 seconds in a microcentrifuge at full speed (12,000xg) at 4˚C. Discard the supernatant.

7 - Resuspend the nuclear pellet in 2/3 PCV of Buffer D. Incubate on ice at 4˚C with stirring for 30 minutes.

- The addition of Buffer D will immediately precipitate the chromatin. If you are following the needle prep protocol, no Buffer C should be required. However, if you are releasing many nuclei (over 500µl) I suggest adding Buffer C prior to adding Buffer D dropwise. The ratio of Buffer C to Buffer D should be 1/3 Buffer C to 2/3 Buffer D with the final volume equaling the 2/3 PCV.
- To help release the chromatin without forming a “goo ball”, I use a 5ml beaker with a mini stir bar that is stirring rapidly. I then add in the Buffer D dropwise to slowly precipitate the Buffer D and allow for full incorporation. Once fully added, I then start the 30-minute incubation. By the end of the 30-minute time-period, the extract will be viscous but pipette-able.

8 - Pellet the nuclear debris by spinning at 12000xg in a microcentrifuge at 4˚C for 30 minutes.

9 - Dialyze the supernatant against Buffer E for two hours at 4˚C. Make sure to pipette up and down inside of the dialysis cups every 30 minutes. Eliminate precipitates by centrifugation at 12,000xg in a microcentrifuge at 4˚C for 15 minutes post-dialysis. Aliquot and store nuclear extract at -80˚C. I tend to aliquot amounts needed per gel to avoid freeze thaw. Extract will lose activity with each thaw with a maximum of two thaws.

Table AI.1 – Cell Type Homogenization Protocol

Hela – Needle Prep	Hela – Dounce Prep	Hek293T – Dounce Prep
Homogenize the cells by rapidly pushing them through a narrow-gauge hypodermic needle as follows: * Attach a 25-g needle to a 1 ml hypodermic syringe. * To avoid air bubbles, fill the needle and syringe with buffer A and displace the buffer from the syringe as fully as possible. * Draw the cell suspension slowly into the syringe. * Eject the cell suspension from the syringe with a single rapid stroke. * Repeat the last two steps five times. * Check that the cell lysis is 80% or more by looking at the microscope.	Homogenize the cells by douncing with a 2ml glass Dounce with a tight pestle that have been baked and chilled on ice: * Transfer the cells to the Dounce * Dounce the cells 30-40 times with swift steady strokes on ice * Check that the cell lysis is 80% or more by looking at the microscope.	Homogenize the cells by douncing with a 2ml glass Dounce with a tight pestle that have been baked and chilled on ice: * Transfer the cells to the Dounce * Dounce the cells 30-40 times with swift steady strokes on ice * Check that the cell lysis is 80% or more by looking at the microscope.

Table AI.2 – Mini-NE Prep Buffers

	Tris, pH 7.9	$MgCl_2$	KCl	EDTA	Glycerol	H_2O	DTT	PMSF
Buffer A	10mM	1.5mM	10mM				1mM	
10 mL	100 µl 1M	15 µl 1M	100 µl 1M			9.775 mL	10 µl 1M	
Buffer C	20mM	1.5mM	20mM		25%		1mM	0.5mM
10 mL	200µl 1M	15µl 1M	200µl 1M		2.5mL	7.275 mL	10µl 1M	
Buffer D	20mM	1.5mM	420mM	0.2mM	25%		1mM	0.5mM
10 mL	200µl 1M	15µl 1M	840µl 5M	4µl 0.5M	2.5mL	6.431 mL	10µl 1M	
Buffer E	20mM		100mM	0.2mM	20%		1mM	0.5mM
10 mL	200 µl 1M		1mL 1M	4µl 0.5M	2mL	6.786 mL	10µl 1M	

*Tris can be substituted for HEPES, pH 7.9 as needed for downstream applications

HEK293T Whole Cell Splicing Extract Protocol
Adapted from Kataoka and Dreyfuss, 2004

Can be used for other cell lines besides HEK293T

Buffer E
20 mM Tris, pH 7.9
100 mM KCl
0.2 mM EDTA
20% Glycerol
1mM DTT

2-3 10-15cm plates ~80% confluency

1. Wash plates 2x with 10mL ice-cold 1X PBS
2. Using a rubber policeman, scrape cells off the plate and rinse and collect with 5ml ice-cold 1X PBS
3. Combine all the cells into a 15mL falcon and spin down at 3000xg for 5 minutes at 4˚C.
4. Aspirate off all the 1X PBS without disturbing the cell pellet.
5. Resuspend the pellet in ice-cold Buffer E: about 200µl per plate
6. Sonicate 3x in cold room (30% power, each continuous 5 second bursts. 30 seconds on ice between sonication bursts)
7. Transfer the sample to Eppendorf tubes
8. Spin at max speed at 4˚C for 20 minutes
9. Remove the supernatant away from the cellular debris and aliquot into whatever size is needed.
10. Store at -80˚C until use.

Glycerol Gradient Protocol – 10-30% 12 mL gradients SW-41 Rotor

Note: HeLa nuclear extract is made at a final glycerol concentration of ~20%. Therefore, it needs to be diluted to effectively run a gradient.

Gradient Prep

10X Gradient Buffer
1.5M KCl
200mM Tris, pH 7.9
20 mM MgAc

10% Glycerol – 4˚C
10% glycerol
150 mM KCl
20mM Tris, pH 7.9
2mM MgAc
*Filter sterilize 0.22µM

30% Glycerol – 4˚C
30% glycerol
150 mM KCl
20mM Tris, pH 7.9
2mM MgAc
*Filter sterilize 0.22µM

Gradient glycerol buffers can be made utilizing the 10X gradient buffer to 1X in glycerol. It can be useful to start with 50% glycerol to then be diluted as necessary.

Note: Gradients should be poured and mixed no earlier than 1 hour prior to loading and starting the ultracentrifugation run.

To a 9/16 x 3.5-inch (12mL) Seton open-top polyclear centrifuge tube (cat.7030) approximately 6.1mL of 10% glycerol was added (approximation made utilizing SW41 marking block) utilizing a 10mL syringe connected to a metal 4-inch cannula. Ensure that there are no bubbles present. If so, carefully remove before proceeding.

Add another 6.1mL of the 30% glycerol by carefully submerging the cannula to the bottom of the tube and floating the 10% glycerol ensuring no bubbles are introduced and minimal mixing of the two percentages.

Once poured, remove the cannula, and carefully add the cap by pointing the cap hole “up” and ensuring that it is the last portion of the cap to be inserted therefore preventing air from being trapped inside the gradient and excess air/glycerol escaping through the hole in the lid.

Gradient Mixer

Once all gradients are poured and capped, carefully place all gradients into the magnetic rack and proceed to the gradient mixer.

Turn on the Biocomp Piston Gradient Fractionator and select the program: SW41 10-30% w/v

Place the magnetic rack onto the platform and select go.

Once done rotating, place at 4˚C until use.

Ultra-Centrifuge

Prior to loading samples into the gradients, pre-chill the SW41 rotor and buckets at 4˚C for at least 4 hours.

Gradients should gently be placed into the ultracentrifuge buckets and the samples gently loaded on the very top of the gradient.

All buckets should be loaded and fitted onto the rotor according to manufacturer guidelines. In addition, all buckets must be weighed to ensure equal weight.

To separate the 12S and 17S U2 snRNPs from one another, the ultra should be run at 29,000 rpm for 18 hours at 4˚C.

Fractionation

After the run is complete, carefully place all buckets onto ice.

Fractionate each gradient into 500µL fractions starting from the top and make sure to save anything at the bottom of the gradient if there is an excess.

Northern Blot Protocol

Antisense snRNA Probe Transcription Recipe

ATP (10 mM)	2µl
CTP (10 mM)	2µl
UTP (10 mM)	0.25µl
GTP (10 mM)	2µl
DTT (1M)	0.5µl
10x txn buffer	5µl
DNA (total of 40ng/µl)	
U1 snRNA - (pBSU1 cut w/ EcoR1, T3 Pol)	
U2 snRNA - (pBSU2 cut w/BamH1, T7 Pol)	
U4 snRNA - (pBSU4 cut w/HindIII, T7 Pol)	
U5 snRNA - (pBSU5 cut w/ XbaI, Sp6 Pol)	
U6 snRNA - (pBSU6 cut w/ EcoR1, T3 Pol)	
ddH_2O	remainder up to 50µl
T3/T7/Sp6 Pol	5µl
UTP ^{32}P	4µl

Place at 37˚C for 3 hours.

After transcribing, purify each probe on separate G-25 columns. Take a column and place 800 µl of Sephadex G-25 beads and pack down by spinning at 2300 rpm for 1 minute. After transcription, add the reaction to the column and spin for 2300 rpm for 2 minutes.

Recipes for 20X SSC and 10X Denhardts (50X)

20X SSC – 500 mL
Sodium Chloride - 87.65g
Sodium Citrate - 44.1g
pH to 7 with 10N NaOH (few drops)
up to 500 mL with ddH_2O
Autoclave

50X Denhardts – 100 mL
Ficoll 400 - 1g
Polyvinylpyrrolidone (PVP) - 1g
Bovine Serum Albumin (Fraction V) - 1g

Filter-sterilize with 0.22 µM filter. Place at -20˚C indefinitely. Prior to use, thaw overnight at 4˚C and store at 4˚C until complete. Do not freeze thaw.

Denaturing Gel

Purify RNA samples and resuspend in FEB.

Pour a denaturing 15% acrylamide (29:1), 7M Urea, 1X TBE gel with a 26-lane comb (transcription spacers, thick). Allow the gel to polymerize for 1 hour and then pre-run at 30W for 20 min prior to loading (1X TBE as running buffer).

Heat RNA samples at 95˚C for 2 min prior to loading on the gel.

Load the gel and run at 30W for 1hour and 30 minutes.

Transfer to Nylon

Equilibrate materials for transfer

- Nylon needs to be gently laid on water to prevent bubbles from getting in the pores of the membrane. Once on the water, carefully submerge into the water and allow the nylon to equilibrate for at least 20 minutes.
- After submerging in water, transfer the wet nylon to 1X TBE (or running buffer) for 10 minutes prior to transfer

After the gel is done running, all snRNAs will be between the cyan blue and the combs. Therefore, carefully trim the gel down to accommodate the area that has your RNAs of interest. Once cut, place the gel into 1X TBE (or whatever transfer buffer to equilibrate) until transfer.

- Prepare transfer sandwich
    - Stack on black face of cassette in this order:
        - Sponge, pre-wet in transfer buffer
        - Whatman paper
        - Gel
        - Nylon
        - Whatman paper
        - Sponge, pre-wet in transfer buffer

    Before placing the top sponge, smooth the Whatman, gel, nylon stack with a rolling pin to remove any air bubbles.

Place in the transfer rig (black side of cassette towards the black side/anode of the transfer box)

Run for 16 hours at 4˚C stirring at 12V

Crosslinking

Carefully remove the nylon from the cassette and flip the side that was towards the gel/RNA side up to be exposed onto a piece of glass and allow to dry to a semi-dry state. Mark this side with a pencil.

Place the gel into the UV Stratalinker 1800 and crosslink.

- Turn on and hit the auto-crosslink button followed by start. The machine will then begin to crosslink the RNA to the nylon counting down from 1200kJ to 0.

Pre-hybridization (Pre-hyb)

Pre-hyb Buffer
6X SSC
2X Denhardts
0.1% SDS
100µg/ml of salmon sperm DNA (prior to adding, boil for 5min and then immediately place on ice, then supplement into pre-hyb)

Place the crosslinked nylon into a hybridization tube with the crosslinked/RNA side facing the inside of the tube. Ensure that there is no overlap of the membrane within the tube.

Using a serological pipette, add in 50mL of pre-hyb buffer. It is important to never pour liquids directly on the membrane.

Allow the nylon to rotate in the hybridization oven for 2 hours at 68˚C in pre-hyb buffer.

- If only running a single blot, ensure to balance the oven with an additional tube and liquid volume.

Probing the Blot

Remove the pre-hyb buffer from the hybridization tube and add the probe reaction directly to the buffer.

- Depending how intense you want the bands, you can add half of the reaction and save the rest in the -20˚C. Be warned. Once the probe is added to the pre-hyb, it will start to degrade and therefore should be used promptly. Do not freeze if you intend to re-use. Store at 37˚C.

Once the probe is added, gently invert the probe/pre-hyb to mix and then add back into the hybridization tube with the nylon, RNA facing inward with a serological pipette. Place back into the hybridization oven overnight rotating at 42˚C.

Washing the Blot

Wash 1- 50mL	Washes 2-4- 50 mL
1X SSC	0.2X SSC
0.1% SDS	0.1% SDS

The nylon will be washed for a total of 4 washes (50mL each). Each wash will rotate for 20 min at 42˚C in the hybridization oven.

Gently pour off the pre-hyb with probe and add Wash 1 with a serological pipette.

Remove the 1st wash and proceed with Wash 2 by adding it to the hybridization tube and nylon with a stereological pipette.

Repeat Wash 2, 2 more times for a total of 3 washes with the Wash 2-4 wash buffer recipe.

Visualization

After the final wash, remove the nylon from the hybridization tube and place RNA side up on a piece of Whatman paper. Gently cover with saran wrap and place in a phosphor cassette. After the appropriate amount of time, image on the typhoon. The timing will depend on the intensity of the rad signal which can be gauged with a Geiger counter prior to insertion in the cassette.

Appendix II – Chapter 4: Isolation of spliceosomal complexes via tripartite purification

Early Spliceosome Complex Purification
Adapted from Melissa Jurica and Melissa Moore 2004

4sU Transcription Reaction - 50 µl

ATP (10mM)	2 µl
CTP (10mM)	2 µl
UTP (10mM)	2 µl
GTP (10mM)	1 µl
CAP or GMP (10mM)	4 µl
4sU (10mM)	0.33 µl
DTT (1M)	0.5 µl
10x Txn Buffer	5 µl
T7Pol	5 µl
^{32}P-UTP	1 µl
Linear DNA	40ng/µl final concentration
ddH_2O	up to 50 µl total volume

Take a 1:100 for concentration determination.

Mix at RT and place at 37˚ C for 3 hours.

Gel Purification

Mix the txn reaction with equal volume of FEB. Heat sample for 2min at 95˚ C prior to loading.

Pour a txn gel (5% acrylamide (1:29), 7M Urea, 1X TBE) and pre-run at 45W for 20 min.

Load the gel (50 µl per lane) and run gel at 45W for 45 min.
Expose gel to x-ray and excise the band. From now on, the RNA should be protected from light and should be covered in foil and worked with under dim lights to prevent crosslinking from occurring due to UV exposure.

Splicing/Assembly Reaction

Tube #1

RNA Substrate (Final concentration of 20nM)
Heat at 95˚C for one minute, then place on ice.
Add 27 µl MS2:MBP (37 µM) => get from Melissa
Incubate on ice for 5 minutes

Tube #2 (1.2 mL) – Add in the following order

H_2O	up to 1200mL
KGlu (1M)	60µl
MgAc (100mM)	30µl
ATP (100mM)	30µl
CP (250mM)	30µl
tRNA (10mg/mL)	15µl
RNAsin	15µl
Tube #1	Entire Contents of Tube #1
NE – B2'' (40%)	600µl

Combine all reagents and then aliquot into 100µl reactions in separate Eppendorf tubes.

Place at 30˚C for 10 min to assemble

UV Crosslinking
Adapted from Manny Are's Crosslinking Protocol

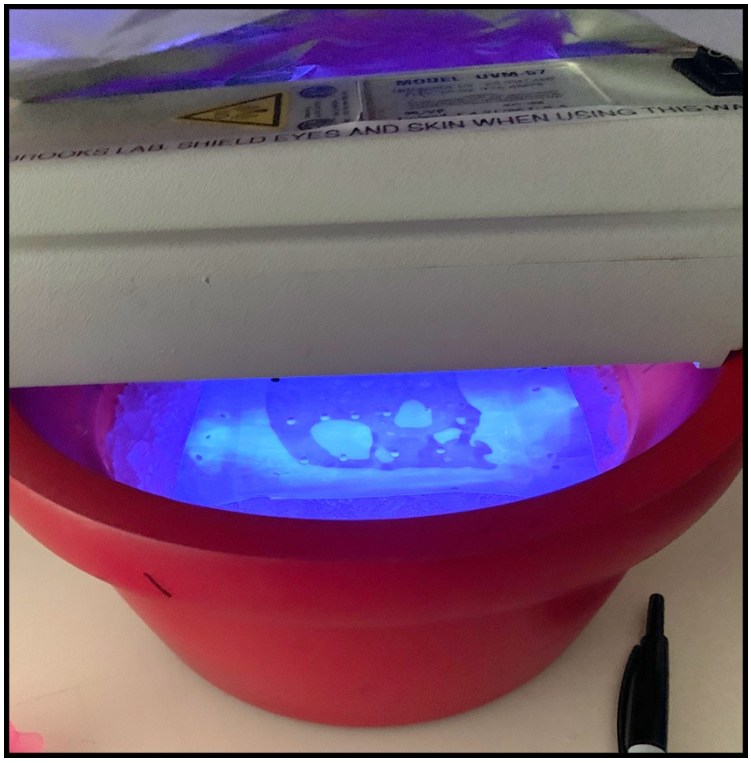
Set up crosslinking apparatus. (See Below)

Room lights should be turned off and the room should be in total darkness

Carefully pipette each splicing reaction onto a piece of parafilm that has been chilled on a piece of glass.

Turn on UV crosslinker to 302nM after placing it 2 cm above the samples and allow to crosslink for 5 min. Covered the entire set up with foil.

After 5 minutes, turn off the crosslinker and combine all samples into one Eppendorf tube. Take a sample for testing (initial crosslinked splicing reaction).



Size Exclusion Chromatography (SEC) – 6mL Sephadex 300

5X SCB2 – 50mL

100mM Tris, pH 7.9
750mM KCl
25mM EDTA (can leave out the EDTA for downstream purposes – Must remove for EM grids)
ddH_2O

1X SCB2-N – 50mL Keep at RT (Make fresh each prep)

5X SCB2 - 10 mL
1M DTT - 50 µL
10% NP-40 - 250 µL
10% Glycerol - 5 mL
ddH_2O - 32.7 mL

1. Equilibrate sizing column with 5 mL 1X SCB2-N
2. Place crosslinked reactions over the sizing column and once fully into the column added 1X SCB2-N until 24 500µL fractions had been collected.
3. Take 50µL samples of each fraction for further analysis.
4. Count the fractions with a Geiger counter and pool the five fractions associated with the first peak.

Amylose Column Purification

1X SCB2-M – 1mL Keep at 4˚C (Make fresh each prep)

5X SCB2 - 200 µL
50% Glycerol - 20 µL
100 mM MgAc - 20 µL
0.5M Maltose - 20 µL
ddH_2O - 740 µL

Pour a 200µL amylose column in a Mobicol and equilibrate with SCB2-N. Attach a 10mL syringe via Leuer lock to the top and place at 4˚C.

Add the combined SEC fractions to the amylose column and allow the complexes to bind.

Wash the column with 10mL of SCB2-N.

Remove the syringe and elute the complexes in 20 µL aliquots of SCB2-M. Utilize the Geiger to determine which fractions have the maximal amount of RNA/complexes.

Immunoprecipitation

*** Ensure you know what downstream process you are doing as that will determine what type of elution you will perform at the end of the IP.

Day 1

Sample Prep
Take amylose eluted sample and remove 5µl for an input sample.

Mix 2µg of anti-V5 antibody with the amylose column elution sample. Cover in foil. Rotate at 4˚C for 13.5 hours.

Bead Prep
Wash 10µL of Protein A magnetic beads with 50µL of 1X PBS cold. Pellet beads by placing on a cold magnet for 2 minutes, then remove the supernatant and take beads off the magnet to resuspend and to wash again. Repeat 3X and then place in 50 µl 1X PBS. Store at 4˚C until the next day.

Wash Buffer
100mM Tris pH 7.5
1mM EGTA
120mM KCl
1% NP40

Day 2- Everything from this point on needs to be on ice

- Remove 1x PBS from beads and resuspend them in the amylose elution incubated with antibody. Cover the tube in foil and rotate at 4˚C for 4 hours.
- After rotating, place the sample on a cold magnet for two minutes. Remove the supernatant (and save it "flow through sample"). Wash beads with 50µL wash buffer (ice cold). Allow the beads to sit in the wash buffer for 1 minute prior to placing on to the magnet for 2 minutes. Remove the wash (Wash 1) and repeat the wash step 2 more times for a total of 3 washes (save all wash buffers in individual tubes) for downstream analyses.
- Note: for downstream analyses other than western blot, will perform 6 washes to remove all unbound proteins.

TEV Elution – Received from C. Partch's lab

Prepping TEV-His Protease
- Dilute 20µl TEV protease into 30µl solution (50mM KCl, 20mM Tris, pH 7.9)
- Dilute this further by taking 20 µl of the initial dilution and diluting 1:2 in 50mM KCl, 20mM Tris, pH 7.9 (Will use 40µl per sample)

Elution

Add 40µl TEV solution directly to the beads and rotate at 4˚C for 1 hour. After 1 hour, place beads on a magnet for 1 minute and remove the supernatant. Save 5µl for analyses (labeled Elution 1, E1) as well as the protein A beads. Resuspend His-beads (30µl slurry, equilibrated and washed with 50mM KCl, 20mM Tris, pH 7.9) with the supernatant and allow the beads to again rock for 1 hour at 4˚C to remove the TEV protease from the sample. After 1 hour, placed sample on a magnet for 1 minute and remove the supernatant (Elution 2, E2).

To the Prot. A beads and His-Beads: perform a soft elution (see below) for downstream analysis.

Soft Elution – 0.1M Glycine, pH 2.5

Add 10µL of 0.1M glycine pH 2.5 after the final wash step. Let the beads sit on ice for 5 minutes. Place on the magnet for 2 minutes and remove the elution to a new tube. Perform a second elution and combine the two elutions into one tube (totals 20 µl). Add an equal volume of 1M Tris, pH 8 to quench the glycine (20µl).

Hard Elution – 5X SDS Sample Buffer

After the soft elution, add 16µL of ddH20 directly to the beads + 4µL of 5x SDS-sample buffer. Boil for 5 minutes at 95˚C. Load the sample directly onto PAGE gel for analysis. Should see IgG bands in the hard elution but not the soft elution.

Downstream Analyses:

EM Grids – Perform TEV elution and purification without EDTA

Mass Spec – Perform Soft Elution

Western Blot – Perform Soft or Hard Elution. Decision based on whether protein runs near IgG.