Investigation of the effects of the splicing inhibitor Pladienolide B and cancer-related mutations in the SF3b1 protein on pre-mRNA splicing *in vitro*

Dissertation

for the award of the degree

“Doctor rerum naturalium”

of the Georg-August-Universität Göttingen

within the doctoral program Molecular Biology

of the International Max Planck Research School for Molecular Biology

and the Georg-August University School of Science (GAUSS)

submitted by

Sebastian Ludwig

from Leverkusen

Göttingen 2019

Members of the Examination Board:

Prof. Dr. Reinhard Lührmann (1st Referee)

Department of Cellular Biochemistry

Max Planck Institute for Biophysical Chemistry, Göttingen

Prof. Dr. Ralf Ficner (2nd Referee)

Department of Molecular Structural Biology

Georg-August-Universität Göttingen

Prof. Dr. Patrick Cramer

Department of Molecular Biology

Max Planck Institute for Biophysical Chemistry, Göttingen

Prof. Dr. Henning Urlaub

Research Group Bioanalytical Mass Spectrometry

Max Planck Institute for Biophysical Chemistry, Göttingen

Prof. Dr. Ralph Kehlenbach

Department of Molecular Biology

Georg-August-Universität Göttingen

Prof. Dr. Detlef Dönecke

Department of Molecular Biology

Georg-August-Universität Göttingen

Affidavit

Herewith I declare, that I prepared the Doctoral Thesis “Investigation of the effects of the splicing inhibitor Pladienolide B and cancer-related mutations in the SF3b1 protein on pre-mRNA splicing *in vitro”* on my own and with no other sources and aids than quoted.

Göttingen,

For my family.

Index

Abstract

Pre-mRNA splicing is catalyzed by the spliceosome, a highly dynamic, macromolecular RNP complex formed out of the U1, U2, U4, U5 and U6 snRNPs and multiple snRNP proteins. Spliceosomes are assembled stepwise on an intron and they can initially form across an exon, generating a cross exon A-like complex if the length of the intron separating the 5' and 3' splice sites is very long, or alternatively across an intron, yielding an A complex if this distance is short. During the early stages of spliceosome assembly, the 5’ splice site is first bound by the U1 snRNP and subsequently, during A complex formation, the branch point sequence (BPS) of the intron is recognized and bound by the U2 snRNP. In the subsequent step the U4/U6.U5 tri-snRNP docks generating the pre-B complex and after its stable integration, the pre-catalytic B complex is formed. In the following steps, the spliceosome undergoes compositional and structural rearrangements leading to its catalytic activation and then catalysis of the pre-mRNA splicing reaction. SF3b is a heptameric subunit of the U2 snRNP that is required for the stable base pairing of the U2 snRNA with the BPS during A complex assembly, and it plays a key role in the recognition and selection of the branch site adenosine. The U2/BPS helix is clamped, and thereby stabilized, within the C-terminal HEAT domain of SF3b1, the largest protein of the SF3b complex. SF3b1, together with the SF3b subunit PHF5A, form the binding pocket for the branch site adenosine, the nucleophile for the first catalytic step of splicing.

SF3b1 is a target of many natural pre-mRNA splicing inhibitors, including Pladienolide B (PlaB), that also exhibit anti-tumor activity and lead to changes in alternative splicing patterns. PlaB stalls spliceosome assembly prior to the formation of a pre-catalytic B complex but currently the exact mechanism responsible for this stalling and the effect of PlaB on the structure and composition of the spliceosome are unknown. To investigate this, I analyzed spliceosome assembly in HeLa nuclear extract in the presence of PlaB. I could show that PlaB does not affect the kinetics of cross-intron A complex formation, but prevents the subsequent docking of the tri-snRNP and thus blocks the formation of a pre-B complex. I could also show that cross-exon A-like complexes form in the presence of PlaB, but that the subsequent docking of the tri-snRNP is also inhibited. My studies further confirmed that the A

complexes formed in the presence of PlaB are less stable, dissociating more readily in the presence of heparin. To investigate the basis for these changes, I affinity purified cross-intron A complexes stalled by PlaB and investigated their proteome by mass spectrometry. Compared to wild-type A complexes, the PlaB stalled complexes contained increased amounts of the helicase DDX42, which was proposed to play a role in the biogenesis of the 17S U2 snRNP, and reduced amounts of SPF30. As SPF30 was previously shown to be required for tri-snRNP incorporation into the spliceosome, its absence could contribute to the lack of tri-snRNP recruitment to the PlaB stalled A complexes. Initial attempts to obtain a high resolution structure of the PlaB stalled A complex and, for comparison, the wild-type human A complex, by performing cryo-EM were not successful. Thus, future studies will be needed to determine what effect PlaB has on the structure of the spliceosome.

Point mutations in the SF3b1 HEAT domain, including the exchange of the lysine at amino acid 700 with glutamic acid (K700E), are linked to various cancers and lead to the utilization of alternative branch sites and 3' splice sites in vivo. The exact mechanism responsible for these changes in alternative splicing is currently unknown. To investigate the consequences of cancer-linked SF3b1 mutations on splicing, I first analyzed their effects on the 17S U2 snRNP, using K562 cells expressing FLAG-tagged, wild-type (wt) SF3b1 or the SF3b1^{K700E} mutant. Interestingly, affinity purification of human 17S U2 snRNPs containing SF3b1^{K700E}, followed by mass spectrometry and immunoblotting, revealed the nearly complete absence of the U2-associated DEAD-box helicase hPrp5 and a significant reduction in the abundance of the hTAT-SF1 protein, compared to 17S U2 snRNPs containing wt SF3b1. Protein-protein crosslinking of 17S U2 snRNPs revealed crosslinks between the SF3b1 HEAT domain and hPrp5, consistent with the idea that conformational changes due to the K700E mutation may directly affect Prp5 binding to the U2 snRNP. As hPrp5 plays a key role in the stable binding of U2 during A complex formation, its absence may directly lead to alterations in BPS selection by U2 containing SF3b1^{K700E}. Initial experiments also showed that the mutant, FLAG-tagged 17S U2 snRNPs are incorporated into spliceosomal A complexes, indicating that they are still functional, and that it may be possible to isolate A complexes containing SF3b1^{K700E} for future biochemical and functional studies.

Finally, I set out to establish in vitro splicing conditions that recapitulate the use of alternative BPS and 3' splice sites in cells expressing SF3b1^{K700E} that is observed in vivo, in order to study in detail the effect of this mutation on splice site selection. I first tested the splicing of ZDHHC16 pre-mRNA containing a cryptic branch site and 3' splice site in nuclear extract from wt and SF3b1^{K700E} K562 cells; however, spliceosome assembly was stalled at the A complex stage. As an alternative to cross-intron splicing, I designed exon RNA substrates with regions of the introns of ZDHHC16, TMEM14C or ENOSF1 that contain cryptic branch sites and 3' splice sites that are used in the presence of SF3b1^{K700E} in vivo. All tested RNAs supported cross-exon A-like formation in both HeLa and K562 nuclear extracts, albeit very inefficiently in K562 nuclear extract. Unexpectedly, exon complex formation on the cryptic branch sites of these substrates, upon mutation of the canonical BPS, was not dependent on the presence of SF3b1^{K700E}. Nonetheless, these in vitro studies lay the groundwork for future investigation of the composition and structure of splicing complexes containing the SF3b1^{K700E} mutation. Taken together, my studies provide new insights into the mechanism whereby the SF3b1 modulator PlaB potentially leads to a block in spliceosome assembly, and also into the effects of SF3b1 cancer-related mutations on the composition of the 17S U2 snRNP. They also pave the way for future in vitro studies to improve our understanding of how alternative branch site usage arises as a consequence of single amino acid substitutions in SF3b1, and thus potentially aid our understanding of how SF3b1 mutations lead to cancer.

1 Introduction

The central dogma of molecular biology defines the flow of genetic information from DNA via messenger RNA (mRNA) to protein. First, the DNA is transcribed into RNA during the process called transcription; in the subsequent step, translation, the RNA is translated into a protein with the corresponding amino acid sequence. While this dogma holds true for prokaryotes, the genomic organization of eukaryotic organisms makes additional steps necessary to translate genetic information into proteins. Eukaryotic genes are discontinuous; the coding information (exons) is separated by non-coding elements, the introns. Transcription in eukaryotes produces a premature messenger RNA transcript (pre-mRNA) that must undergo multiple processing steps prior to translation. The most complicated maturation step is the removal of introns and joining of exons, which is known as pre-mRNA splicing. Pre-mRNA splicing is a complex process governed by a macromolecular ribonucleoprotein complex called the spliceosome. The spliceosome identifies the exon-intron boundaries and joins the two exons flanking an intron while simultaneously removing the intron.

1.1 Pre-mRNA processing

Eukaryotic RNA-polymerase II transcripts undergo extensive co- and post-transcriptional modifications in addition to the removal of introns, before being exported into the cytoplasm for translation (Kelly and Corbett, 2009). The first modification is the addition of a N^7-methylguanosine (m^7G-cap) to the 5’ end of the growing transcript by formation of a 5’-5’ triphosphate bridge and 2’O-methylation of the first two pre-mRNA nucleotides. The capping of pre-mRNA is required for the nuclear export mechanism and to protect the RNA from degradation by 5’-3’ exonucleases. After transcription, the 3’ end of the pre-mRNA is cleaved, and a poly(A)-tail is added; it increases the stability of the pre-mRNA and later helps initiate the translation of the pre-mRNA into proteins (Jackson and Standart, 1990). In addition to these modifications at the 5’ and 3’ end of the RNA molecule, the intronic sequences of the transcribed gene are removed by pre-mRNA splicing. The process of pre-mRNA splicing takes place during and after transcription of the RNA molecule and has to be completed prior to nuclear export. Splicing of pre-mRNA carried out by the spliceosome will be described in detail in the following chapters.

1.2 Identification of splice sites

The most important step to ensure a precise and reproducible outcome of intron removal is the correct identification of splice sites within the introns. The analysis of the pre-mRNA transcriptome of various organisms has led to the identification of short conserved *cis*-acting motives within the introns that are recognized by the spliceosome (Burge et al., 1999). These elements can be found near the exon-intron junctions and they are required for the identification of splice sites by the spliceosome. They are known as the 5' splice site (5'SS), the branch point sequence (BPS) and the 3' splice site (3'SS) (Figure 2.1). Unlike yeast, higher eukaryotes also contain a polypyrimidine tract located between the BPS and the 3'SS.

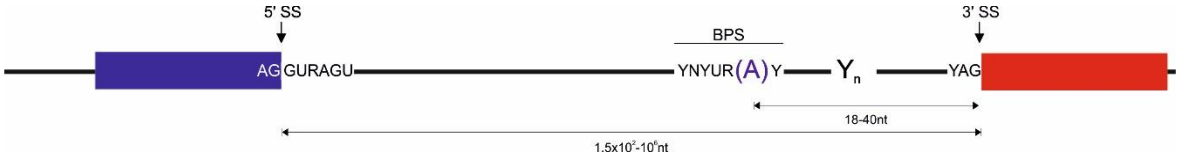


Figure 1.1- Scheme of a human pre-mRNA. The 5' exon is displayed in blue and the 3' exon is in red. Indicated are the important recognition sites for spliceosome assembly with their respective consensus sequences: the 5' splice site (5'SS), the 3' splice site (3'SS), the branch point sequence (BPS) with the branch point (marked in blue) and the poly-pyrimidine tract between BPS and 3'SS.

The 5'SS defines the border between the 5' exon and the downstream intron and has a consensus sequence of AG|GURAGU ("|" marks the 3' end of the exon) in *Homo sapiens*. The 3'SS defining the border between the intron and the downstream exon has a sequence of CAG| or UAG| ("|" marks the 5' end of the exon). The consensus sequence for the branch point region in human pre-mRNAs is defined as YNYURAC (underlined adenosine nucleotide acts as a branch point). The branch point is commonly found between 18 and 40 nucleotides upstream of the designated 3'SS. As mentioned earlier, in humans, a polypyrimidine rich region known as the pY-tract is found between the branch point and the 3'SS; it is used as an additional interaction region for spliceosomal factors and aids in spliceosome assembly.

The *cis*-acting elements required for splicing of pre-mRNAs show different degrees of sequence conservation (Table 2.1). While in *Saccharomyces cerevisiae* the sequences of the individual elements are precisely defined, these sequences in higher eukaryotes are much more degenerate. In addition to the aforementioned

classical introns, which are spliced by the major spliceosome (U1/U2 type), higher eukaryotes also contain a small subset of introns (~1%) that possess different, more highly conserved consensus sequences (Table 2.1). These introns are recognized and spliced-out by a second, so-called minor spliceosome (U11/U12-type) (Hall and Padgett, 1994, 1996; Patel and Steitz, 2003).

Table 1.1 – **Splice site consensus sequences for yeast and higher eukaryotes.**

	Higher eukaryotes		**Yeast**
	U2-type	U12-type	U2-type
5’ splice site	AG|GURAGU	|AUAUCCUUU |GUAUCCUUU	AG|GUAUGU
Branch point sequence	YNYUR**A**C	UCCUUA**A**C	UACUA**A**C
3’ splice site	YAG|G	YAC| YAG|	CAG|G

1.3 Chemical mechanism of pre-mRNA splicing

All pre-mRNA introns are removed via the same two-step transesterification reactions (Figure 2.2) (Moore and Sharp, 1993). The process of pre-mRNA splicing itself does not require any external energy input. However, the rearrangements of the spliceosomal components require energy in the form of ATP, making the overall process ATP-dependent. The first step of splicing is the nucleophilic attack of the branch point adenosine (BPA) 2’ hydroxyl-group on the phosphate of the terminal exonic nucleotide. This creates a free 5’ exon and an intron lariat. The intron lariat is formed due to the creation of 2’-5’ phosphodiester bond between the first intronic nucleotide (a guanine) and the BPA. During the second step, the now free 3’ OH group of the 5’ exon performs a nucleophilic attack on the phosphate group of the last intronic nucleotide. This creates a new phosphodiester bond between the last nucleotide of the 5’ exon and the first nucleotide of the 3’ exon while simultaneously releasing the intron lariat. The intron lariat is subsequently degraded while the ligated exons are released.

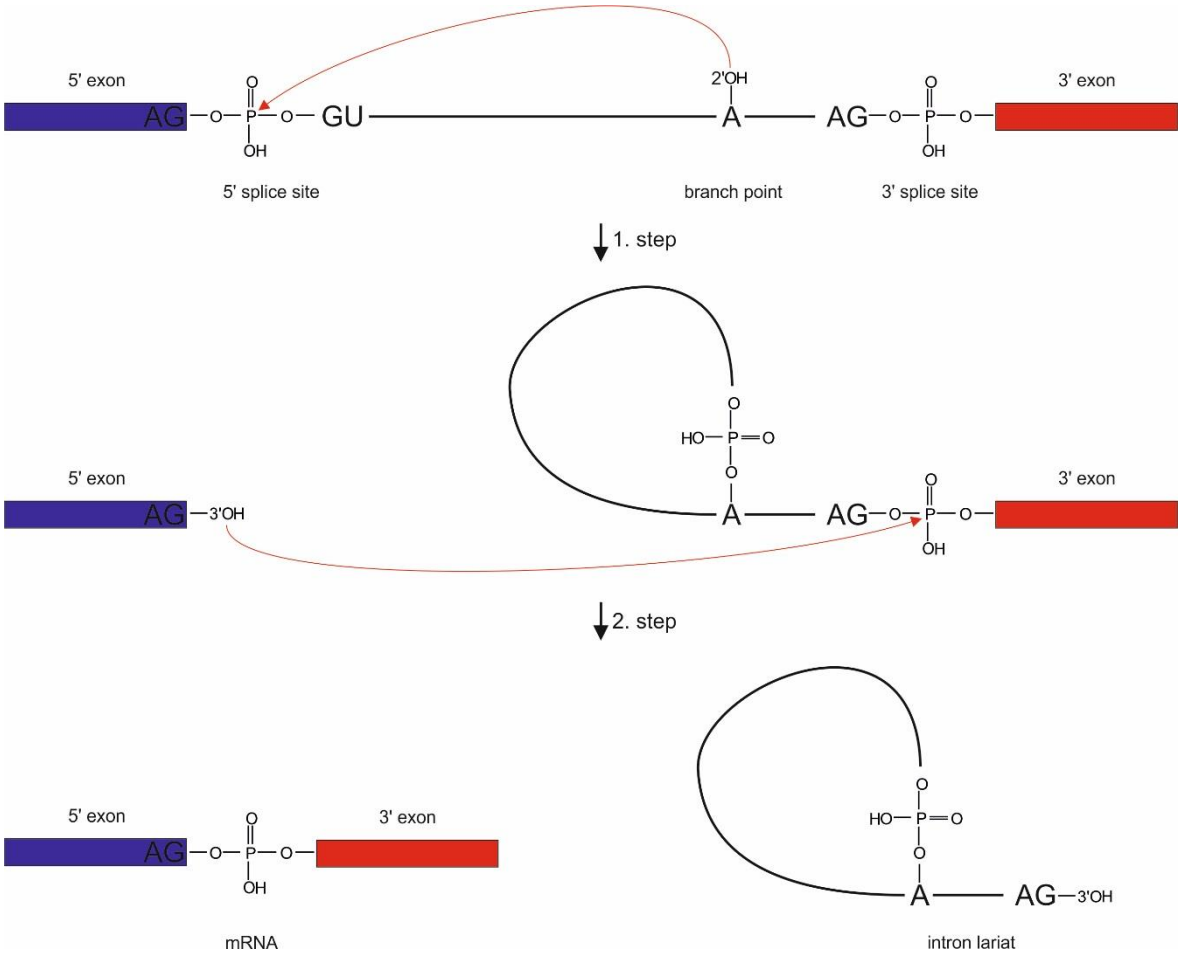


Figure 1.2 – **Schematic representation of the two-step splicing reaction.** The chemical process of pre-mRNA splicing is divided into two transesterification reactions in which the 5' exon is first disconnected from the intron and then ligated with the 3' exon while the intron lariat is removed.

1.4 Stepwise assembly of the human spliceosome

Spliceosome formation activation and catalytic action is a multistep process involving the assembly and disassembly of various complexes as well as precise RNP remodeling events. During these assembly events, multiple distinct spliceosomal complexes can be identified based on biochemical and structural data. The E, A, pre-B, B, B^{act}, B*, C, C* and P spliceosomal complexes are formed (in this order) during the splicing cycle (Will and Luhrmann, 2011). During the activation stage, the catalytic center of the spliceosome is formed and subsequently both transesterification reactions are carried out, as described above. After splicing catalysis, the individual components are recycled to be used in a new assembly cycle. The transitions between the various stages are mediated by energy-dependent DExH/D helicases, described in detail in chapter 2.9.

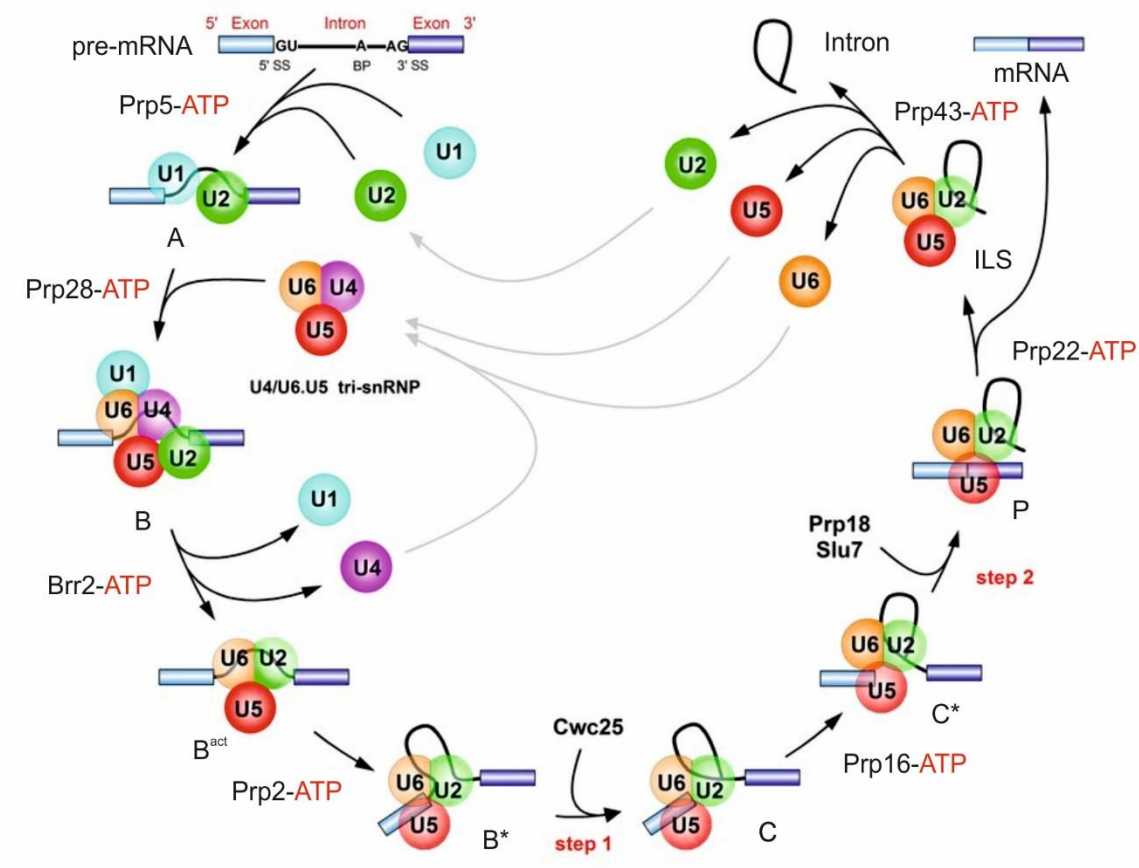


Figure 1.3 – **Stepwise assembly, catalytic activation and actalytic activity pathway of the spliceosome during splicing.** The splicing of pre-mRNA begins with the association of U1 and U2 with the pre-mRNA. After formation of the A complex the tri-snRNP is bound and the B complex is formed. Then the U1 and U4 snRNAs and associated proteins are removed to form the activated B complex (B^{act}). Further structural rearrangements form the catalytically activated B* complex that can then complete the first step of splicing giving rise to the C complex. The C complex is then rearranged for the second step of splicing (C*). Once the second step is completed the spliceosome is disassembled and the individual components are recycled.

The assembly of the spliceosome begins with the recognition of the 5'SS and the BPS/3'SS region. In the spliceosomal E complex the 5'SS is recognized by the U1 snRNP (Figure 2.3). During this recognition process, the U1 snRNA base pairs with the intron nucleotides immediately downstream of the 5'SS. The U1/5'SS interaction is stabilized by the U1-70K protein and U1-C proteins (Heinrichs et al., 1990; Kohtz et al., 1994). In parallel, the BPS region is recognized by the branch point binding protein (BBP), whereas U2AF65 associates with the poly-pyrimidine tract and U2AF35 with the 3'SS. At this stage, the U2 snRNP is also loosely associated. During A complex formation the U2 snRNP displaces the BBP and now stably binds the branch point region. U2 snRNA base pairs with the pre-mRNA forming the U2/BPS helix. Additionally, multiple SF3a and SF3b proteins of the U2 snRNP make contacts with the pre-mRNA and stabilize the U2 binding (Gozani et al., 1996). For instance, SF3b1 binds the pre-mRNA up- and downstream of the branch point

adenosine (Gozani et al., 1998). The branch point adenosine itself is contacted by the SF3b6 (SF3b14a) protein (Will et al., 2001). Besides interacting with the pre-mRNA, SF3b1 also contacts the U2AF65 protein, which is bound to the polypyrimidine tract (Gozani et al., 1998). At this stage, multiple proteins interact with their binding sites on the pre-mRNA and/or other spliceosomal proteins, forming the aforementioned network of protein-RNA and protein-protein interactions necessary for the high fidelity of the splice site selection.

The U5.U4/U6 tri-snRNP first binds loosely to the A complex and a pre-B complex results (Boesler et al., 2016). Upon stable association of the tri-snRNP with the spliceosome, the B complex is formed (Figure 2.3). During the association of the tri-snRNP, the helicase Prp28 destabilizes the interactions between U1 and the 5'SS, thereby enabling the interaction of the U6 ACAGA box with the 5'SS (Staley and Guthrie, 1999). During this process, the kinase Prp4 phosphorylates the tri-snRNP-associated factors Prp31 and Prp6, and the SF3b1 protein is hyper-phosphorylated by an unknown kinase. These phosphorylations are required for the splicing progression (Girard et al., 2012; Schneider et al., 2010a). During the next remodeling event of the spliceosome, catalyzed by Brr2 helicase, the U4 snRNP is removed from the B complex. In parallel, multiple B^{act} associated proteins are recruited and major structural remodeling events in the RNA-RNA network take place forming the activated B complex (B^{act}). Then Prp2 further remodels the spliceosome, creating the B* complex containing the catalytic center required for the first transesterification reaction. After completion of the first step of splicing, the C complex is formed, which contains the products of the first step of splicing, the free 5' exon and the intron lariat-3'exon. With the help of Prp16, the C complex can be rearranged into the C* complex that catalyzes the second step of splicing. During the second step of splicing, the 5' and 3' exons are ligated, and the intron lariat is excised. The second step of splicing results in the formation of the post-catalytic spliceosome, which is then disassembled by Prp22, resulting in the release of the mRNA and of the intron lariat bound to the remaining U snRNPs. The U snRNPs are then recycled with the help of Prp43, whereas the intron lariat is degraded.

1.5 Exon definition

Recent studies have shown that the average size of an intron is in the range of 5400 nucleotides compared to the much smaller exons with an average size of 170 nucleotides (Sakharkar et al., 2004). The initial assembly of spliceosomes across individual introns (Figure 2.4), a process known as intron definition, is limited to introns shorter than 300 nucleotides (Sterner et al., 1996). Pre-mRNA introns that exceed this length utilize an alternative pathway called exon definition, in which the first spliceosomal complex forms across a comparably short exon (Berget, 1995). The assembly of a cross-exon complex starts with the base pairing of the U1 snRNA to the 5'SS downstream of an exon, while the U2 snRNP interacts with the BPS and U2AF with the 3'SS located upstream of the exon (Figure 2.4). This process defines the 5' and 3' borders of an exon (hence the name exon definition) instead of an intron. Enhancer and silencing elements within the exon can also influence this recruitment process by allowing SR proteins and hnRNPs to interact with their respective ESE or ESS sequences. This regulation of spliceosome assembly at the stage of exon complex formation most likely has an impact on alternative splicing.

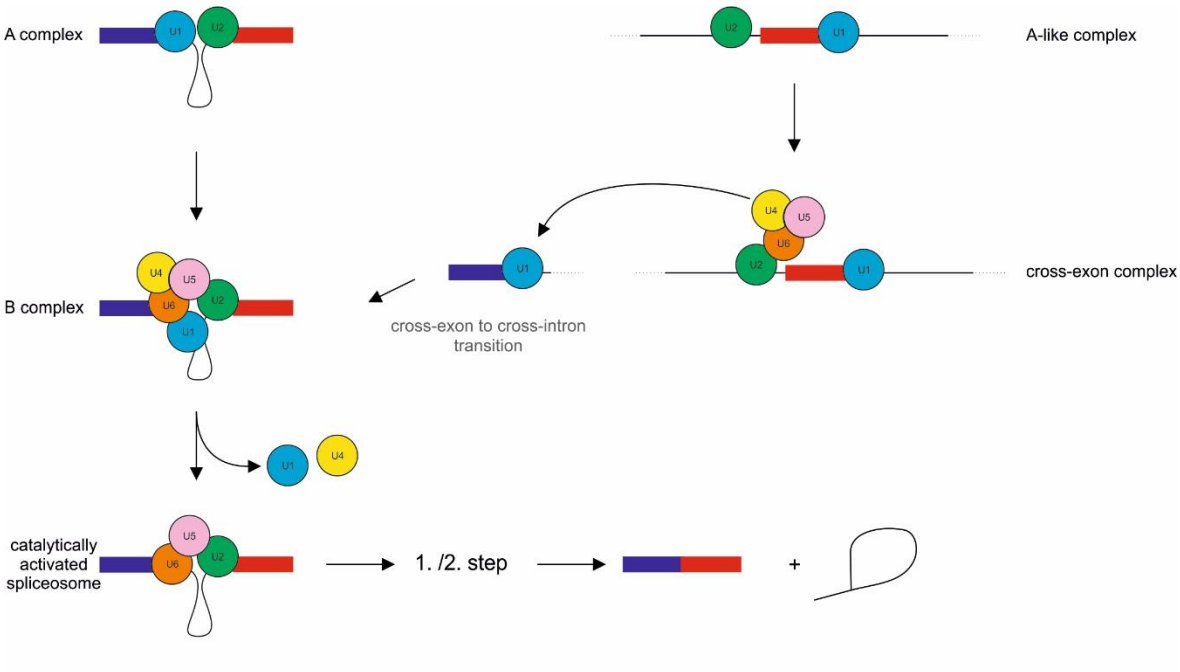


Figure 1.4 – **Splicing of pre-mRNAs after cross-intron or cross-exon assembly of the spliceosome.** Spliceosome assembly can take place via formation of a cross-intron A complex or by assembling a cross-exon A-like complex. Both complexes then recruit the tri-snRNP and form the pre-B or cross-exon complex respectively. Once formed the cross-exon complex can interact with the upstream 5'SS and transition into the normal B complex. After formation of the B complex splicing progresses normally. Exons are depicted in blue,

for the upstream exons and red for the downstream exon. The introns are indicated as black line and various U snRNPs are shown as colored circles.

In order for a cross-exon complex to take part in pre-mRNA splicing, a transition into a cross-intron complex must take place because the catalytic steps of splicing can only occur in this complex. Previous studies indicate that the cross-exon complex can convert into the cross-intron complex by the recruitment of an upstream 5'SS, leading to the pairing of cross-intron splice sites (Schneider et al., 2010b). This means that the cross-intron and the cross-exon assembly of the spliceosome converge into one common pathway during the B complex stage (Figure 2.4).

1.6 The spliceosome and its building blocks

The precise identification of exon-intron borders and the correct ligation of individual exons is crucial for the cell viability to prevent the formation of aberrant gene products. Even small errors in splicing can have significant consequences for the cell by introducing frameshift mutations into the mRNA or by including/ excluding sequences in the spliced mRNA. The cis-acting sequences defining the exact position of splicing sites on pre-mRNAs are relatively short and only weakly conserved in higher eukaryotes. Therefore, these sequences are not sufficient to specify splice sites with the required fidelity on their own nor do they allow for the juxtapositioning of the reactive nucleotides for catalysis. Instead, additional *trans*-acting components are needed to increase the overall fidelity of the splicing process by recognition of different cis-acting elements and to achieve the three dimensional RNA architecture required for splicing (Will and Luhrmann, 2011). These *trans*-acting factors can be grouped into the U snRNPs and non-U snRNP proteins.

1.6.1 Uridine-rich small nuclear ribonucleoproteins

The spliceosome is a multi-megadalton RNA/protein (RNP) assembly that undergoes major structural changes during the splicing cycle, which are accompanied by significant changes in its protein and RNA composition. The major spliceosome consists of five U snRNPs (U1, U2, U4, U5 and U6) that interact with the pre-mRNA during the splicing cycle. Each snRNP contains a uridine-rich snRNA that acts as a binding platform for multiple different interacting proteins. The minor spliceosome, responsible for 1% of splicing events in humans, is also comprised of

five U snRNPs; however, except for U5, it contains a different set of U snRNPs (U11, U12, $U4_{atac}$, U5 and $U6_{atac}$) with analogous functions to their counterparts found in the major spliceosome. The snRNA molecules of the snRNPs adopt defined three dimensional structures (Figure 2.5), which act as interaction sites for different RNA-binding proteins and also base pair with nucleotides of the 5' splice site and BPS (see below). The structures of the individual snRNAs are highly conserved but they undergo significant changes during the splicing process. The most notable secondary structure formed by the snRNAs is the U4/U6 snRNA duplex. U4 and U6 snRNA share two regions of complementarity, the stem I and stem II, and thereby form a di-snRNP via base pairing (Brow and Guthrie, 1988).

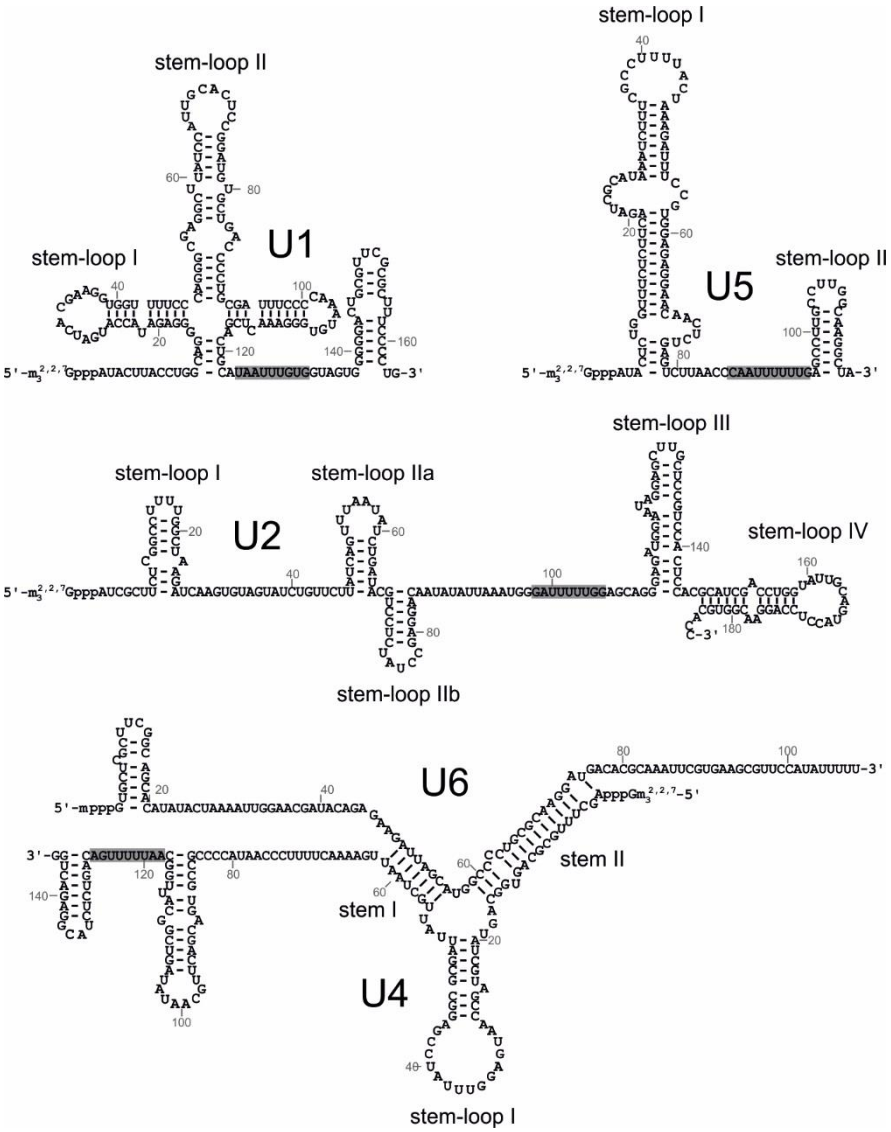


Figure 1.5- **Sequence and structure of the human spliceosomal snRNAs.** Sm protein binding sites are indicated in grey.

The U1, U2, U4 and U5 snRNAs all contain a so called Sm-site (Figure 2.5). These RNAs are synthesized by the RNA polymerase II and initially contain an m^7G-cap. During maturation, they are exported to the cytoplasm. Here the Sm-proteins B, D1, D2, D3, E, F and G bind the uridine-rich Sm-site, forming a ring. Once the Sm-ring is formed, the m^7G undergoes hypermethylation to the 2, 2, 7 tri-methyl-guanosine-cap and the 3' end of the RNA molecule is trimmed (Mattaj, 1986). The formation of the Sm-ring and the hypermethylation of the 5'-cap induce nuclear re-import of the snRNAs. In the nucleus, the remaining snRNA-associated proteins interact with their respective binding sites on the RNA and the mature snRNP is formed. The U6 snRNA in contrast is transcribed by RNA polymerase III and does not contain a Sm-site. Instead, the LSm proteins 2-8 assemble into a ring-like structure by binding to a uridine stretch at the 3' end of the RNA (Achsel et al., 1999). The 5' end of the U6 snRNA contains a γ-monomethyl cap, as a result of the transcription by RNA polymerase III. The maturation of U6 snRNA is completed entirely within the nucleus of the cell (Mroczek and Dziembowski, 2013).

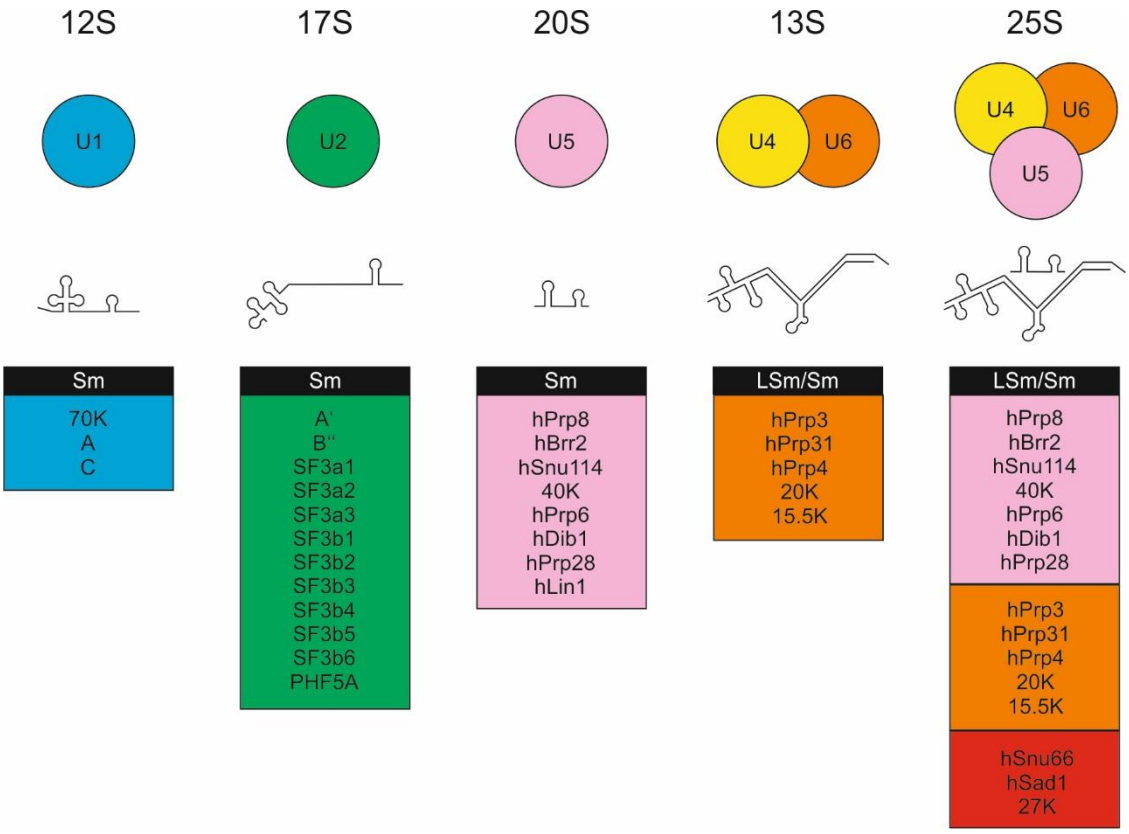


Figure 1.6 – **Protein composition of the human U snRNPs.** Spliceosomal snRNPs are represented as colored circles with their sedimentation value (Svedberg units) indicated above. The snRNA structure and the individual associated protein composition is depicted underneath, with colored boxes indicating the origin of the proteins for multi-particle assemblies.

As mentioned above, each of the U snRNAs adopts a three-dimensional fold and interacts with a specific set of proteins in addition to the Sm-/LSm-ring proteins (Figure 2.6). For instance, the U1 snRNA interacts with three additional proteins, U1-70K, U1-A and U1-C. The 17S U2 snRNP is comprised of U2A' and U2B'' and the two heterologous protein complexes SF3a and SF3b. SF3a is comprised of 3 proteins (SF3a1, SF3a2 and SF3a3) whereas SF3b contains seven proteins (Will and Lührmann, 2002). Proteins in both complexes help the U2 snRNP to recognize and select the correct BPS. Additional proteins associated with the 17S U2 particle are hPrp43, SPF45, SPF30, SR140, CHERP, PUF60, hPrp5, hTAT-SF1 and U2AF1/U2AF2 (Will and Lührmann, 2002). The U5 snRNP contains eight proteins (hPrp8, hBrr2, hSnu114, hPrp6, hPrp28, hLin1, 40K and hDib1) (Bach et al., 1989) and it is responsible for tethering of the 5'SS and the 3'SS after the first step of splicing, bringing the two exons into proximity for the second step of splicing (Newman, 1997). Prp8 is the largest snRNP protein of the spliceosome and acts as a scaffold for the catalytic RNA network in the spliceosome (Fica et al., 2017; Galej et al., 2016; Nguyen et al., 2015; Nguyen et al., 2016). hPrp28 and Brr2 are two DExD/H-helicase helicases that are required for the formation of the pre-catalytic spliceosome and the subsequent rearrangements leading to an activated spliceosome (see chapter 2.9) (Boesler et al., 2016; Mathew et al., 2008; Möhlmann et al., 2014; Price et al., 2014). The activity of hBrr2 is regulated by the U5-associated proteins hPrp8 and hSnu114 (Mozaffari-Jovin et al., 2012; Mozaffari-Jovin et al., 2014; Small et al., 2006).

The U4 and U6 snRNP do not exist as individual snRNP particles but are found as parts of the U4/U6 di- or U4/U6.U5tri-snRNP. The U4/U6 di-snRNP contains five different proteins, hPrp3, hPrp31, hPrp4, CypH and 15.5K (Behrens and Luhrmann, 1991; Lauber et al., 1997). The tri-snRNP is formed after association of the U5 snRNP with the di-snRNP, and it contains all proteins previously associated with the di-snRNP, excluding hLin1. The tri-snRNP also recruits three additional proteins (hSnu66, hSad1, 27K) that were not part of either the di-snRNP or U5 snRNP (Behrens and Luhrmann, 1991). The U5 protein hPrp6 and the di-snRNP associated protein hPrp31 play an important role in the formation of the tri-snRNP, as depletions of either protein prevent its formation *in vivo* and *in vitro* (Makarova et al., 2004; Schaffert et al., 2004). The interactions between the di-snRNP and the U5 snRNP

in the mature tri-snRNP are exclusively based on protein-protein contacts, no interactions between the snRNA molecules are known.

1.7 The RNA network of the spliceosome

The pre-mRNA splicing is catalyzed by an active center of the spliceosome entirely comprised of RNA molecules. This active center remains largely unchanged for the first and second catalytic step of splicing. In order to form the active center, extensive re-organizations in the pre-mRNA-snRNA and snRNA-snRNA network of the spliceosome are required (Wahl et al., 2009; Will and Luhrmann, 2011). The first RNA-RNA interaction formed during the splicing reaction is the base pairing of the U1 snRNA with the 5' splice site. During U2 binding, the so-called U2/branch point helix is formed, between the 5' end of the U2 snRNA and the BPS of the intron, and the branch point adenosine is bulged out. The next step is the recruitment and integration of the tri-snRNP to form a pre-B complex. During integration of the tri-snRNP, the 3' end of U6 snRNA forms a short helical structure with the 5' end of the U2 snRNA (U2/U6 helix II) (Nilsen, 1994). The U4/U6 RNA interactions remain largely unchanged at this stage; they still interact via the stem I and stem II regions of complementarity (Brow and Guthrie, 1988).

The spliceosomal RNAs and group II self-splicing introns contain multiple shared structural motives and act via the same chemical reaction mechanism. Comparing the structure of the RNA in the catalytic center of the spliceosome with the group II self-splicing RNAs reveals that both form a similar catalytic center (Toor et al., 2008). However, the catalytic RNA structure of the spliceosome only forms in the presence of proteins that stabilize the RNA-RNA interactions required for catalysis. To form the active center of the spliceosome required for the first step of the splicing reaction to occur, major remodeling events within the RNA structure of the spliceosome must take place. The U1 snRNA/pre-mRNA interaction is disrupted and the U6 snRNA forms base pairs near the 5'SS via its ACAGAGA-box (Staley and Guthrie, 1999). After U6 base pairs with the 5' SS, the U4/U6 duplex is disrupted by Brr2 and the U4 snRNP is displaced from the spliceosome. Once U4 is displaced, U2 and U6 snRNA interact to form new base-paired helices (Ia and Ib). The interactions between the U2 snRNA and the U6 snRNA are very dynamic and lead to a three-

dimensional RNA structure crucial for catalysis (Fica et al., 2013). The U6 snRNA also forms a new internal stem-loop (ISL) upon dissociation of the U4/U6 snRNA dimer. The ISL is required to coordinate the two magnesium ions needed for catalysis (Yean et al., 2000). This formation of essential components of the catalytic center after removal of the U4 snRNA suggests a regulatory function for the U4 snRNA; the U4/U6 RNA interactions could prevent premature formation of the active center. The U5 snRNA interacts with both the 5' and the 3' exons during catalysis and these interactions keep the two exons in close proximity for exon ligation during the second step of splicing (Sontheimer and Steitz, 1993).

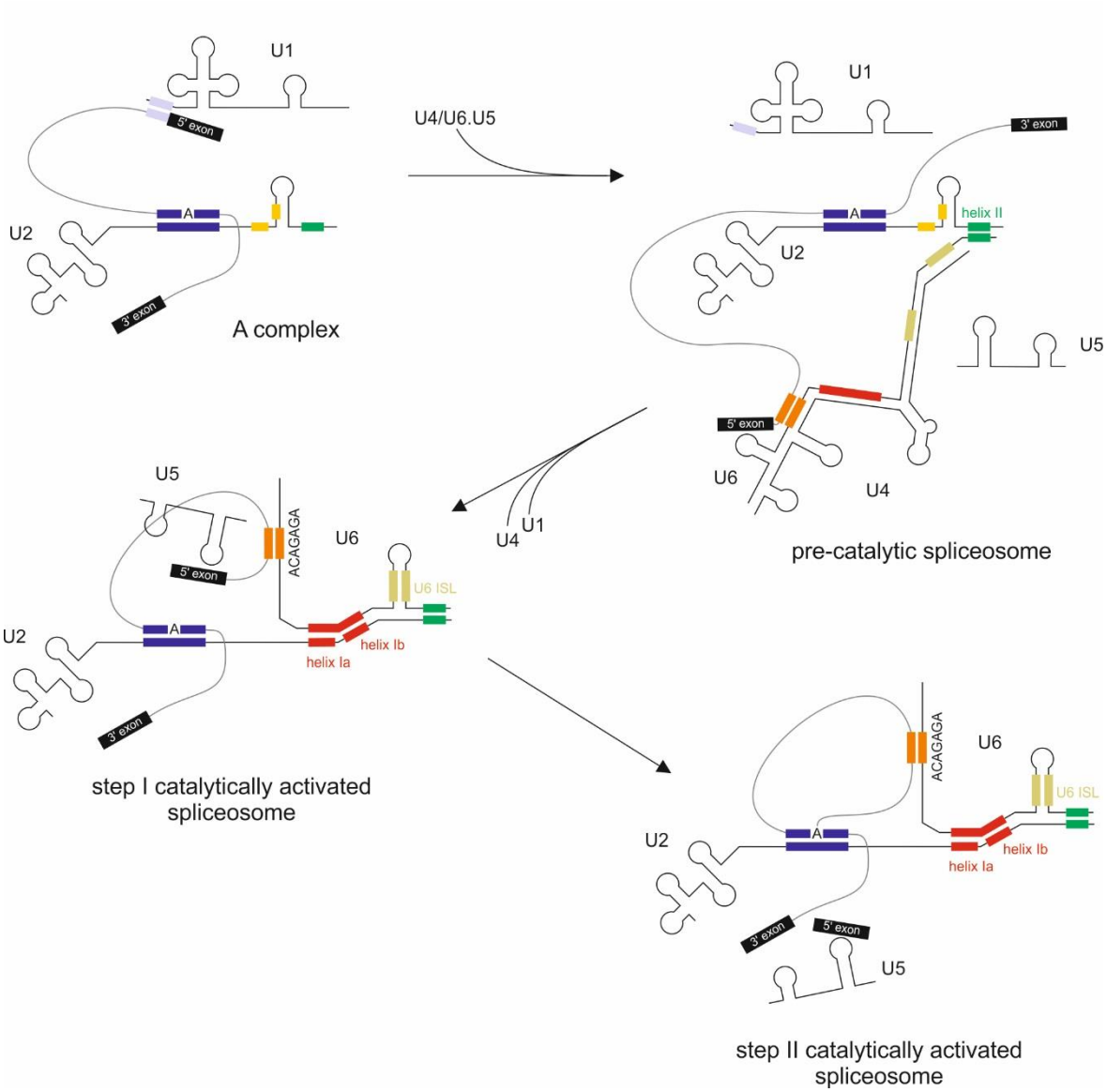


Figure 1.7 – **Changes in the spliceosomal RNA network during pre-mRNA splicing.** Schematic representation of the dynamic changes in the spliceosomal RNA network during splicing. In the A complex, the U1 snRNA base pairs with the 5' splice site (light blue) and the U2 snRNA forms the branch point helix with the pre-mRNA branch site (dark blue) and bulges the branch point adenosine out. After stable integration of the tri-

snRNP, the U2 snRNA and the U6 snRNA form contacts via base pairing in the helix II (green). After Prp28 displaces U1 the U6 ACAGAG base pairs with the 5'SS (orange). To form the step I catalytically activated spliceosome, U4 is displaced; the base pairing between U4 and U6 is disrupted, allowing U6 snRNA to form an internal stem-loop (ISL). The removal of the U4 snRNA also allows additional interactions between U6 and the U2 snRNA in the form of helices Ia and Ib. In addition to interactions with the U6 snRNA, the U2 snRNA remains in contact with the branch point region of the pre-mRNA. After the completion of the first step of splicing, internal rearrangements reposition the 5' and 3' exon into the catalytic center for the second step. The pre-mRNA exons are depicted as black boxes and the intron is represented by a grey line. Relative distances and sizes are not to scale.

1.8 Protein composition of the spliceosome

Most of the spliceosome's mass can be attributed to proteins. Spliceosomal proteins play crucial roles in the formation and maintenance of the catalytic RNA structure and facilitate all conformational and structural rearrangements of the spliceosome. The elucidation of the exact protein composition of the individual splicing complexes provides hints as to when a given protein functions. The ability to assemble spliceosomes *in vitro* and their purification made it possible to elucidate the individual protein composition of distinct stages (Wahl et al., 2009). Approximately 170 proteins associate with human spliceosomes throughout splicing (Wahl et al., 2009). Individual complexes like B, B^{act} and the C complex each contain around 120 proteins. There is a large exchange of proteins are most steps of the spliceosomal cycle (Figure 2.8) (Agafonov et al., 2011). The ca. 170 proteins associated with the human spliceosome are grouped into U snRNP associated proteins, which were described in Chapter 2.6, and the non-U snRNP associated proteins. The non-U snRNP proteins interact with the spliceosome at distinct stages either as individual proteins or as a part of larger sub-complexes. Some are associated with the spliceosome only transiently for distinct steps of splicing, and others remain bound throughout the splicing cycle. Major non-U snRNP complexes that interact with the spliceosome are the Prp19/CDC5L complex, the RES complex, the intron binding complex (IBC) and the exon junction complex (EJC).

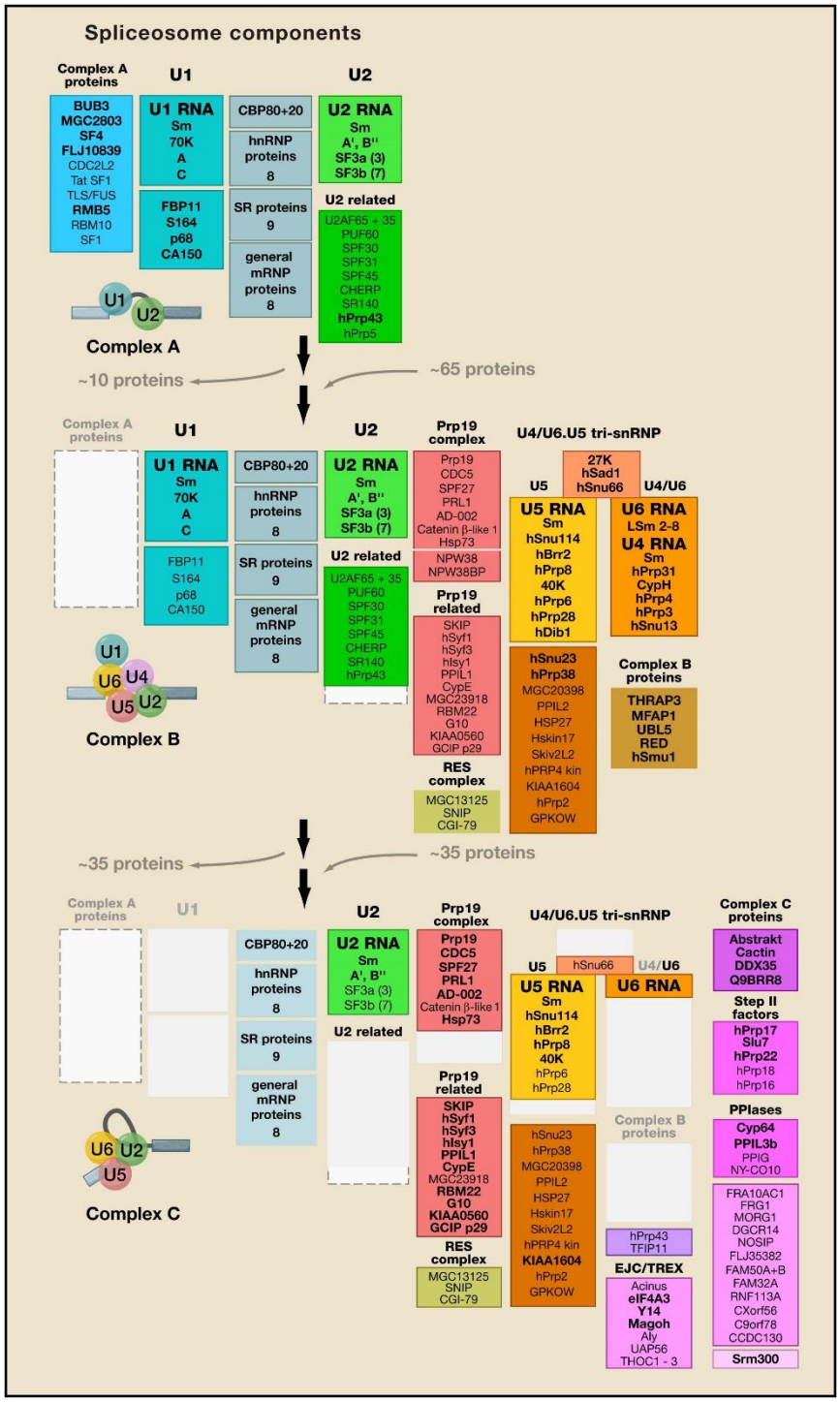


Figure 1.8 - **Spliceosomal protein dynamics during the splicing cycle**. The protein composition of the human spliceosomal A, B and C complex assembled in vitro. The proteins are grouped based on their snRNP and complex association, or function. Taken from (Wahl et al., 2009)

The Prp19/CDC5L complex is formed out of eight core proteins (PRP19, CDC5L, HSP73, β-catenin-like-1 CTNNBL1, PRL1, AD002 and SPF27) (Makarova et al., 2004) and is not only involved in the splicing of pre-mRNAs but also in protein degradation, the cellular DNA damage response and genome maintenance

(Chanarat and Strasser, 2013). The Prp19/CDC5L complex is recruited during the formation of the B complex and is essential for the catalytic activation of the spliceosome (Chan et al., 2003). The RES complex associates with the human spliceosomes during the formation of the B, B^{act} and C complex (Bessonov et al., 2010). It is required for splicing and plays a key role during spliceosome activation (Bao et al., 2017; Dziembowski et al., 2004). In addition, the RES complex functions in the nuclear export of mature mRNAs (Dziembowski et al., 2004). Another complex that binds the spliceosome during B^{act} formation is the pentameric Intron Binding Complex (IBC), which contains the helicase Aquarius (De et al., 2015). The EJC is a protein complex that assembles upstream of exon-exon junction during pre-mRNA splicing and consists of the proteins eIF4AII, Mago and Y14; the spliceosomal factor CWC22 is essential for EJC recruitment to the spliced RNA (Akhtar et al., 2019).

Many individual proteins also associate with the spliceosome at defined stages of the splicing cycle. For example, proteins of the SR family (see chapter 2.10.1), which promote spliceosome assembly, hnRNP proteins, and DExD/H-box RNA helicases that catalyze structural changes (see below). Other spliceosomal proteins include, among others, so-called Prp19/CDC5L-related proteins, B-specific proteins, and step I and step II factors that are required for the catalytic steps of splicing. The human spliceosome has many proteins that are missing in yeast (Fabrizio et al., 2009). These human-specific proteins bind at a particular stage (i.e. B-specific, C-specific proteins) and include PPIases and additional helicases, like Aquarius, not found in the yeast spliceosome.

1.9 The role of RNA helicases during splicing

Al least eight RNA helicases are required for pre-mRNA splicing: UAP65, Prp5, Prp28, Brr2, Prp2, Prp16, Prp18, Prp22 and Prp43 (Cordin et al., 2012). Their sequential activity is directly involved in triggering the structural and proteomic rearrangements in the spliceosome during splicing. All helicases involved in splicing belong to the superfamily two (SF2) helicases (3 DEAD-box, 4 DEAH-box and 1 Ski2-like helicase) (Cordin et al., 2012; Fairman-Williams et al., 2010; Koonin, 1991).

Helicases UAP56 and Prp5 are involved in the selection of the branch point sequence. UAP56 is recruited to the branch site by its interaction with U2AF65; UAP56 binding triggers ATP hydrolysis leading to the dissociation of the dimer. Upon the dimer release, the U2 snRNP can access the branch site (Shen et al., 2008). Prp5 is required for A complex formation and appears to remodel the 17S U2 snRNP so that it can bind stably to the BPS (O'Day et al., 1996; Ruby et al., 1993; Will and Lührmann, 2002). In yeast Prp5 activity is also required for the proofreading of the U2 branch site interactions and for the displacement of specific proteins (e.g. Cus2, the yeast homolog of hTAT-SF1) prior to U2 association (Perriman et al., 2003; Perriman and Ares, 2007). After initial tri-snRNP docking, Prp28 disturbs the interactions between the 5' splice site and U1 snRNA by removing stabilizing protein factors. The removal of the U1/5' SS interactions is directly connected to the establishment of the U6 snRNA/5' SS interactions and might point towards a proofreading function for Prp28 (Chen et al., 2001; Staley and Guthrie, 1999). The helicase Brr2 is a part of the U5 snRNP and tri-snRNP. It unwinds the U4/U6 snRNA in the B complex and in this way starts the catalytic activation process of the spliceosome (Absmeier et al., 2016). During B^{act} complex formation, Prp2 is recruited together with GPKOW/Spp2, a cofactor that is required for the activity of Prp2 (Silverman et al., 2004). Prp2 action leads to the destabilization of the interactions between the SF3a and SF3b complex with the branch point sequence (Warkocki et al., 2009). As a result, Prp2 remodels the catalytic center of the spliceosome and liberates the reactants of the first splicing step from their protein scaffolds (Bao et al., 2017; Silverman et al., 2004; Warkocki et al., 2009). The helicase activity of Prp16 is required during the transition from the step one spliceosome (the C complex) to the step two activated spliceosome (C*). After the helicase Prp16 triggers the conformational rearrangements required for the second step of splicing, Slu7, Prp18 and Prp22 are recruited to the spliceosome (Aronova et al., 2007; Warkocki et al., 2009). Prp18 and Slu7 are involved in the catalysis of the second step of splicing by holding the exons in position for the transesterification reaction (Aronova et al., 2007). Prp22 joins the spliceosome during the second step of splicing and it is necessary for the release of the spliced mRNA and U5 proteins from the disassembling spliceosome (Company et al., 1991; Schneider et al., 2002). After release of the mature mRNA from the spliceosome, Prp43 disassembles the

remaining post-splicing complex and releases the intron lariat (Martin et al., 2002; Tsai et al., 2007; Tsai et al., 2005).

1.9.1 Proofreading of spliceosome assemblies

As discussed above, the correct identification of splice sites is of crucial importance for cell viability. Therefore, the correct definition of splice sites is checked multiple times during spliceosome assembly. Interestingly, the helicases responsible for crucial structural rearrangements during splicing also appear to be responsible for fidelity control in some cases (Koodathingal and Staley, 2013). The identification of splice sites is directly connected to the ATPase activity of RNA helicases: correct splice site recognition increases the speed of ATP consumption, whereas incorrect splice site recognition decreases the ATPase activity of helicases (Egecioglu and Chanfreau, 2011), thus increasing the chance of the spliceosome being disassembled due to persistent stalling. In yeast, this phenomenon could be observed for multiple helicases at different stages of spliceosome assembly. For example, Prp5 appears to proofread the branch site selected during A complex assembly and Prp16 proofreads the U2/ BPS helix prior to catalytic step 2 (Burgess and Guthrie, 1993; Xu and Query, 2007).Taken together, the RNA helicases involved in splicing are not only involved in the structural rearrangements required to form the catalytic center and thereby complete RNA splicing, but they also have proofreading functions during these transitions.

1.10 Alternative splicing

The average pre-mRNA of a vertebrate is made up of multiple short exons separated by much larger introns. Exons are on average between 140 and 400 nucleotides long while introns can be up to several hundred thousand bases long (Sakharkar et al., 2004). Alternative splicing is the process of generating multiple different mature mRNA molecules out of one pre-mRNA transcript. The function of alternative splicing is to increase the genome-coding capacity of higher eukaryotes without increasing the genome size. This finding is supported by the observation that the prevalence of alternative splicing increases with genome complexity of the organism (Bush et al., 2017). In yeast only 5% of genes contain introns and only one example of functional alternative splicing is known (Juneau et al., 2009),

whereas in humans almost all genes are spliced and 92-94% of them were shown to have one or more alternative splicing products (Wang et al., 2008).

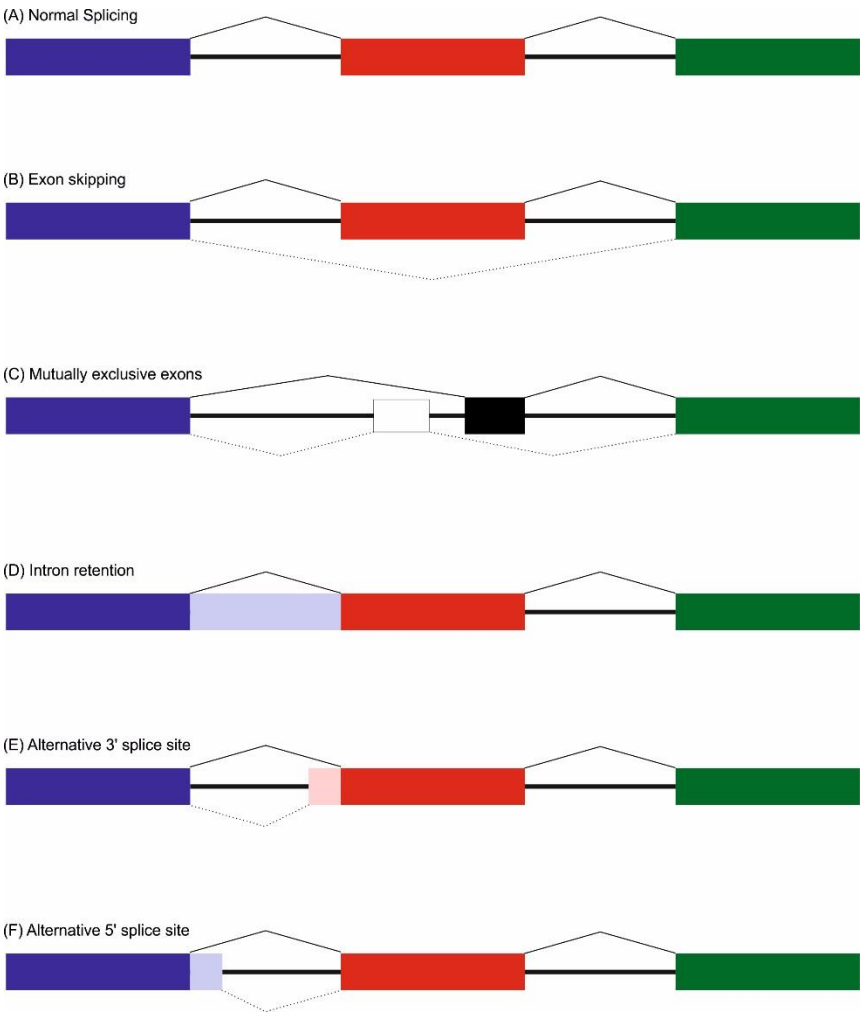


Figure 1.9 – **Different forms of alternative splicing of pre-mRNA.** Scheme of different alternative splicing events found in humans. Exons are depicted as colored boxes, or in case of the mutually exclusive exons as black and white box while introns are depicted as thick black lines. Splice site pairings are indicated using thin lines. (A) Normal splicing of pre-mRNA without alternative splicing events. (B) Exon skipping, in which an entire exon is skipped during splicing and does not appear in the mature mRNA. (B) Mutually exclusive exons entail that either exon is included into the mature mRNA (mutually exclusive). (D) Intron retention, during which an entire intron is kept as a part of the mature mRNA. (E/F) Alternative 5' and 3' splice sites can be used to create new mRNAs from pre-mRNA molecules. This figure was adapted from (Keren et al., 2010).

The most common mechanism of alternative splicing is exon skipping, during which an entire exon is not included into the mature mRNA (Fig. 2.9b). A rare special case of exon skipping is the presence of mutually exclusive exons in the pre-mRNA. Here only one of two exons can be present in the mature mRNA while the other one is skipped entirely (Fig 2.9c). In addition to excluding exons from the mRNA, it is also possible that introns are not removed from the processed mRNA; this event is called

intron retention. Finally, along with omitting intended splice sites entirely, alternative 5’ and 3’ splice sites also can be used. They are found in spatial proximity to the canonical exon borders and cause alterations in the exon/intron borders of the pre-mRNA without including or removing exons completely (Figure 2.9e/f).

1.10.1 Regulating alternative splicing

Alternative splicing is tightly regulated by multiple different factors. First, splice sites in human cells are much less conserved than their counterparts in yeast suggesting that splice sites could show varying affinities to their respective splicing factors. Second, RNA secondary structures may shield splice sites making them unavailable for splicing under certain conditions (Keren et al., 2010). In addition, various regulatory proteins interact with short RNA sequences and influence the recruitment of spliceosomal factors (Figure 2.10). These proteins are known as SR-proteins and hnRNPs (Lin and Fu, 2007). SR proteins contain arginine-serine repeats in their RS domain. This domain can interact with the RS domains found in multiple splicing factors and enhance their recruitment and subsequent complex assembly by forming or stabilizing protein-protein interactions. SR proteins typically bind to intronic or exonic splice enhancers (ISE/ESE). The hnRNPs are antagonists of SR proteins, they bind to RNA silencing RNA elements known as the intronic and exonic splice silencers (ISS/ESS), and negatively influence spliceosome recruitment and assembly (Martinez-Contreras et al., 2007). The effect of both SR proteins and hnRNPs is confined to their surroundings and can thereby selectively enhance or decrease the usage of a particular splice site.

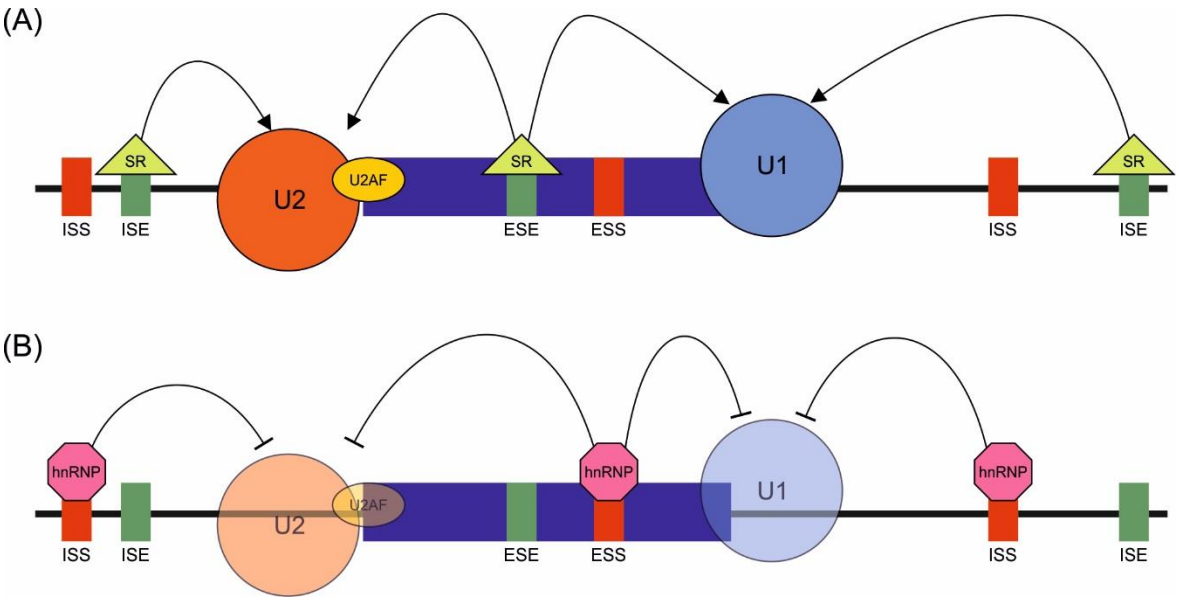


Figure 1.10 – **Regulation of alternative splicing.** The presence of splicing regulators in form of exonic splice enhancers (ESE) and silencers (ESS) as well as their intronic counterpatrts (ISE/ ISS) can modulate the assembly of the spliceosme on the splice sites. SR proteins (light green triangles) enhance the recruitment of snRNPs and proteins and stablize splicesome assembly, while hnRNPs (pink octagons) interfere with snRNP recruitment and spliceosome assembly.The exon is represented as blue box while the flanking introns are shown as thick black lines. The spliceosomal snRNP particles are represented as circles (U1 – light blue, U2 – orange) and the poly-pyrimidine binding protein complex U2AF is shown in yellow. Fading of bound proteins in (B) is used to indicate decreased recruitment.

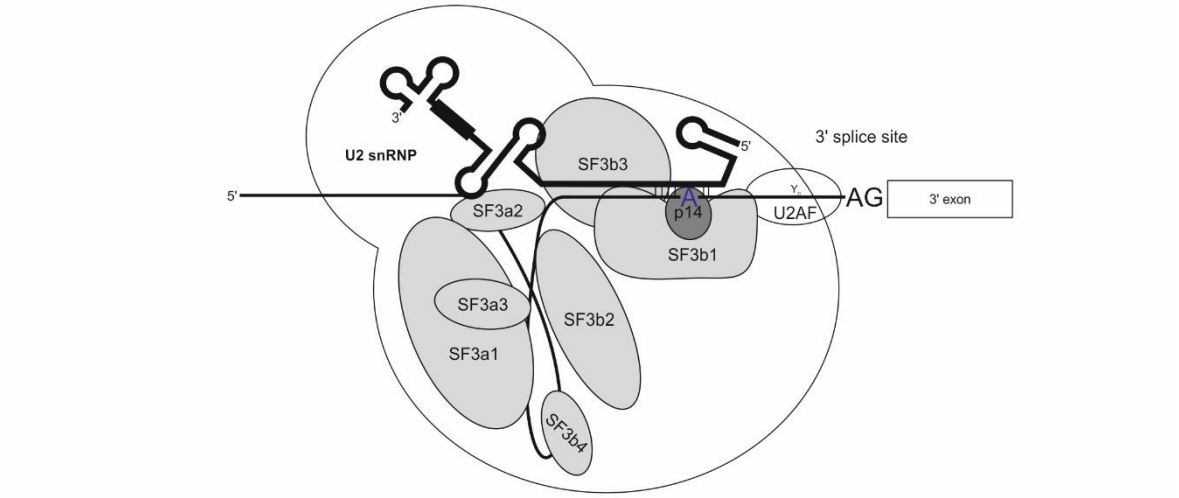
In addition to the pre-mRNA internal regulatory elements, the transcription speed of RNA polymerase II is also thought to have a direct influence on alternative splicing. Transcription speed and pausing events impact the splice sites availability for spliceosome formation and thereby significantly influence the splice site pairing possibilities available during spliceosome formation.

1.11 Structure of SF3b and the U2 snRNP

The U2 snRNP plays a key role during A complex formation and the selection of the branch point. The 17S U2 particle is assembled in a stepwise process, which gives rise to two intermediate particles prior to the formation of the mature 17S U2 particle. The first assembly step is the formation of the Sm-ring at the Sm-site and the recruitment of the U2A' and U2B'' complexes to the stem-loop IV of the U2 snRNA, thus forming the 12S U2 particle (Lührmann et al., 1990). After the formation of the 12S U2 particle, the SF3b protein complex binds and the 15S U2 particle is assembled, followed by the binding of the SF3a complex giving rise to the mature 17S U2 particle (Nesic and Kramer, 2001). The U2 snRNP proteins form an

extensive network of protein-protein and protein-RNA interactions that are necessary for the stable integration of the U2 snRNP in the A complex.

(A)



(B)

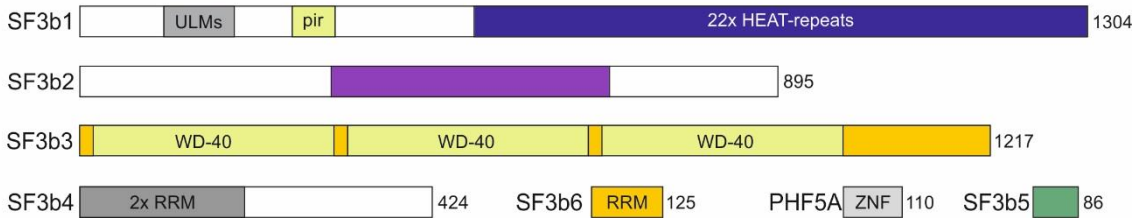


Figure 1.11 – **Model of the U2 snRNP interaction with the intron and domain structure of human SF3b proteins.** Schematic representation of the SF3a and SF3b proteins (grey) in the 17S U2 particle that contact the pre-mRNA (Gozani et al., 1996). The U2 snRNA (thick black line) and the 3' end of the pre-mRNA (thin black line) are shown schematically. The 3' exon is indicated as box and 3' splice site (AG) and branch point adenosine (blue A) are shown as letters. The polypyrimidine tract is indicated as Y_n. Adapted from (Will et al., 2001). (B) Domains of Sf3b proteins: Znf, Zinc finger; RRM, RNA recognition motif; pir, SF3b6 interacting region; ULM, U2AF ligand motif. The number at the C-terminus indicates the number of amino acids.

Upon binding of U2 to the pre-mRNA, the SF3a and SF3b proteins interact with the RNA at the so-called anchoring site, roughly 25 nucleotides upstream of the branch point (Gozani et al., 1996). The SF3a proteins SF3a3 and SF3a2 contact the RNA directly via their zinc-finger domains and both interact with SF3a1 (Nesic and Kramer, 2001). The SF3b complex is a heptameric complex (SF3b1, SF3b2, SF3b3, SF3b4, SF3b5, PHF5A and SF3b6) (Figure 2.11b). Several SF3b proteins contact the pre-mRNA upstream of the branch site, while SF3b1 can be cross-liked to RNA nucleotides up- and downstream of the branch site (Gozani et al., 1996; Will et al., 2001). SF3b1 also forms protein-protein interactions with U2AF65. In addition, the SF3b protein p14 is in direct contact with the branch point adenosine (Will et al.,

2001). Together, these interactions play an important role in the stabilization of the U2/BPS helix and tether the U2 snRNP to the BPS.

Recently, the crystal structure of an isolated SF3b complex, which lacked SF3b2, SF3b4 and the N-terminal domain of SF3b1, was reported (Cretu et al., 2016). The C-terminal HEAT-domain of SF3b1 contains 20 tandem repeats that adopt a super helical structure (Cretu et al., 2016) (Figure 2.12a). The conformation of HEAT domain is directly influenced by its binding partners (Grinthal et al., 2010; Zachariae and Grubmuller, 2008). The HEAT repeats of SF3b1 are in direct contact with SF3b3, SF3b5 and PHF5A. PHF5A contacts the concave (inside) surface of the HEAT domain and bridges the N-terminal and C-terminal HEAT repeats (H2+H3). The WD-40 protein SF3b3 and SF3b5 also make multiple contacts with the SF3b1 HEAT domain and influence its conformation.

Advances in cryo-electron microscopy enabled the visualization of the near atomic structure of the spliceosome, making it possible to observe the structural changes accompanying the transition into different stages of assembly (Fica and Nagai, 2017). High resolution cryo-EM models of the human pre-B, B, B^{act}, C, C* and P complexes have been obtained (Bertram et al., 2017a; Bertram et al., 2017b; Charenton et al., 2019; Haselbach et al., 2018; Zhang et al., 2017; Zhang et al., 2018). At present no high-resolution structures of the human A complex or the 17S U2 snRNP are available. The cryo-EM structure of the human B^{act} complex reveals how the SF3b1 HEAT domain and other SF3b proteins interact with the U2/BPS helix and 3’ region of the intron (Figure 2.12). The U2/BPS helix is clamped between the N-terminal and C-terminal HEAT repeats and the BPS is bound in a pocket formed by PHF5A and the SF3b1 HEAT domain. Numerous other proteins interact with the HEAT domain within the spliceosome.

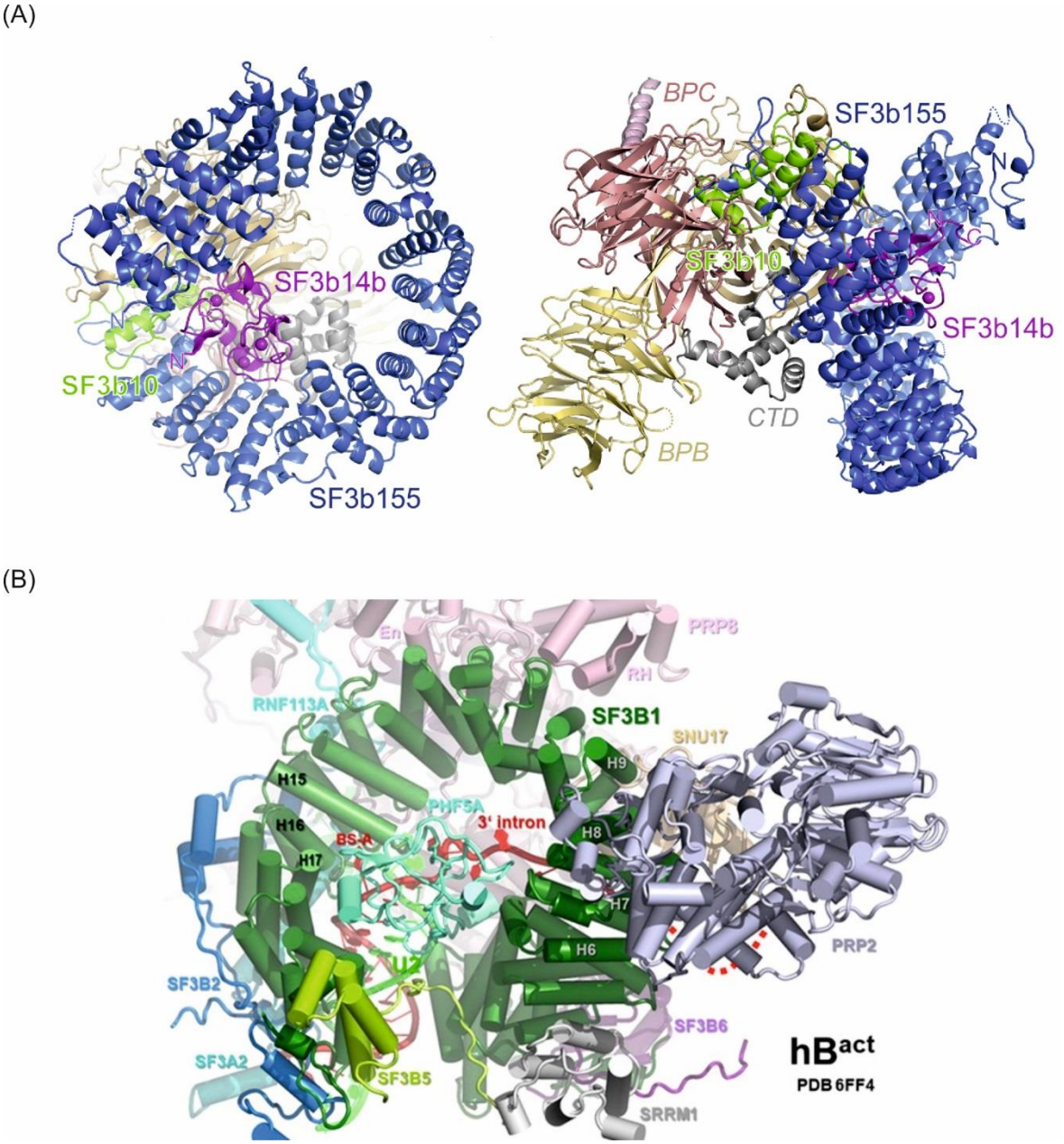


Figure 1.12- **Pseudo atomic model of SF3b interactions with the pre-mRNA in the human B^{act} complex.** (A) Crystal structure of the SF3b core. Depicted are the HEAT-domain of SF3b1 and then bound SF3b10 and SF3b14 proteins. (B) Scheme of the SF3b1 HEAT domain as seen in the human B^{act} complex. The intronic RNA and the U2 snRNA form the U2/ BPS helix and is sequestered by the N- and C-terminal end of the HEAT domain. The intronic RNA downstream of the branch point adenosine then leaves the HEAT domain in the region between HEAT repeat 6 and 7. Image taken from (Cretu et al., 2016) and (Kastner et al., 2019).

Comparison of the structure of the SF3b1 HEAT domain in the isolated complex with its structure in the human B^{act} complex revealed a major structural rearrangement within the HEAT-domain of SF3b1 (Figure 2.13). While the HEAT-domain of SF3b1 adopts an open conformation in the isolated SF3b complex in the B^{act} complex the HEAT-domain is closed. The closed HEAT-domain, which also appears to be present in the human B complex is thought to clamp the RNA in place,

to stabilize the U2 interaction and to sequester the BPS adenosine to prevent premature splicing catalysis.

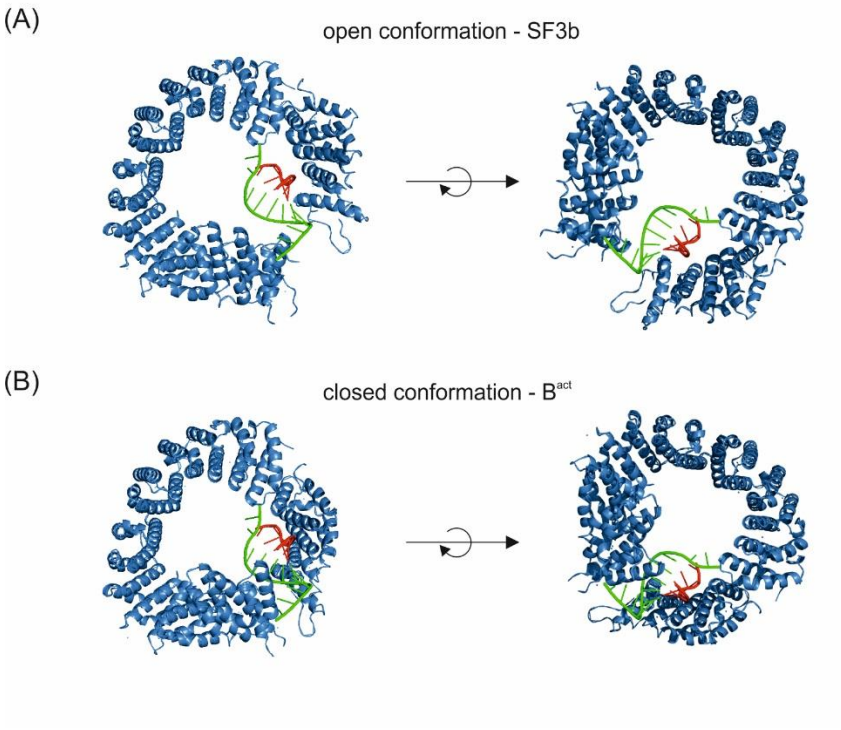


Figure 1.13- **Open and closed conformation of SF3b1 during splicing.** Reconstruction of SF3b1 in an open conformation as found in the SF3b complex and in a closed conformation as seen in the B^{act} complex. The HEAT domain of SF3b1 is shown in blue, the U2 RNA in green and a part of the pre-mRNA involved in the formation of the branch point helix in red. Structures were extracted from the SF3b structure (Cretu et al., 2016) and the human B^{act} structure (Haselbach et al., 2018).

1.12 Spliceosomal mutations in cancer

Misregulated splicing of pre-mRNA is related to carcinogenesis and could be often observed in cancer cells (David and Manley, 2010). In addition, up to 50% of cancer-causing mutations are found in genes whose products have splicing-related functions (Cartegni et al., 2002). With the emergence of affordable transcriptome sequencing, multiple mutational hotspots in splicing-associated proteins were discovered in samples obtained from patients with hematological malignancies (Yoshida and Ogawa, 2014). The mutational targets were spliceosomal proteins associated with the initial assembly steps of the spliceosome, for instance, U2AF1, SRSF2, ZRSR2 and SF3B1. All four of these proteins are directly involved in the definition of splice sites during early spliceosome assembly and their mutations

occur in patients in a mutually exclusive manner, suggesting lethality of multiple mutations.

SF3B1 is a common mutational target in hematopoietic cancers like myelodysplastic syndromes (MDS), chronic myelomonocytic leukemia (CMML) and acute myeloid leukemia (AML) (Yoshida and Ogawa, 2014). Up to 70% of cancer patients contain one of the hot-spot mutations. The SF3b1 protein can be divided into 3 structural domains. The N-terminal region of the protein is an unstructured stretch of 528 amino acids, followed by a structured HEAT-domain (529-1201aa) comprised of 20 HEAT repeats. The C-terminal end is also comprised of an unstructured stretch of 102 amino acids (1202-1304aa). The mutations of SF3b1 associated with cancer mainly cluster within the HEAT-repeat domain of the protein with the most common mutations being K700E/Q699_K700del, K666N/T/E/R/Q/M and H662Q/D/Y (Yoshida and Ogawa, 2014) (Figure 2.14a). In the three-dimensional structure of the SF3b1 protein, most of its mutations could be seen clustered within a specific region that is in close proximity to the sixth HEAT-repeat where the intron exits the HEAT domain (Cretu et al., 2016) (Figure 2.14b). The nature of the amino acid substitutions suggests that the surface charge of the SF3b1 protein might change in the mutated regions.

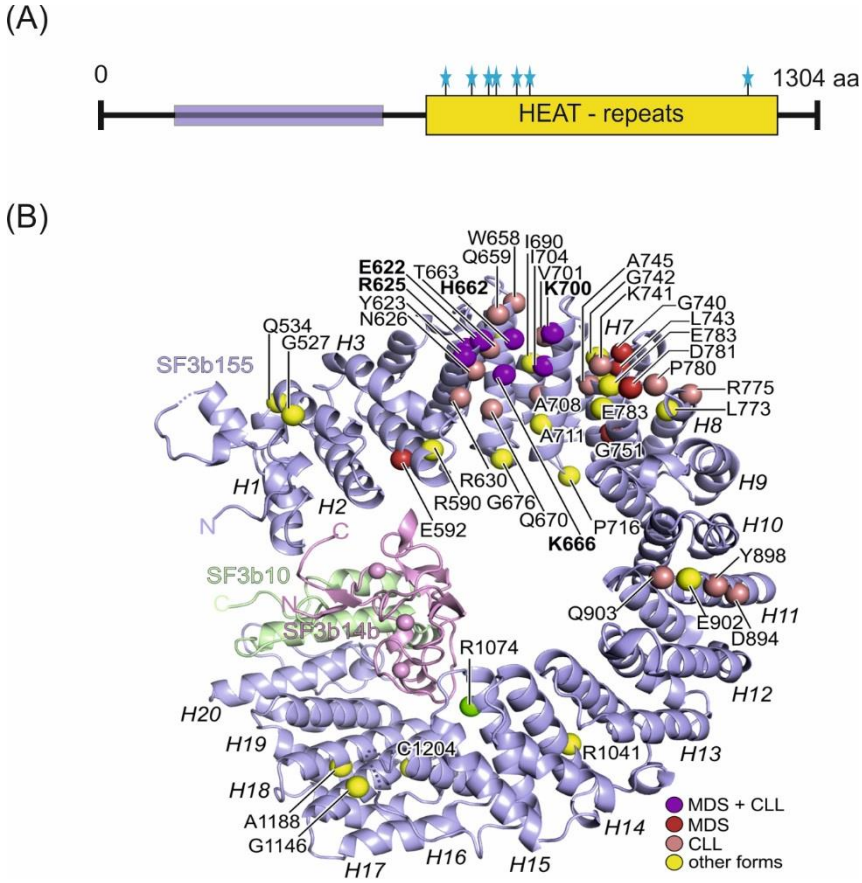


Figure 1.14 – **Location of the cancer related mutations in the 3D-structure of SF3b1.** (A) Schematic representation of SF3b1 domains, the unstructured N-terminal region containing the TP-repeats (transparent blue), the target region for phosphorylation and the structured HEAT-repeat domain (yellow). (B) The 3D-structure of the HEAT-domain of SF3b1. Balls indicate the position of various cancer-associated mutations found in SF3b1. Taken from (Cretu et al., 2016).

The presence of SF3b1^{K700E} in the cell leads to the emergence of aberrantly spliced pre-mRNA molecules. The alterations of splicing mainly affect the selection of the 3' splice site, with cryptic 3' splice sites being used only in the presence of the K700E mutation (Agrawal et al., 2017). These cryptic splice sites are found upstream of the canonical splice site and the poly-pyrimidine tract and are associated with their own distinct branch point sequence (Darman et al., 2015). Experimental evidence suggests that the usage of the cryptic splice site is directly coupled to the presence of a cryptic branch point, as removal of the latter results in the disappearance of cryptic splice products (Darman et al., 2015).

Table 1.2 – **Usage of cryptic splice sites is dependent on SF3b1^{K700E} expression**

	SF3b1^{wt} + SF3b1^{wt}	**SF3b1^{wt} + SF3b1^{K700E}**
Canonical splice sites	spliced	spliced
Cryptic splice sites	not spliced	spliced

Bioinformatics analysis of total cellular RNA further identified that the cryptic branch points are located 12-18 nucleotides downstream of their respective branch point while the canonical AG dinucleotide is found more than 18 nucleotides downstream of the branch point adenosine (Deboever et al., 2015). However, the characteristic features of the cryptic branch point and 3' splice site are poorly conserved and do not suffice to explain the changes in the splicing patterns of the affected cells. Cancer related mutations in SF3b1 have been proposed to alter the conformation of the HEAT domain and in this way destabilize the U2/BPS interaction, which could affect the usage of alternative BPS and 3' splice sites. As an alternative, the mutations were proposed to alter the path of the 3' end of the intron and in this way affect 3' splice site usage (Jenkins and Kielkopf, 2017).

1.13 Small molecule inhibitors of splicing that interact with SF3b1

High-throughput screens have led to the identification of multiple compounds that can inhibit pre-mRNA splicing *in vitro* and/or *in vivo*. These compounds interfere with the normal function of the spliceosome causing general or gene specific changes in splicing. The mechanism used to inhibit pre-mRNA splicing is varied and ranges from interference during spliceosome assembly to steric modifications of the RNA. Many natural splicing inhibitors interfere with the assembly of the spliceosome and cause global changes in splicing patterns leading to cell-cycle arrest and apoptosis (Effenberger et al., 2014a; Effenberger et al., 2013; Folco et al., 2011; Hasegawa et al., 2011; Kaida et al., 2007; Kotake et al., 2007; Pawellek et al., 2014; Vigevani and Valcárcel). Interestingly these inhibitors all interact with the SF3b1 protein and stop spliceosome assembly at an early assembly stage. The small molecule inhibitors targeting SF3b1 can be grouped into four classes based on their chemical structures (Bonnal et al., 2012; Effenberger et al., 2017). These are the pladienolides (A-G, E7107), spliceostatins (A-G, FR901464, thainlanstatins, meayamycin), herboxidiene (GEX1A), and sudemycins (C1, D6, E, F1). It was proposed that while chemically rather diverse, all inhibitors of SF3b1 share a common pharmacophore defined by a distinct set of functionally required groups (Lagisetti et al., 2008). Importantly, these compounds have been shown to selectively inhibit cancer cell growth.

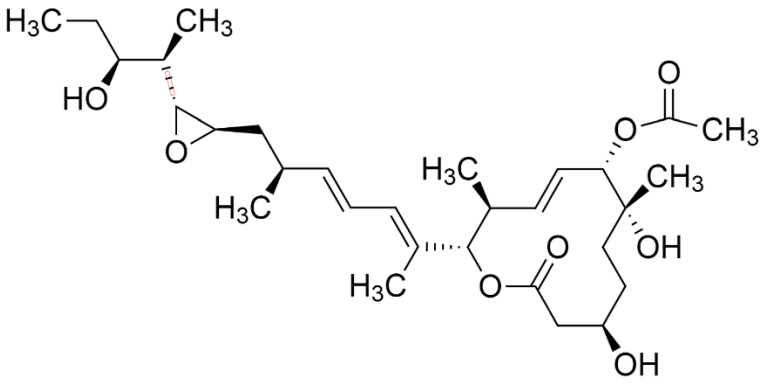


Figure 1.15 – **Structure of Pladienolide B.** The Pladienolide B molecule is composed of a macrolide group connected to an aliphatic arm by a diene group acting as linker.

The recent crystal structure of Pladienolide B bound to SF3b1 revealed the exact binding location of Pladienolide B and could show that the inhibitor binds to the open conformation of the protein (Cretu et al., 2018). To date it is still not known how exactly the presence of Pladienolide B influences the assembly of the spliceosome and which mechanism leads to the aforementioned changes in the cellular splicing pattern. A better understanding of the effects of Pladienolide B on the splicing machinery could help improve therapeutic approaches using these compounds.

1.14 Aim of this thesis

SF3b is a heptameric, U2 snRNP subunit that is essential for the stable interaction of the U2 particle with the branch point sequence of an intron during early spliceosome assembly and therefore aids in the recognition and selection of the branch site adenosine. SF3b1 is the largest protein of the SF3b complex and it is the target of several small molecule splicing inhibitors, including Pladienolide B, that have anti-tumor properties and alter alternative splicing patterns. At the same time, mutations in SF3b1 have been linked to a variety of cancers and cause changes in alternative splicing events that lead to malignancies. Thus, modulating the function of SF3b can have both positive and negative effects on modulating the onset or severity of a variety of human diseases. However, the molecular mechanisms how this is achieved are not completely understood.

I wanted to address two major questions in my thesis regarding the function of Pladienolide B as an inhibitor of spliceosome assembly. First, what are the consequences of Pladienolide B binding on the assembly of the spliceosome? Secondly, what compositional and structural changes arise due to the presence of the inhibitor? To investigate this, I set out to affinity purify human spliceosomal complexes stalled by Pladienolide B and also wild-type complexes for comparison, and check for differences in their protein compositions using mass spectrometry. I also hoped to obtain a high resolution cryo-EM structure of the wild-type, human spliceosomal A complex, which is currently not available, and the Pladienolide B stalled A complex, that would reveal any structural differences especially in the conformation of the SF3b1 HEAT domain. These studies should clarify if indeed Pladienolide B prevents a conformational change in SF3B1 that was proposed to occur when U2 interacts during A complex assembly and leads to clamping down of the HEAT domain on the U2/BPS helix.

In a second set of experiments, I also wanted to investigate how cancer-linked point mutations in SF3b1 lead to the recently observed changes in alternative splicing patterns in various cancers. To date the exact mechanism responsible for these changes could not be identified. I initially aimed to purify 17S U2 snRNPs containing the SF3b1 mutation K700E, one of the most common cancer-related mutants, and investigate its effect on the composition, structure and function of the U2 particle, and then to expand this to both cross-intron and cross-exon spliceosomal complexes containing the mutated SF3b1 protein. Another goal was to recapitulate in vitro, alternative splicing events dependent on SF3b1 mutations that were observed in vivo, in order to dissect the mechanisms that lead to the use of cryptic branch sites and 3' splice sites. These investigations will improve our understanding of how exactly the usage of cryptic branch points arises as a consequence of the single amino acid substitutions in SF3b1. These studies could provide valuable information concerning how changes in the function of SF3b1 due to single point mutations lead to cancer. Understanding the exact mechanism underlying these changes could potentially help the creation of targeted and improved disease therapies.

2 Material

2.1 Chemicals

β-Mercaptoethanol	Roth, Germany
Accutase	ThermoFischer, USA
Acetic acid	Merck, Germany
Agarose (low melting point)	ThermoFischer, USA
Agarose	ThermoFischer, USA
Ammonium peroxidisulphate (APS)	Merck, Germany
Ampicillin	Sigma-Aldrich, Germany
Anacardic acid	Alexis biochemical, USA
Bromophenol blue	Merck, Germany
Coomassie brilliant blue G-250	Serva, Germany
Creatine phosphate	Sigma-Aldrich, Germany
DMEM	ThermoFischer, USA
Dimethylsulphoxide (DMSO)	Roth, Germany
Dipotassiumhydrogenphosphate	Merck, Germany
Dithioerythritol (DTE)	Roth, Germany
Dithiothreitol (DTT)	Roth, Germany
DNA-molecular weight marker	Gibco, New Zealand
EDTA (Disodium salt dihydrate)	Roth, Germany
Ethanol	Merck, Germany
FBS	ThermoFischer, USA
FLAG-peptide (3x)	Sigma-Aldrich, Germany
Formaldehyde	Merck, Germany
Formamide	Merck, Germany
GlutaMax	ThermoFischer, USA
Glycerol	Merck, Germany
Glycine	Merck, Germany
Glycoblue	Ambion, USA
Heparin (sodium salt)	Roth, Karlsruhe
HEPES (N-2-Hydroxyethylpiperazin-N-2-ethansulfonic acid)	Calbiochem, USA
IMEM	ThermoFischer, USA
Imidazole	Merck, Germany
Maltose	Merck, Germany
Methanol	Merck, Germany
MOPS buffer	ThermoFischer, USA
Milk powder	Roth, Germany
Nonidet P-40 (Igepal CA-630)	Sigma-Aldrich, Germany
Penicillin/Streptomycin	ThermoFischer, USA
Pladienolide B	ChemCruz, USA
Pluronic F-86	ThermoFischer, USA
PMSF (Phenylmethylsulfonylfluoride)	Roche, Germany
Potassium Chloride	Merck, Germany
Potassiumdihydrogenphosphate	Merck, Germany
Pre-stained protein-molecular weight marker	Bio-Rad, Germany
Roti-Phenol-Chloroform-Isoamyl alcohol (PCI)	Roth, Germany
Rotiphorese Gel 30 solution	Roth, Germany
Rotiphorese Gel 40 solution	Roth, Germany
Sodium Chloride	Merck, Germany
Sodiumacetate	Merck, Germany
Sodiumdodecylsulfate (SDS)	Serva, Germany
Sucrose	Merck, Germany

TEMED (N, N, N', N'-Tetramethylethylendiamine)	Sigma-Aldrich, Germany
Tris-(hydroxymethyl)aminomethane (Tris)	Roth, Germany
Triton X-100	Merck, Germany
Tween 20	Sigma-Aldrich, Germany
Urea	Merck, Germany
Xylene cyanol	Fluka, Switzerland

2.2 Chromatography materials and consumables

Amylose resin	New England Biolabs, USA
Cassettes for film exposure	Kodak, USA
Centrifuge tubes	Beranek, Germany
Chromatography columns	Bio-Rad, Germany
Concentrator	Millipore, USA
HisTrap HP, Ni-NTA-Sepharose (5 ml)	GE Healthcare, Germany
NuPAGE™ gels (1.5 mm, 4-12%)	ThermoFischer, USA
Parafilm	Roth, Germany
Poly-Prep columns	Bio-Rad, USA
Protran BA 83 nitrocellulose	Whatman, UK
ProbeQuant™ G-25/ G-50 micro columns	GE Healthcare, UK
Slide-A-Lyzer dialysis units (MWCO 6 kDa)	Pierce, USA
SnakeSkin dialyisis membranes (MWCO 6000-8000 Da)	Thermo Fischer Scientific, USA
Sterile filters (0.2 Mm or 0.45 Mm)	Sarstedt, Germany
Surgical blades	Martin, Germany
Whatman 3MM Paper	Whatman, UK
X-ray films (BioMax MR)	Kodak, USA

2.3 Cell culture materials

0.2l, 1l, 6l spinner flasks	Bellco, USA
T-75 flasks	Sarstedt, Germany
T-175 flasks	Sarstedt, Germany

2.4 Machines

ÄKTA Start	GE Healthcare, UK
Amersham Imager 680	GE Healthcare, UK
Autoclaves	H+P Labortechnik, Germany
Axiovert 25	Zeiss, Germany
Biofuge fresco	Kendro, USA
Biofuge pico	Kendro, USA
FLA-7000	Fuijfilm, Japan
Gel documentation unit	Bio-Rad, Germany
Gel electrophoresis apparatus	in-house
Gel dryer Model 583	Bio-Rad, Germany
Gradient Master Modell 106	BioComp Instruments
Head-over-tail Rotor	Cole-Parmer, USA
Heating blocks	Eppendorf, Germany
HeraSafe	Kendro, USA
Hybridization oven Hybaid	Biometra, Germany
Incubator BD6220	Haraeus, Germany
Invitrogen Safe Imager Blue-Light Transilluminator	ThermoFischer, USA

Megafuge 1.0R	Kendro, USA
Milli-Q-water supply apparatus	Millipore, USA
NanoDrop	ThermoFischer, USA
PerfectBlue Semi-Dry Electro blotter	PeqLab, Germany
pH-Meter	Mettler Toledo, Switzerland
Phosphorimager Typhoon 8600	Amersham Pharmacia, UK
Power supply EPS 2A200	Hoefer, USA
Power supply EPS 3501/XL	Amersham Pharmacia, UK
Q-ToF Ultima mass spectrometer	Waters, USA
Scintillation counter LS 1701	Beckman/Packard, USA
Sorvall HB-6 rotor	Kendro, USA
Sorvall SA800 AT4 Rotor	Kendro, USA
Sorvall SS-34 Rotor	Kendro, USA
Sorvall TH660 Rotor	Kendro, USA
Sorvall TST41.14 Rotor	Kendro, USA
Speed Vac Konzentrator 5301	Eppendorf, Germany
Spectrophotometer Nanodrop ND-1000	Thermo Scientific, USA
Spectrophotometer Ultrospec 3000 pro	Amersham Pharmacia, USA
SS-34 Rotor	Kendro, USA
Tabletop centrifuge 5415D	Eppendorf, Germany
Titan Krios	FEI, USA
Trans-Blot Cell	Bio-Rad, USA
Ultracentrifuge Discovery 90	Sorvall, Kendro
Ultracentrifuge Discovery M150	Sorvall, Kendro
Variomag Biomodul 40B	H + P Labortechnik, Germany
Vortex	Janke & Kunkel, Germany
X-ray film developer X-Omat 2000	Kodak, USA

2.5 Nucleotides

Nucleoside – 5’ – triphosphate (100 mM): ATP, CTP, UTP, GTP	Pharmacia, Sweden
Deoxynucleoside – 5’ – triphsophate (100 mM): dATP, dCTP, dTTP, dGTP	Pharmacia, Sweden
m^7G(5’)ppp(5’)G-cap (7-monomethyl-diguanosine triphosphate)	Kedar, Poland

2.5.1 Radiolabeled nucleotides

α-^{32}P-UTP [10 µCI/µl, 3000 Ci/mmol]	Perkin Elmer, USA

2.5.2 Oligonucleotides

2.5.2.1 Inserted tag DNA oligonucleotides for SF3b1

Tag	**Sequence (5’-3’)**
N-terminal SBP tag for SF3b1/SF3b1^{K700E}	TAGATCTCTCGAGGTTAACGAATTCGAGGATCCAGGCGGCA GCGGCGGCAGCATGGACGAGAAGACCACCGGCTGGAGAGG CGGCCACGTGGTGGAGGGCCTGGCCGGCGAGCTGGAGCA GCTGAGAGCCAGACTGGAGCACCACCCCCAGGGCCAGAGA

C-terminal His-tag for SF3b1/SF3b1^{K700E}	GAGCCCGGCGGCAGCGGCGGCAGCGAGAACCTGTACTTCC AGGGCGGCGGCAGCGGCGGCAGCTACCCCTACGACGTGCC CGACTACGCCGCTGGAGCATACCCCTACGACGTGCCCGACT ACGCCGCTGGAGCATACCCCTACGACGTGCCCGACTACGC CTGAGTCTAGGCGGCCGCCAATTCTACCGGGT GGACTACATCCTGCACCACCATCACCACCATCATCACCACCA CTAAGTCGACACCTGCGGCCGCTAACATGGCGCGCC

2.5.2.2 RNA oligonucleotides

5'splice site RNA oligonucleotide	5'-CUGUUCAGGUAAGUAU

2.6 Commercial kits

BCA protein assay kit	Pierce, USA
ECL Western Blot Detection Kit	GE Healthcare, USA
Invitrogen® PureLink Maxiprep	Invitrogen, USA
PCR & Gel Clean-Up Kit	Macherey-Nagel, Germany
TMTduplex™ Isobaric Mass Tagging Kit	ThermoFischer, USA

2.7 Enzymes

cOmplete protease inhibitor, EDTA-free	Roche, Switzerland
Proteinase K	Sigma-Aldrich, USA
Restriction endonucleases	New England Biolabs, USA
RNasin (RNase inhibitor) [40U/µl]	Promega, USA
RQ1 DNase [1U/µl]	Promega, USA
T4 Polynucletide kinase (PNK)	New England Biolabs, USA
T7 RNA polymerase	New England Biolabs, USA
YIPP [0.1U/µl]	New England Biolabs, USA

2.8 Antibodies

Anti-FLAG M2	Sigma-Aldrich, USA
Anti-SF3b1.1	AG Lührmann
Anti-SF3b1 (13E12)	AG Lührmann
Anti-hPrp5	AG Lührmann
Anti-hTAT SF1	AG Lührmann

2.9 Plasmids

pMINX	Pre-mRNA construct (MINX) in pUC18-vector under control of a T7 promoter, Amp^R. (Zillmann et al., 1988)
pMINX-exon	MINX exon RNA construct in pSP65-vector under control of a T7 promoter, Amp^R. (Schneider et al., 2010b)
pMINX-exon variants	MINX exon RNA derivate synthesized externally into the pUC57 vector under control of a T7 promoter, Amp^R.
pMS2-MBP-StrepII	Bacterial MS2-MBP-StrepII expression vector, Amp^R.
pWPI	2nd gen bicistronic lentiviral vector allows for simultaneous expression of a transgene and EGFP marker to facilitate tracking of transduced cells, Amp^R.
psPAX2	2nd generation lentiviral packaging plasmid, Amp^R. (Didier Trono)

pMD2.G	VSV-G envelope expressing plasmid, Amp^R. (Didier Trono)

2.10 Cell lines

2.10.1 Bacterial cell lines

XL-10 gold	Agilent, USA

2.10.2 Human cell lines

HeLa S3	ATCC, USA
HEK293T	ATCC, USA
K562	ATCC, USA
K562 FLAG-SF3b1^{wt}	Dr. James Manley
K562 FLAG-SF3b1^{K700E}	Dr. James Manley

2.11 Buffers and Media

2.11.1 Media

LB-Medium		
	Yeast extract	0.5% (w/v)
	Tryptone	1% (w/v)
	NaCl	1% (w/v)
	pH	7.0

SOC-medium		
	Tryptone	2% (w/v)
	Yeast extract	0.5% (w/v)
	NaCl	10 mM
	KCl	2.5 mM
	MgCl2	10 mM
	Glucose	20 mM

DMEM-medium complete		
	DMEM	
	Fetal bovine serum	10% (v/v)
	L-glutamine	2 mM
	Penicillin G	100 U/mL
	Streptomycin	100 U/mL

IMEM-medium complete		
	IMEM	
	Fetal bovine serum	10% (v/v)
	Penicillin G	100 U/mL
	Streptomycin	100 U/mL
	GlutaMAX	0.5x

IMEM-medium complete (spinner)		
	IMEM	
	Fetal bovine serum	10% (v/v)
	Penicillin G	100 U/mL
	Streptomycin	100 U/mL
	Pluronic F-86	1x
	GlutaMAX	0.5x

2.11.2 Buffers

Coomassie staining solution		
	Coomassie brilliant blue G250	100 µM
	HCl	0.13 mM

6x DNA loading buffer		
	glycerol	60% (v/v)
	Tris-HCl, pH 7.5	10 mM
	EDTA, pH 8.0	60 mM
	bromphenol blue	0.05% (w/v)
	xylene cyanol	0.05% (w/v)

10x G-75		
	Hepes-KOH, pH 7.9	200 mM
	NaCl	750 mM
	$MgCl_2$	15 mM

10x G-150	
Hepes-KOH, pH 7.9	200 mM
NaCl	1.5 M
$MgCl_2$	15 mM

5x HTRO buffer	
Hepes-KOH, pH 7.5	1 M
$MgCl_2$	160 mM
Spermidin	10 mM
DTT	200 mM

1x MOPS buffer	
MOPS	40 mM
NaOAc	10 mM
EDTA, pH 8.0	1 mM

5x native gel loading buffer	
Tris	90 mM
boric acid	90 mM
EDTA, pH 8.0	2.5 mM
glycerol	30% (v/v)
bromphenol blue	0.05% (w/v)

PBS	
NaCl	137 mM
KCl	2.7 mM
Na_2HPO_4	10 mM
KH_2PO_4	1.8 mM
$CaCl_2$	1 mM
$MgCl_2$	0.5 mM

2x PK buffer	
Tris-HCl, pH 7.5	200 mM
EDTA, pH 8.0	25 mM
SDS	2% (w/v)

2x RNA loading buffer		
	formamide	80%(v/v)
	EDTA, pH 8.0	1 mM
	bromphenol blue	0.05% (w/v)
	xylene cyanol	0.05% (w/v)

1x Röder C buffer		
	glycerol	25% (v/v)
	Hepes-KOH, pH 7.9	20 mM
	NaCl	420 mM
	$MgCl_2$	1.5 mM
	EDTA, pH 8.0	0.2 mM
	DTT	0.5 mM
	PMSF	0.5 mM

1x Röder D buffer		
	glycerol	10% (v/v)
	Hepes-KOH, pH 7.9	20 mM
	NaCl	100 mM
	$MgCl_2$	1.5 mM
	EDTA, pH 8.0	0.2 mM
	DTT	0.05 mM
	PMSF	0.5 mM

6x SDS-loading buffer		
	Tris-HCl, pH 6.8	25 mM
	glycine	192 mM
	SDS	1% (w/v)

4x separating gel buffer		
	Tris	1.5 M
	SDS	0.4% (w/v)
	adjust to pH 6.8	

4x stacking gel buffer		
	Tris	0.5 M
	SDS	0.4% (w/v)
	adjust to pH 6.8	

Stripping buffer		
	Tris-HCl, pH 6.7	62.5 mM
	β-mercaptoethanol	100 mM
	SDS	2% (w/v)

SLAB4		
	Tris	50 mM
	Glycine	380 mM
	SDS	3.5 mM

10x TBE		
	Tris	0.89 M
	boric acid	0.89 M
	EDTA, pH 8.0	25 mM

10x TBS		
	Tris	200 mM
	NaCl	1.5 M
	adjust to pH 7.5	

10x TBS-T		
	Tris	200 mM
	NaCl	1.5 M
	Tween-20	1% (v/v)
	adjust to pH 7.5	

Western-Blot transfer buffer		
	SLAB4	50% (v/v)
	MeOH	20% (v/v)

3 Methods

3.1 General Molecular Biology methods

3.1.1 Quantification of nucleic acids solutions

To quantify the concentration of nucleic acids in aqueous solutions spectrometry was used. The sample was applied to a NanoDrop® spectrophotometer and the extinction coefficient of the solution was measured. This was then compared to the extinction coefficient of the aqueous solution and OD at a wavelength of 260 nm could then be used to calculate the concentration of nucleic acid present in the solution. Due to variations in the optical properties the following equations were used to determine the concentration:

$1OD_{260\ nm}$= 50 µg/ml double-stranded DNA

$1OD_{260\ nm}$= 33 µg/ml single-stranded DNA

$1OD_{260\ nm}$= 40 µg/ml single-stranded RNA

The purity of the nucleotide solution could be evaluated by calculating the ratio of $OD_{260\ nm}/OD_{280\ nm}$. Values around 1.8 indicate pure DNA, while values close to 2.0 indicate pure RNA. Lower than expected ratios indicate contamination of the sample with phenol and make further treatment of the sample necessary (e.g. an additional phenol-chloroform extraction).

3.1.2 Phenol-Chloroform-Isoamyl extraction

Phenol-Chloroform-Isoamyl alcohol extraction (PCI) was used to isolate nucleic acids from samples containing proteins. During PCI, if the sample volume was smaller than 200 µl water was added to obtain a final sample volume of 200 µl, then an equal amount of phenol-chloroform-isoamyl alcohol (25:24:1) was added to the sample and mixed by vigorous shaking for 5 to 10 min at room-temperature (RT). After shaking the organic and aqueous phase was separated by centrifugation (13 krpm, 5 min, RT) in a tabletop centrifuge. After centrifugation, the aqueous upper phase is carefully removed, while avoiding to come into contact with the organic or interphase. The aqueous phase is then mixed with 0.1 volume of NaOAc (pH 5.2) and 2.5 volumes of 100% EtOH to precipitate the nucleic acids present in it. For precipitation the samples are stored at -20°C for 1 h or overnight and then

centrifuged at 4°C, max. rpm for 1 h in a table top centrifuge. The resulting pellet is then carefully preserved, while the supernatant is aspirated, and washed once with 70% (v/v) ethanol. After washing the pellet is dried and resuspended in the desired solution.

3.1.3 Proteinase K digestion

In order to recover RNA from samples containing high amounts of proteins, or to release RNA bound in stable protein complexes, a proteinase K (PK) digest is performed. To digest samples with proteinase K, the sample is incubated in 1x PK-buffer containing 0.2 µg proteinase K per µl of sample for 45 min to 1 h at 45°C. After PK digest the RNA can be isolated using the aforementioned PCI protocol followed by resuspension in water and storage at -20°C.

3.1.4 Protein precipitation

To precipitate proteins from aqueous solutions two different approaches were used. For acetone precipitations the protein solution was combined with 5 volumes of ice cold acetone, mixed and then kept at -20°C overnight. Afterwards the proteins were precipitated from the solution by centrifugation (30 min, 4°C, 21000 rcf) and the supernatant was carefully aspirated. The pellet was then washed once with 80% (v/v) ethanol, and after removal of the supernatant, air dried for 10 min. The protein pellet could then be resuspended in protein sample buffer. Ethanol precipitation was used as an alternative to the precipitation of protein with acetone. For this approach, the protein containing solution was combined with 2.5 volumes of ethanol and 0.1 volume of sodium acetate (pH 5.2). After mixing, the sample was kept at -20°C overnight and the proteins were precipitated by centrifugation (30 min, 4°C, 21000 rcf). The pellet was again washed with 80% (v/v) ethanol and dried prior to resuspension in sample buffer. For both precipitation approaches, a small amount (1 µl) of GlycoBlue can be added to the sample prior to precipitation in order to increase the visibility of the pellet after precipitation.

3.1.5 SDS-PAGE electrophoresis

Proteins were resolved by SDS-PAGE electrophoresis using commercial 4-12% Bis-Tris gels (lifeTechnologies). The samples were mixed with commercial 4x LDS-loading dye and incubated at 70°C for 10 min prior to loading. The gels were run in

1x MOPS buffer at 200 V for 55 min. Afterwards the proteins were visualized by Coomassie staining.

3.1.6 Coomassie staining of proteins

To stain SDS-PAGE gels with Coomassie blue, the gel was first fixated for 30 min in 40% (v/v) methanol; 10% (v/v) acetic acid and then washed twice for 15 min with ddH_2O. To stain the proteins, heated sensitive Coomassie blue solution was added to the gel and kept overnight at RT. To destain the gel, it was washed multiple times with ddH_2O, until the background staining reached an acceptable level.

3.1.6.1 SPYRO Ruby staining of proteins

To stain proteins with SPYRO Ruby the proteins were first separated on an SDS-PAGE gel by gel electrophoresis. The gel was then fixated overnight using fixation solution [30% (v/v) ethanol, 10% (v/v) acetic acid] and after fixation the gel was equilibrated for staining by three incubations in 20% ethanol for 30 minutes each. In parallel, the dye stock solution was diluted in water to a final concentration of 1 µM and added to the gel. The gel was exposed to the dye for 6 hours, while being kept dark. After the staining was completed, the gel was equilibrated in water for 10 minutes twice and then destained for a minimum of 16 h using a destaining solution [40% (v/v) ethanol, 10% (v/v) acetic acid]. Once destaining was completed, the stained proteins could be visualized by fluorescence measurements using the SPYRO Ruby filter on the FLA-7000.

3.1.7 Western Blot

To transfer proteins from a SDS-PAGE gel onto a nitrocellulose membrane (GE healthcare Amersham Protran 0.2 µM nitrocellulose) a tank-blot procedure was used. For blotting, a sandwich of a foam pad, three layers of Whatman 3 mm paper, SDS gel, nitrocellulose membrane, three additional layers of Whatman 3 mm paper and another foam pad was assembled in a gel holder cassette of the BioRad Trans-Blot Cell system. The gel holder cassette was then submerged in blotting buffer. The transfer was facilitated by applying a constant voltage of 70 V for 2 h at 4°C. Once the transfer was completed the nitrocellulose membranes were removed and blocked by incubating in 5% (w/v) milk-TBST for 1 h at room temperature. Detection of bound proteins was then achieved by incubation with specific primary antibodies diluted in 5% (w/v) milk-TBST overnight at 4°C. After incubation with the primary

antibody the membrane was washed three times in TBST for 10 min at room temperature and to detect bound primary antibody the blot was exposed to secondary antibodies conjugated with horseradish peroxidase diluted 1:40000 in 5% (w/v) milk-TBST for 1 h at room temperature. The secondary antibody was removed by three 10 min washing steps of TBST, followed by detection with the ECL Western Blot Detection Kit (GE Healthcare) according to the manufacturer's instructions. Images were acquired by exposing to X-ray films or digitally with the Amersham imager 580 (GE).

3.1.7.1 Striping of Western Blots

If a western blot membrane was used for multiple different primary antibodies, it was necessary to remove any previously bound antibodies. For this, the membrane was incubated with stripping-buffer for 30 min at 50°C and then washed with TBST twice for 15 min at room temperature. The stripped membrane was then incubated with to primary antibody after blocking as described above.

3.1.8 Restriction digest of plasmid DNA

To digest plasmid DNA with restriction endonucleases plasmid DNA was incubated with the appropriate endonuclease/buffer combination. The exact reaction setup is described in the following table:

Table 3.1 – **General restriction digestion reaction**

Component	**Amount**
Plasmid DNA	1 µg
Restriction enzyme	10 U
10x buffer	1x buffer
H_2O	Ad 50 µl

The reaction was then incubated for 2 h at 37°C. After restriction digestion the DNA fragments can be separated by agarose gel electrophoresis and gel purified using a gel extraction kit.

3.1.9 SYBR gold staining of RNA

To stain RNA separated on SDS-PAGE gels, SYBRgold® (ThermoFischer) staining was performed. The SDS-PAGE gel was fixated with 40% MeOH, 10% HOAc for 30

min and then washed with water once for 15 min followed by two 15 min washing steps with 1x TBE. After the final 1x TBE washing the SYBRgold dye was diluted in 1x TBE 1:10000 and added to the gel. The gel was then protected from light and incubated with the staining solution for 1h. After the incubation period has passed, RNA can be visualized using the TAMRA-filter on a FLA-7000 machine. If desired the gel can now directly be stained with Coomassie blue to visualize proteins.

3.1.10 Agarose gel electrophoresis

To separate DNA fragments by size, agarose gel electrophoresis was performed. Depending on the desired resolution 0.5-2% (w/v) of agarose were dissolved in 1x TBE buffer, mixed with 1x SYBR Safe DNA stain, and used to cast a gel. The gel was then placed in an electrophoresis chamber and the sample, mixed with loading dye, were loaded into the wells. Gels were run for ~1 h at 150 V in 1x TBE buffer. After electrophoresis, the DNA was visualized using an Invitrogen Safe Imager Blue-Light Transilluminator.

3.1.11 Extraction of DNA from agarose gels

To extract DNA fragments from agarose gels the DNA fragments were separated by size using agarose gel electrophoresis and the fragments were visualized using the Invitrogen Safe Imager Blue-Light Transilluminator. The bands corresponding to the desired fragment size were cut out of the gel using a sterile scalpel and placed in Eppendorf tubes. The commercial PCR& Gel extraction kit of Macherey-Nagel was then used according to the manufacturer’s instructions to extract the DNA.

3.1.12 Ligation of DNA fragments

To ligate DNA fragments ligation with T4 DNA ligase were performed. Fragments were mixed with a molar ratio of 3:1 for sticky-end ligations and 7:1 for blunt end ligations. In total 100 ng of plasmid fragments were combined with 1x T4 DNA-ligase buffer and 1 µl of T4 DNA ligase per 10 µl of ligation reaction. The assembled reaction was then incubated at 16°C overnight. After ligation the reaction containing the ligated fragments could be used for *E.coli* transformation directly.

3.1.13 Generation of chemically competent *E.coli*

To generate chemically competent *E.coli* cells, 3 ml LB-medium were inoculated with the appropriate bacterial strain (for this thesis XL-10 gold) and incubated at

37°C overnight. The starter culture was then added to 500 ml SOB medium and grown to an OD_{600nm} between 0.4 and 0.6. After the desired OD was reached, the culture was placed on ice for 10 minutes and then the cells were pelleted by centrifugation (10 min, 5900 rpm, 4°C in F14-14 rotor). After centrifugation the pellet was resuspended in ice cold TB-buffer and incubated on ice for 10 minutes. Afterwards the cells were pelleted by centrifugation (10 min, 3400 rpm, 4°C in F14-14 rotor) again and then resuspended in 18.6 ml TB and 1.4 ml DMSO. The cell suspension is then incubated on ice for another 10 min. Then the competent cells were aliquoted in to 50 µl aliquots and frozen in liquid nitrogen. The competent cells can then be stored at -80°C for several years.

3.1.14 Transformation of chemically competent *E.coli*

To transform chemically competent *E.coli* the cells were thawed on ice and then 100 ng of plasmid DNA were added to 100 µl of competent bacterial cells. The DNA-bacteria mixture was incubated on ice for 30 min followed by a 45 second heat-shock at 42°C and a 2 minute regeneration period on ice. After the regeneration period, the transformed bacterial cells were combined with 2 ml of LB-medium without antibiotics and incubated in a shaker at 37°C for 1 hour prior to being plated onto LB-Agar plates containing the appropriate antibiotic selection marker. The plates were then placed in a 37°C incubator to grow the bacterial colonies. After single colonies formed, individual colonies were picked using a sterile toothpick and used to inoculate 2 ml cultures.

3.1.15 Purification of plasmid DNA from *E.coli*

To isolate plasmid DNA from *E.coli* culture two commercial kits were used. For small preparations (up to 2 ml culture volume), the Roche Plasmid Mini Kit was used according to the manufacturer's instructions. For large scale preparations with up to 500 ml culture volume the MaxiPrep Kit from lifeTechnologies was used according to the manual. The purified precipitated plasmid was resuspended in TE buffer.

3.2 Construct design

3.2.1 Design of the ZDHHC16-MS2 plasmid

The ZDHHC16 pre-mRNA transcription plasmids were designed by extracting the exon 9, intron 9 and exon 10 from the ZDHHC16 gene and placing them under

control of a T7 RNA polymerase promoter. In addition, three MS2-loops were added to the 3' end of the exon 10. The whole plasmid was then synthesized by GeneScript in a vector (pUC57) containing an ampicillin resistance.

3.2.2 Design of MINX-exon complex plasmid and its derivatives

The MINX exon plasmid was designed as a derivative of the original MINX-MS2 plasmid. To generate the MINX exon plasmid, the 5' exon of MINX and part of the adjacent intron were removed and replaced with a T7 promoter sequence. The remaining 3' side of the intron and the 3' exon were left unchanged. A 5'splice site was added to the 3' end of the exon followed by a 15 nt spacer in between the 5'SS and the MS2-loops. The derivatives used for the analysis of the cross exon complex assembly on genes sensitive to the SF3b1^{K700E} mutation were designed by replacing the region from 25 nt upstream of the cryptic branch point to the last most 3' nucleotide with the same region from the affected intron of the gene of interest. The plasmids were then synthesized externally, by GeneScript, and subsequently used for transcriptions.

3.2.3 Design of lentiviral plasmids and transient expression vectors

3.2.3.1 Generation of a tagged SF3b1 plasmid

To generate a tagged SF3b1 expression construct the cDNA sequence of SF3b1 was synthesized (GeneScript) and ligated into an empty vector. The N- and C-terminal tags for SF3b1 were then added by restriction digestion and ligation. For this dsDNA oligomers, containing the appropriate restriction sites, were restriction digested and ligated into the similarly prepared vector. The ligated vector was then transformed into competent *E.coli* and integration of the tags was verified by DNA sequencing.

3.2.3.2 Generation of lentiviral expression vectors

The vectors used for the lentiviral expression of tagged SF3b1 were created by ligating a SF3b1 expression construct into the multiple cloning site of the respective vector. For lentiviral experiments and transient expressions, the pWPI backbone was used to create an expression vector. The SF3b1 cDNA was excised from its plasmid using Pac I and Pme I restriction enzymes and the pWPI vector was digested with the same two restriction enzymes. The DNA from both reactions was

separated on an agarose gel and the correct fragment was isolated using agarose gel extraction. The resulting fragments were ligated using T4 DNA ligase and cloned into competent *E.coli*.

3.2.4 Sequencing of plasmid DNA

In order to obtain sequence information 1 µg of plasmid DNA was combined with 3 µl of 10 µM sequencing primer and sent to MicroSynth® for Sanger sequencing. The resulting sequence files were aligned using the ClustalOmega tool for multiple sequence alignments.

3.3 Cell culture

3.3.1 General cell maintenance

HeLa S3 cells (Computer Cell Culture Center, Belgium) were grown in suspension in S-MEM media supplemented with 5% (v/v) newborn calf serum, 50 µg/ml penicillin and 100 µg/ml streptomycin. Cultivation and harvesting of the cells was done as described in (Dignam et al., 1983).

Adherent HEK293T cells were grown in DMEM (Gibco) supplemented with 10% (v/v) fetal bovine serum (Gibco) and 100 U/ml of both penicillin and streptomycin. The cells were maintained in a non-confluent monolayer and split at a confluence of ~80%. To detach the cells from the cell culture flask/dish, the medium was aspirated and the cells were washed with 1xPBS, then Accutase (Gibco) was added and the cells were incubated for 5 min at room temperature. After incubation, the cells were carefully detached by carefully washing them off with culture medium. The cell suspension was then counted and could be used to seed new culture containers. The seeding density for new plates/flasks was kept at roughly 10^5 cells/cm².

K562 suspension cells were cultivated in IMEM medium supplemented with 10% (v/v) heat inactivated fetal bovine serum (Gibco) and 100 U/ml of both penicillin and streptomycin. During culture in flasks, the cell density was kept between $3x10^5$ and $1.5x10^6$ cells per ml. During cultivation in spinner flasks, the cells were constantly stirred at 40 rpm and 1x Pluronic F-86 (Gibco) was added to the medium in addition to the normal supplements used for culturing K562 cells in flasks. The cell density was maintained in the range of $3x10^5$ to $1.2x10^6$ cells per ml. To split K562 cells, the cell suspension was transferred into 50 ml Falcon tubes and centrifuged for 5 min,

500 rcf at room temperature. After centrifugation the supernatant was aspirated and the cell pellet was carefully resuspended in preheated medium. Then the cell density was measured and an appropriate amount of cells was added to the new culture. For cells growing in large (3l and 6l) spinner flasks the cell density was kept in an acceptable range by increasing the culture volume without replacing the culture medium completely. This was done to minimize the risk of contamination.

3.3.2 Cell counting

To evaluate the cell density an improved Neubauer chamber was used. The Neubauer chamber allows the counting of cells in a defined volume, therefore the cell density can be calculated, based on the observed cell count and the used dilution-factor (d).

$$cell\ density = d \times \frac{cell\ count}{n} \times 10^4$$

3.3.3 Transfection of HEK293T cells

Transient transfections of HEK293T cells were performed with polyethyleneimine (PEI). To transfect cells with PEI the plasmid DNA was mixed with the cationic compound to form particles that precipitate onto the cells and are taken up into the cell nucleus. To form these particles dissolved PEI was mixed with plasmid DNA. PEI and DNA were combined in a ratio of 3 µg of PEI with 1 µg of DNA in 25 µl of OPTI-MEM, commonly 1 µg DNA per 7.5 cm² of adherent cells were used. The PEI-DNA mixture was then vortexed briefly and kept at room temperature for 15 to 30 minutes. Once incubation is completed the PEI-DNA mixture is carefully added to the cells and kept on the cells for 24 h. After 24 h the medium containing the PEI-DNA particles was carefully aspirated and replaced with normal preheated culture medium. To further increase the transfection efficiency the ratio of PEI to DNA can be increased up to 7 to 1.

3.3.3.1 Preparation of PEI

The PEI solution was prepared by dissolving 25 kDa polyethyleneimine to a final concentration of 1 µg/µl in water. To dissolve the PEI in water it is necessary to decrease the pH-value of the solution to approximately 2, at this pH the PEI will go into solution. Once all PEI is dissolved the pH can be readjusted to with sodium

hydroxide. The PEI solution can then be stored at 4°C for one month or at -20°C for a minimum of four years.

3.3.4 Preparation of HeLa S10 nuclear extract

To prepare splicing active nuclear extract from HeLa S10 cells, the protocol described in Dignam et al., 1983 was used with some modifications. Six to eight liters of cells were grown to a density of 5x 10^5 cells/ml and harvested by centrifugation in a Heraeus Cryofuge 6000i (10 min at 2000 rpm). The supernatant was aspirated and the cell pellet was washed thrice with 1x PBS pH 7.4. The pellet was then carefully resuspended in 1.25 volumes of 1x MC buffer containing 2 tablets of cOmplete EDTA-free protease inhibitor (Roche) per 50 ml of buffer. After 5 min incubation on ice the cells were transferred into 100ml douncers and while being kept on ice dounced 18-times. The cell suspension was then transferred to falcon tubes and the nuclei were pelleted by centrifugation in an SS34 rotor (13000 rcf, 5 min and 4°C). After centrifugation the supernatant was removed and the nuclei pellet was resuspended in 1x RC buffer containing 0.5 mM DTE and 0.5 mM PMSF. To disrupt the nuclei they were again treated with 20 strokes of the dounce homogenizer while being kept on ice. The lysate was then stirred for 40 min at 4°C and centrifuged in a SS34 rotor for 30 min at 16000 rpm at 4°C to remove debris. The recovered supernatant could then be frozen in liquid nitrogen for long-term storage or dialyzed against 50 volumes of RD buffer. For dialysis, the supernatant was poured into pre-wetted SnakeSkin® dialysis tubing with a cut-off of 7000 Da and placed in RD for 5 h with an exchange of dialysis buffer after 2.5 h. Once the dialysis was completed, the resulting RD-nuclear extract was again centrifuged in the SS34 rotor (10 min, 16000 rpm, 4°C). After centrifugation the nuclear extract was aliquoted and frozen in liquid nitrogen prior to storage at -80°C.

3.3.5 Preparation of small scale nuclear extracts from K562 cells

To prepare small scale nuclear extracts from K562 cells the aforementioned large scale protocol was used with minor adjustments to accommodate the smaller pellet size. The main difference was using a smaller douncer. In addition, the final centrifugation step was carried out in 2 ml aliquots using a tabletop centrifuge (2 min, max. rpm and 4°C).

3.4 *In vitro* splicing reactions

3.4.1 Template generation for *in vitro* transcription

To obtain specific RNA molecules for biochemical experiments, *in vitro* transcriptions were used. For this purpose, transcriptions plasmids encoding the different pre-mRNAs were isolated from XL-10 gold chemically competent *E.coli* using the Invitrogen® PureLink Maxiprep Kit according to the manufacturer's instructions. Plasmid DNA was then resuspended in ddH_2O and the sequence integrity was validated by sequencing. To generate a template for run-off transcription the plasmid DNA was linearized by restriction digestion with XbaI endonuclease (NEB). For this 25 µg of plasmid DNA were combined with 100 U enzyme and the appropriate buffer, and the reaction was then incubated overnight at 37°C. To isolate the digested plasmid DNA agarose gel-electrophoresis was performed to purify preparative amounts of linearized plasmid. The restriction digestion reactions were combined with 6x loading dye (supplied by NEB) and run on a 1% (w/v) agarose gel at 150 V for 30 min.

The 1% agarose gel was prepared by melting an appropriate amount of agarose in 1xTBE buffer and adding 1xSYBRsafe DNA stain after a short cooling period. The finished mixture was then used to pour a gel. After gel-electrophoresis, the DNA was visualized using blue light (wavelength of 470 nm). The desired bands were carefully cut out of the gel using a sterile scalpel and the DNA was extracted using the Macherey-Nagel Gel and PCR purification kit according to the manufacturer's instructions.

3.4.2 *In vitro* transcription

RNA *in vitro* transcriptions were carried out using T7 RNA polymerase run-off transcription and Xba I-linearized DNA templates. Uniformly ^{32}P labeled RNA was obtained by performing RNA synthesis in the presence of α-^{32}P-UTP. The composition of a transcription reaction is shown in the following table:

Table 3.2 – **Different *in vitro* transcription conditions used for pre-mRNA transcription**

Component	Stock Concentration	hot capped RNA		cold capped RNA		hot uncapped RNA		cold uncapped RNA	
		Final concentration	Volume (µl)	Final concentration	Volume (µl)	Final concentration	Volume (µl)	Final concentration	Volume (µl)
HTRO buffer	5x	1x		1x		1x		1x	
ATP	100mM	7.5mM	1.5	7.5mM	1.5	7.5mM	1.5	7.5mM	1.5
CTP	100mM	7.5mM	1.5	7.5mM	1.5	7.5mM	1.5	7.5mM	1.5
GTP	100mM	1.5mM	0.3	7.5mM	1.5	7.5mM	0.3	7.5mM	1.5
UTP	100mM	1.5mM	0.3	7.5mM	1.5	1.5mM	0.3	7.5mM	1.5
$m^7G(5')ppp(5')G$-cap	174mM	5mM	0.6	5mM	0.6				
linear template	0.5µg/µl	50ng/µl	1	50ng/µl	1	50ng/µl	1	50ng/µl	1
Rnasin	40U/µl	1U/µl	0.5	1U/µl	0.5	1U/µl	0.5	1U/µl	0.5
YIPP	40U/µl	0.5U/µl	0.25	0.5U/µl	0.25	0.5U/µl	0.25	0.5U/µl	0.25
T7 RNA polymerase	20U/µl	1U/µl	1	1U/µl	1	1U/µl	1	1U/µl	1
α-^{32}P-UTP	10mCi/µl; 3000 Ci/mmol		5				5		
ddH2O			8.05		13.05		8.65		13.65
total			20		20		20		20

The transcription reaction was performed at 37°C for 2.5 h. After completion of transcription, the DNA-template was removed by adding RQ1 DNase (1 U/µl, Promega) treatment for an additional 30 min at 37°C. To remove excess unincorporated NTPs from the reaction the sample was either run on a denaturing PAGE gel electrophoresis or applied to Illustra G-50 columns (GE healthcare) according to the manufacturer's instructions. In case of purification via PAGE electrophoresis, the bands on the gel were visualized via autoradiography and then cut out of the gel with a sterile scalpel. To extract the RNA from the gel, excised gel pieces were submerged in gel extraction buffer [20 mM Tris, pH 7.5, 0.2 mM EDTA, pH 8.0, 150 mM NaCl, 0.5% SDS] and kept at 4°C overnight. Afterwards the RNA was isolated from the gel extraction buffer using PCI extraction.

For all RNAs the concentration was determined using the NanoDrop and in case of ^{32}P-radiolabelled RNA molecules, the specific activity was determined by Cherenkov counting.

3.4.3 *In vitro* splicing reactions

To splice pre-mRNA *in vitro* radiolabeled RNA is added to a splice reaction. The splice reaction combines nuclear extract with energy substrates and salt. The reaction is set-up as follows, with the nuclear extract providing the entire potassium chloride and HEPES buffer required for the reaction:

Table 3.3 - **Röder D splice reaction**

Component	**Final concentration**
RD Nuclear extract	40%
ATP	2 mM
Creatine phosphate	20 mM
Magnesium chloride	3 mM
RNA	10 nM

The splice reaction is then incubated at 30°C for the required time.

Splice reactions that are performed using Röder C nuclear extract were carried out with the same general procedure; however the initial reaction setup changes slightly:

Table 3.4- **Röder C splice reaction**

Component	**Final concentration**
RC Nuclear extract	20%
ATP	2 mM
Creatine phosphate	20 mM
Magnesium dichloride	3 mM
Potassium chloride	60 mM
HEPES-KOH (pH 7.9)	20 mM
RNA	10 nM

Additionally it is necessary that the RNA is pre-incubated with MS2-MBP in order to prevent premature RNA degradation.

For experiments carried out in the presence of Pladienolide B, the reaction mixture without RNA was pre-incubated with Pladienolide B on ice for 10 minutes prior to the addition of the RNA.

3.4.4 *In vitro* assembly of cross exon complexes

The assembly of exon complexes was performed with exon RNA and RD nuclear extract under the same conditions that were used to perform *in vitro* splicing reactions outlined above, except the RNA concentration was reduced to 1 nM.

3.4.5 Analysis of splicing complexes by native electrophoresis

To visualize formation of spliceosomal complexes electrophoretic methods can be employed. Due to significant differences in migration speed, spliceosomal complexes can be separated on low percentage agarose gels and the respective position of the complex can then be visualized by autoradiography.

To analyze complex formation on native gels, splice reactions were performed as described above and, after completion of the incubation time, the reaction was mixed with 0.15 µg/µl heparin and stored on ice to stop the reaction. The samples were then supplemented with an appropriate amount of native-gel loading-dye and loaded onto a 1-2% low melting agarose gel prepared with 0.5x TBE buffer. Electrophoresis was then carried out at 80 V for 5 h in 0.5x TBE buffer at 4°C. After electrophoresis, the gel was dried for 5 h at 60°C prior to visualization of complex formation via autoradiography or with a PhosphorImager.

3.4.6 Denaturing polyacrylamide gel-electrophoresis

Polyacrylamide gel-electrophoresis (PAGE) was used to separate RNA molecules by size. For the separation of the RNA-products of *in vitro* splicing denaturing PAGE with 8M urea was used. Depending on the size of the expected RNA-products the gels were prepared containing 6-15% (v/v) acrylamide solution (40:1), 1xTBE and 8M urea. To initialize polymerization of the gels 300 µl of 20% (w/v) ammonium persulfate solution and 60 µl tetramethylethylenediamine (TEMED) were added to 100 ml of gel solution. After polymerization the gel was prepared for the samples by washing the pockets with 1x TBE buffer and a 10 min pre-run, followed by another washing of the pockets. The RNA samples were dissolved in RNA-loading dye and briefly heated to 95°C for 5 min and chilled on ice prior to loading. The electrophoresis was performed at 16 W in 1x TBE. Afterwards the gel was fixed in 40% (v/v) methanol; 10% (v/v) acetic acid, dried for 1 h at 80°C and the radioactive RNA was visualized by autoradiography or with a PhosphorImager.

3.4.7 MS2-affinity selection of splicing complexes

To isolate spliceosomal complexes from *in vitro* splicing reactions an MS2-affinity purification approach was used, as described in (Bessonov et al., 2010). The pre-mRNA was incubated for 30 minutes on ice with a 5 to 10-fold molar excess of MS2-MBP fusion protein in 20 mM HEPES-KOH, pH 7.9 buffer. Then the RNA was added to a splicing reaction of the desired size. After the reaction was completed, it was loaded onto an amylose affinity column. For analytical experiments, the reaction was passed through the column by gravity flow, while preparative experiments used commercial HiTrap MBP columns. Gravity flow columns were prepared with Amylose Resin (NEB) by resuspending the beads and applying an appropriate

amount to a gravity flow column. The column was then washed with 10 bed-volumes of G-150 buffer. The splicing reaction was then added and passed through the column by gravity. After binding of the MS2-tagged complexes, the column was washed with 10 bed-volumes of G-150 buffer and then the bound complexes could be eluted using 10 mM maltose in G-150 buffer. The commercial HiTrap MBP columns were used according to the manufacturer's instructions and the bound complexes were eluted using the same G-150 buffer containing 10mM that was used for the gravity columns.

3.5 Other Complex and protein purifications

3.5.1 Purification of MS2-MBP

E.coli bacteria expressing the MS2-MBP-StrepII protein were grown in LB-medium at 37°C and harvested at an OD of 2. The cells were then harvested by centrifugation and the cell pellet was frozen in liquid nitrogen for storage. To isolate the MS2-MBP protein the pellet was thawed and resuspended in five volumes of buffer A [20 mM Tris-HCl, pH 7.4, 200 mM NaCl, 1 mM DTT, 1 mM $MgCl_2$, 1x cOmplete protease inhibitor] and the cells were disrupted by harsh sonication (70% power, 15 min). The disrupted cell slurry was then clarified by centrifugation (30000 rcf, 30 min at 4°C) and the supernatant was collected. The pooled supernatant was then applied to a Streptactin HiTrap 5 ml column to purify the MS2-MBP-StrepII protein. The columns were prepared for affinity purification according to the manufacturer's protocol and pre-equilibrated with 10 column-volumes of buffer A. The clarified cell lysate was then applied over several hours with a loading speed of 0.2 ml/ minute. After loading the column was washed with 10 column-volumes of buffer A (not containing protease inhibitor or $MgCl_2$). Afterwards the column was washed with 10 column volumes of amylose matrix wash buffer [20 mM HEPES-KOH, pH 7.9, 150 mM NaCl, 0.05% (v/v) NP-40]. The bound protein was then slowly (0.1ml/ min) eluted with heparin elution buffer (20 mM HEPES-KOH, pH 7.9, 150 mM NaCl, 15% glycerol (v/v) and 5 mM desthiobiotin). The elution fractions were quantified by a Bradford assay and the peak fractions were pooled. The pooled elution fractions were then concentrated using an equilibrated Amicron 10 MWCO column. After concentration the final protein concentration was measured on a NanoDrop (extinction coefficient: 83435 $m^{-1}cm^{-1}$ molecular weight of MS2-MBP-

StrepII: 55.4 kDa). The isolated protein was then aliquoted and frozen in liquid nitrogen. For long term storage the aliquots were kept at -80°C.

3.5.2 Anti-FLAG purification of 17S U2 particles

The anti-FLAG affinity column was prepared with Sigma M2 anti-FLAG beads. Prior to application of the pooled gradient fractions, the beads were washed twice with 5 bed-volumes of G-150 buffer containing 2% (w/v) sucrose (denoted washing buffer). After the initial washing step to remove the storage buffer potentially bound contaminants were removed by washing the column with 2.5 bed volumes of 0.1 M glycine (pH 3.5) followed 20 bed-volumes of washing buffer.

To purify 17S U2 particles via the N-terminal FLAG-tag of SF3b1, nuclear extract containing the tagged protein was prepared using the small scale nuclear extract protocol outlined above. The nuclear extract was then dialyzed against Röder D buffer and the resulting dialyzed extract was applied to SureSpin gradients. The gradients were prepared as linear 5-20% (w/w) sucrose gradients containing G-150 buffer. In order to accommodate a larger amount of extract per gradient, 3.3 ml of the nuclear extract were applied to each gradient. The gradients were then centrifuged for 18:15 h at 4°C and 29500 rpm. After centrifugation, the gradient was fractionated into 1.5 ml fractions and aliquots of each fraction were analyzed by PAGE, followed by SYBR gold staining to visualize the distribution of snRNA across the gradient. Fractions in the 17S region of the gradient containing the U2 snRNA were pooled and mixed with an equal volume of G-0 buffer, to reduce sucrose content and salt concentration. The diluted gradient fractions were then applied to the anti-FLAG affinity column and incubated on the column for 2h by rotating head-over-tail.

The column was washed with 5 bed-volumes of wash buffer twice. To elute the bound complexes, 0.5 bed-volumes of elution buffer (0.1 mg/ml 3xFLAG-peptide in wash-buffer) were added to the column and drained by gravity flow. The column was then capped and 1 bed-volume of elution buffer was added and incubated on the beads for 15 minutes. The elution buffer was drained by gravity flow and was collected in an Eppendorf tube. Afterwards four to nine additional elution fractions of 1 bed-volume were collected by gravity flow. The eluted 17S U2 snRNPs were

then analyzed for their snRNA and protein content, frozen in liquid nitrogen and stored at -80°C.

3.5.3 Crosslinking of 17S U2 snRNPs and crosslink identification

Affinity-purified 17S U2 snRNPs were cross-linked with 75 µM BS3 for 30 min at 20 °C. The reaction was stopped by adding Tris-HCl, pH 7.9 to 10 nM and crosslinked U2 was further purified further by centrifugation on 10-25% (w/v) glycerol gradients in G150 buffer. Crosslinked 17S U2 particles sedimenting in the 17S region of the gradient were pelleted by ultracentrifugation and then analyzed as described by (Leitner et al., 2014), with some modifications as follows: pelleted material was dissolved in 4 M urea- 50 mM ammonium bicarbonate, reduced with DTT, alkylated with iodoacetamide, and the mixture was diluted to an end concentration of 1 M urea before digesting with trypsin (1:20 w/w). Peptides were reverse-phase extracted, fractionated by gel filtration on a Superdex Peptide PC3.2/30 column (GE Healthcare) and analyzed on a Thermo Scientific Q Exactive HF mass spectrometer. Protein-protein crosslinks were identified by pLink 1.23 search engine (pfind.ict.ac.cn/software/pLink) and filtered at FDR 1% according to the recommendations of the developer (Yang et al., 2012).

3.5.4 Affinity purification of 17S U2 from HeLa nuclear extract

HeLa nuclear extract was prepared essentially as described by Bertram et al. (2017b), except that the nuclei were extracted with a modified buffer Roeder C buffer containing 250 mM KCl and 8.5 mM $MgCl_2$. After centrifugation to remove non-soluble material, the extract was fractionated on a 17-30% (w/v) sucrose gradient containing 1X GK5M buffer [20 mM Hepes-KOH, pH 7.9, 150 mM KCl, 5 mM $MgCl_2$, 0.2 mM EDTA, pH 8.0]. Peak fractions containing 17S U2 were pooled, loaded onto an anti-SF3B1 affinity column and bound 17S U2 snRNPs were eluted with 0.2 mg/ml SF3B1 peptide (EQYDPFAEHRPPKIAC) in 1X GK5M + 1.5% sucrose.

3.5.5 Anti-FLAG affinity purification of exon complexes

To purify exon complexes from the K562 nuclear extract the complexes were assembled on radiolabeled RNA as described in “*In vitro* exon complex assembly”. The assembled complex was first isolated from the nuclear extract by MS2-selection and then eluted using maltose containing elution buffer. The eluates were pooled and applied to an anti-FLAG affinity column and incubated for 30 minutes. The

column was then washed twice with 5 bed-volumes of wash buffer. To elute the bound complexes, 0.5 bed-volumes of elution buffer (0.1 mg/ml 3xFLAG-peptide in wash-buffer) were added to the column and drained by gravity flow. The column was then capped and 1 bed-volume of elution buffer was added and incubated on the beads for 15 minutes. The elution buffer was drained by gravity flow and was collected in an Eppendorf tube. The proteins were then precipitated and analyzed by mass spectrometry.

3.6 Mass spectrometry

3.6.1 Protein identification by MS

To identify protein by mass-spectrometry, the sample containing the proteins of interest were first separated by 1-D gel electrophoresis (see chapter 4.1.5). The gel was then cut into pieces and the proteins were digested in-gel with trypsin and extracted as described in (Shevchenko et al., 1996). The extracted peptides were identified in a liquid-chromatography coupled electrospray ionization quadrupole time of flight (Q-ToF) or Orbitrap mass spectrometer. Proteins were identified by searching fragment spectra against the UniProt reference database using Mascot as a search engine.

3.6.2 Proteome analysis by MS

To analyze the proteome of K562 cells tandem mass tag-labeling (TMT-labeling) was performed (Thompson et al., 2003). Equal amounts of K562 cells expressing wild-type SF3b1 and K700E SF3b1 were harvested by centrifugation (5 min, 500 rcf, RT) and washed with PBS twice. The washed pellet was then lyzed by adding 5 pellet volumes of lysis buffer (1% SDS in 100 mM TEAB). The lysate was then clarified using centrifugation (10 min, 16000 rcf, 4°C) and the supernatant was transferred into a new Eppendorf tube. The clarified lysate was then precipitated by adding six volumes of acetone and chilling the sample overnight at -20°C followed by centrifugation (8000 rcf, 10 min at 4°C). The protein pellet was resuspended in 100 mM TEAB. RapiGest (in 50 mM TEAB) and DTT were added to a final concentration of 0.1% (RapiGest) and 10mM (DTT) and incubated for 1 h at 55°C. After incubation, the sample was cooled to RT and iodoacetamide was added to a final concentration of 20 mM and incubated at RT (in the dark) for 1 h. Afterwards the proteins were precipitated with acetone. The pellet was resuspended in 100 µl

of 50 mM TEAB and 1x of the 50x trypsin stock solution was added. The sample was then kept at 37°C overnight to digest the proteins.

To prepare the TMT-labels 41 µl of anhydrous acetonitrile were added to 0.8 mg of labeling reagent and the reagent was dissolved for 5 min with occasional vortexing. The dissolved TMT-reagent (TMT 126 for SF3b1^{wt} and TMT 127 for SF3b1^{K700E}) was then added to 100µl of the trypsin digested protein sample and incubated for 1 h at RT. After the incubation, 8 µl 5% hydroxylamine were added to each reaction and incubated for 15 min to quench the reaction. At this stage, the sample could be stored at -80°C if desired. Prior to mass spectrometry the samples were cleaned with C18 spin tips. To clean the sample the C18 tip was pre-washed with 400 µl 100% acetonitrile twice, followed by one wash step with 50% acetonitrile, 0.1% formamide in water and three wash steps with 0.1% formamide in water. The sample was then passed over the column twice followed by two wash steps with 0.1% formamide in water. The bound peptides were then eluted in 200µl 0.1% formamide, 50% acetonitrile in water. The eluted peptides were dried in a vacuum centrifuge and stored at -80°C.

3.7 Electron microscopy

Structural analysis of isolated complexes was done by negative stain and cryo-EM. All EM experiments were done in collaboration with Dr. Norbert Rigo. For EM investigations complexes were purified using the outlined MS2-affinity purification approach. Due to the concentration requirements of cryo-EM preparation reactions were scaled up (up to 400 ml), so that a post-gradient yield of 10-20 pmol of spliceosomes in the peak fractions could be expected. The eluate of the MS2 selection was applied to linear 5-20% (w/v) sucrose gradients (TST41.14 rotor, 27500 rpm, 18 h and 4°C). Post centrifugation the gradient was fractionated by hand and the presence of radioactive pre-mRNA was detected by Cherenkov counting. Peak fractions corresponding to the desired A complex were then pooled and buffer-exchanged on Amicron 50 kDa MWCO columns. Afterwards the sample was used to prepare grids for negative stain or cryo-EM microscopy.

For negative stain, the sample was adsorbed to a thin carbon film for 10-30 min and then transferred to a staining solution [2% (w/v) uranyl formate)] and stained for 2 min. After staining, the grids were air-dried and could be used for image acquisition.

Images were acquired on a CM200 FEG electron microscope (Philips, Netherlands) equipped with a 4k x 4k charge coupled CCD camera at 88000 fold magnification. For each data set, images containing approximately 75000 particles were collected. The particles were then picked using the Warp software package and classified using Relion. Cryo-EM grids were prepared by adsorbing particles to a carbon film followed by removal of excess liquid with a VitroBot (ThermoScientific) prior to freezing in liquid ethane. The frozen grids were then stored under liquid nitrogen. For gold grids, the sample was applied directly to the grid, excess liquid was immediately blotted away, and the grid was plunge frozen in liquid ethane. The gold grids were then also stored in liquid nitrogen until image acquisition took place. Images were acquired using a Titan Krios (FEI) equipped with a K2 camera (Gatan) at a magnification of 110000x. Acquired images were then processed with Warp and picked particles were classified in Relion.

4 Results

4.1 Splicing inhibition by Pladienolide B

4.1.1 Stalling of splicing in the presence of Pladienolide B

Pladienolide B (Pladienolide B) leads to an inhibition of splicing both *in vitro* and *in vivo* (Effenberger et al., 2014a; Kotake et al., 2007). Separate studies have also shown that Pladienolide B can interact with the SF3b1 subunit of the SF3b complex (Yokoi et al., 2011). Here the compound occupies the SF3b1 pocket that normally binds the branch point adenosine and interferes with its normal recognition (Cretu et al., 2018; Cretu et al., 2016). *In vitro* studies indicated that spliceosome assembly is stalled by Pladienolide B at an early stage, at either the A or subsequent pre-B complex stage (Effenberger et al., 2014a). However, it is not clear how Pladienolide B inhibits the spliceosome assembly, in particular whether it alters the composition and/or structure of the spliceosome after binding.

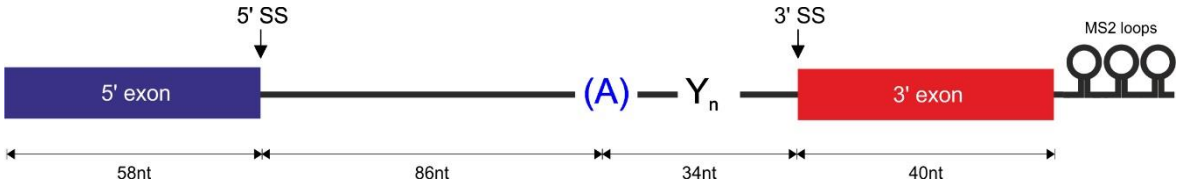


Figure 4.1- **Schematic representation of the MINX pre-mRNA transcript**. The pre-mRNA contains two exons separated by a 120nt intron. The 5' exon has a length of 58nt and the 3' exon is 40nt long. The branch point (A) is located 34 nucleotides upstream of the 3' splice site. In between the branch point and the 3' splice site a poly-pyrimidine tract (Y_n) can be found. At the 3' end of the transcript, there are three MS2-loops for affinity purification.

To investigate in more detail the inhibitory effect of Pladienolide B on pre-mRNA splicing *in vitro,* I transcribed ^{32}P-UTP labeled MINX pre-mRNA *in vitro* and purified it. This pre-mRNA contains two adenovirus derived exons and a single intron, in which a branch site and poly-pyrimidine (Y_n) tract suitable for splicing in human nuclear extracts can be found (Figure 5.1). After the 3' exon three stem-loops that bind the MS2 protein have been added to enable the purification of complexes assembled on the pre-mRNA. Incubation of this RNA under splicing conditions leads to the robust formation of a mature MINX mRNA. I then performed splicing essays in Pladienolide B treated HeLa S10 nuclear extract using 10 nM radiolabeled pre-

mRNA. The nuclear extract was pre-treated with Pladienolide B, in the absence of the pre-mRNA, by incubating the reaction on ice for 10 minutes.

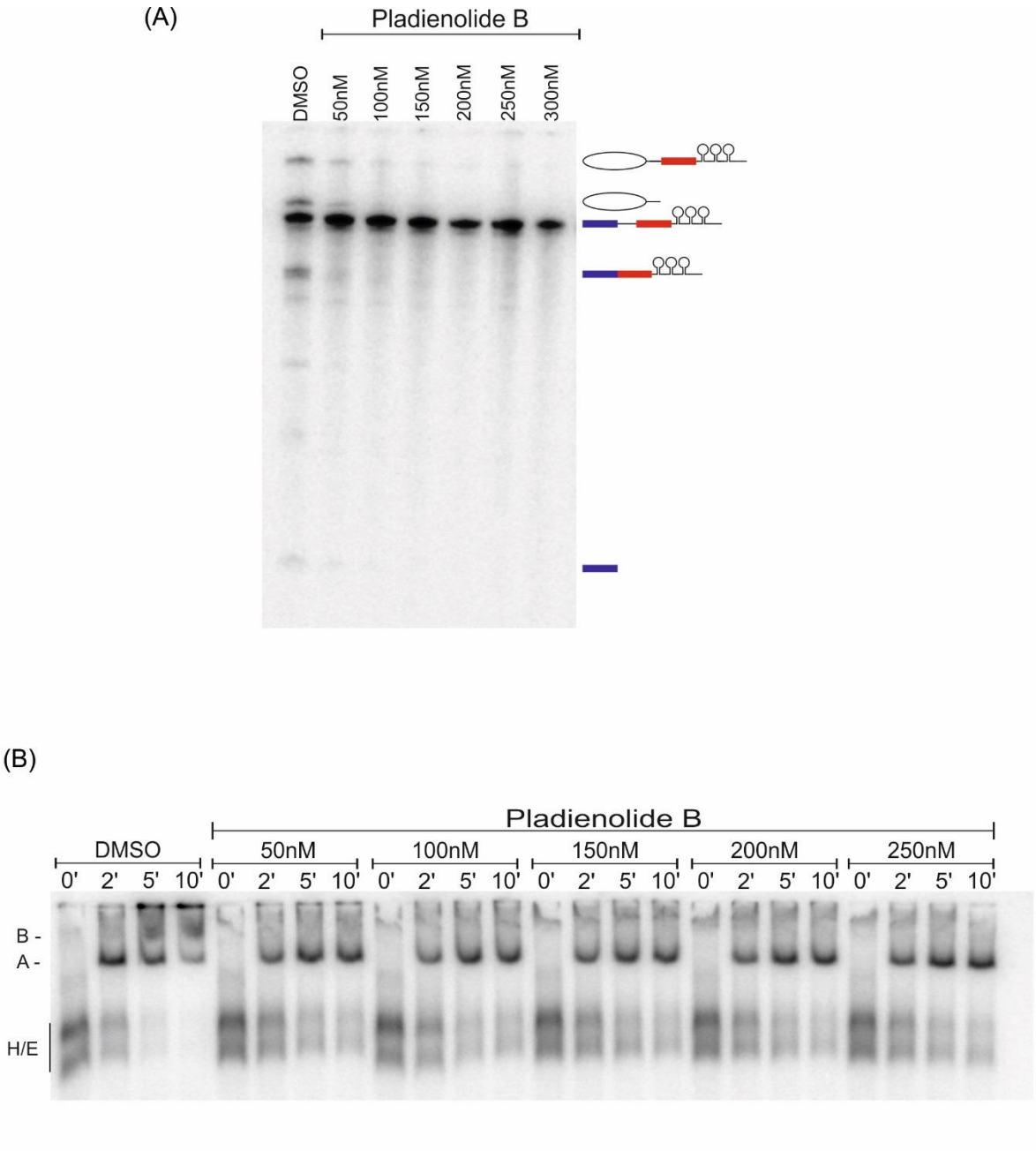


Figure 4.2 – **Effects of increasing concentrations of Pladienolide B on the splicing of ^{32}P-labled MINX-MS2 pre-mRNA in vitro**. Splicing was performed in the presence of various concentrations of Pladienolide B in HeLa nuclear extract for one hour. RNA was analyzed by denaturing PAGE (A). Spliceosomal complex formation was analyzed on a native agarose gel (B). The position of intron-lariat/3'exon, intron-lariat, pre-mRNA, mRNA and the positions of the spliceosomal complexes H/E, A and B are indicated next to the respective gel. Bands were visualized by autoradiography. DMSO, control reaction with the compound solvent.

In the presence of DMSO, the Pladienolide B solvent, the splicing products (mRNA and excised intron) and intermediate products (cleaved 5'exon and intron-lariat/3'exon) are clearly formed. The treatment of nuclear extract with increasing

concentrations of Pladienolide B shows that splicing is inhibited starting at a concentration of 50 nM (Figure 5.2). At 200 nM, both steps of splicing are completely inhibited. That is, not only the mature mRNA is no longer formed in the presence of Pladienolide B, but also no intermediate products of the splicing of MINX pre-mRNA cannot be detected (Figure 5.2a). This leads to the conclusion that the stalling of the splicing reaction takes place prior to the first step of splicing, consistent with previous studies (Effenberger et al., 2014a). In order to investigate the exact point in the splicing cycle at which Pladienolide B inhibits the progression of splicing, *in vitro* splicing reactions were incubated with increasing concentrations of Pladienolide B and analyzed by native gel electrophoresis, which allows analysis of the formation of spliceosomal complexes. In the presence of DMSO, A complexes were seen after two minutes and B complexes after 10 minutes (Figure 5.2b). With Pladienolide B, complex assembly was stalled at the stage of an A complex, already at 50 nM and further increases in the concentration did not impact the formation or the stability of the stalled complex.

To further investigate the exact nature of the stalled complex I performed less stringent gradient centrifugation experiments. The presence of heparin during the native gel electrophoresis disrupts weak interactions. For example, the tri-snRNP could potentially be loosely bound and thus a pre-B complex might form, but would be disrupted by the presence of heparin and not detected on the native gel. For this purpose a normal splicing reaction (3 minutes incubation time) and a reaction stalled by the addition of 250 nM Pladienolide B (10 minutes incubation time) were performed and subjected to linear 10-30% glycerol gradient centrifugation. The gradient profiles show that in the absence of Pladienolide B, complexes peak in fractions 14-17 and 20-24 (Figure 5.3a). Analysis of the RNA present in these fractions indicates they are A and pre-B/ B complexes, respectively. The complexes assembled in the presence of Pladienolide B on MINX pre-mRNA peak at the position of an A complex. The snRNA profile confirms that the peak observed in the gradient of the Pladienolide B stalled splicing reaction indeed corresponds to an A complex, indicated by the presence of both the U1 and the U2 snRNA, as well as MINX pre-mRNA (Figure 5.3c). No U4, U5 or U6 snRNA can be detected and the appearance of a second peak cannot be observed with the Pladienolide B treated sample. Thus the tri-snRNP is not associating with the A complex, even though the

incubation time is three times as long. We conclude that the presence of Pladienolide B in the nuclear extract interferes with the association of the tri-snRNP with the A complex, blocking assembly before the formation of the pre-B complex.

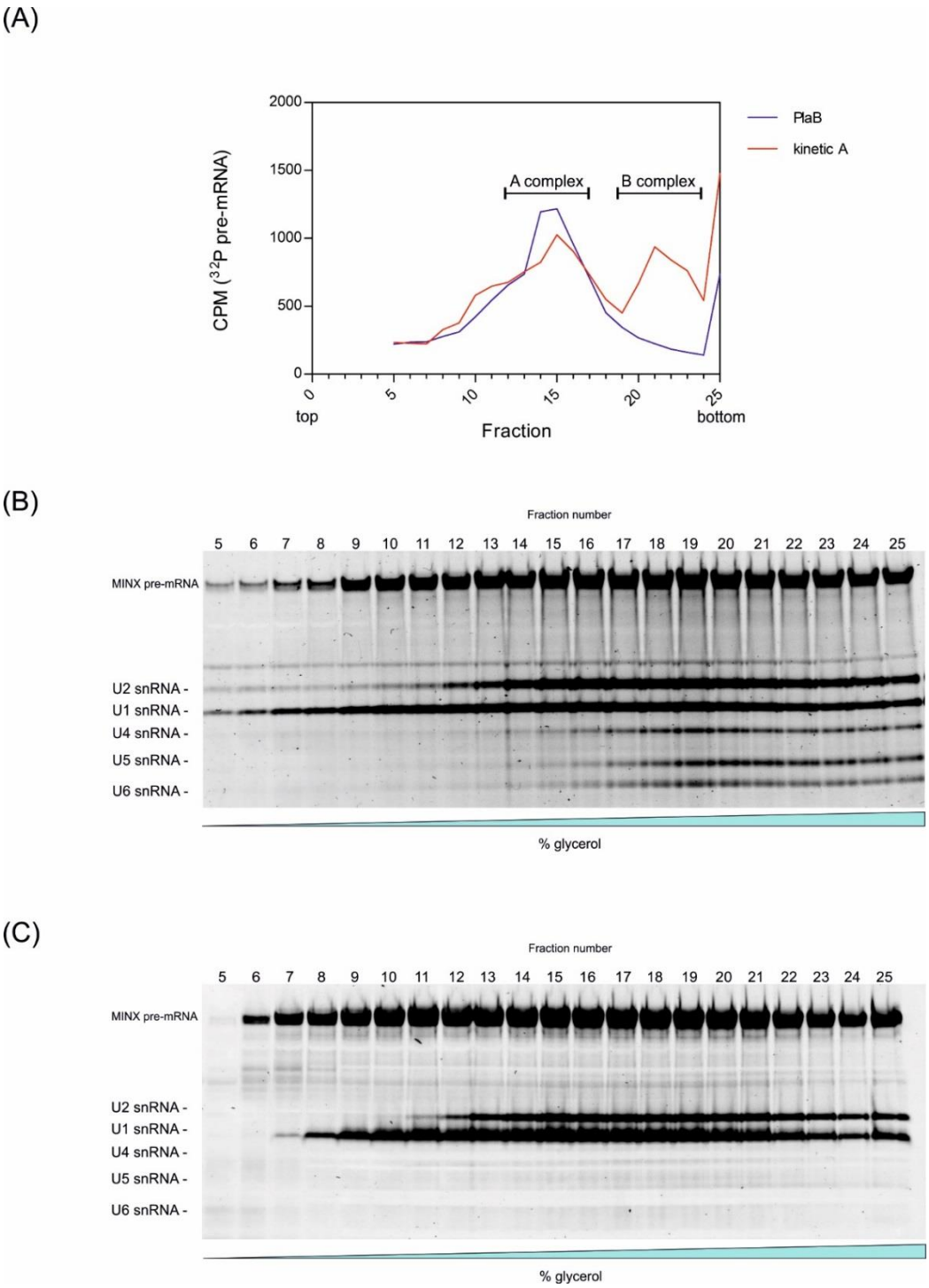


Figure 4.3 – **Glycerol sedimentation profiles of splicing complexes formed in the presence of 250nM Pladienolide B or DMSO.** In vitro splicing was performed in the presence of DMSO for 3 minutes or 250nM Pladienolide B for 10 minutes. The spliceosomal complexes formed on radiolabeled MINX pre-mRNA were

separated on a linear 10-30% (w/v) glycerol gradient containing 150mM sodium chloride (1x G-150 buffer), and fractionated from the top (A). The radioactivity of each gradient fraction is measured by Cherenkov counting and plotted. Peaks corresponding to the A and B spliceosomal complexes are indicated. Splicing extract plus DMSO (B) or Pladienolide B (C) were subjected to 10-30% (w/v) gradient centrifugation. RNAs were isolated from gradient fractions, separated by PAGE and visualized by SYBR gold staining. RNA identities are indicated at the left.

4.1.2 A complexes formed in the presence of Pladienolide B are unstable

The presence of Pladienolide B in the SF3b core was proposed to lock the HEAT-domain of the SF3b1 protein in an open conformation (Cretu et al., 2018). However cryo-EM structures of the human and yeast spliceosome show that during later stages of the splicing cycle SF3b1 adopts a “closed” conformation, tightly grabbing the U2 snRNA/branch point helix (Haselbach et al., 2018; Rauhut et al., 2016; Yan et al., 2016; Zhang et al., 2018). Since the presence of Pladienolide B potentially leads to the arrest of SF3b1 in an open conformation this could impact the stability of the assembled complex. In order to evaluate the stability of Pladienolide B stalled A complexes compared to kinetically assembled A complexes, I performed a heparin titration. Heparin dissociates protein complexes in a concentration dependent manner and therefore can be used to analyze complex stability.

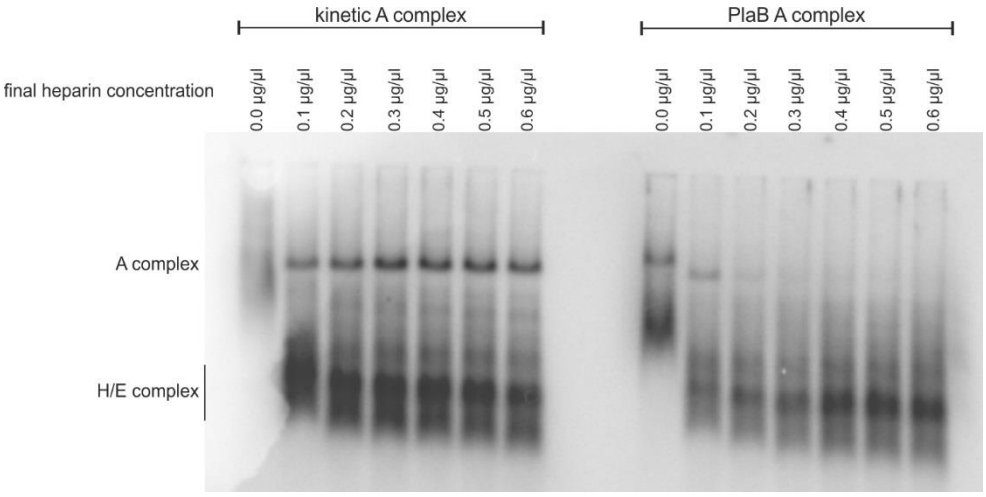


Figure 4.4 - **Stability of A complexes in the presence of increasing concentrations of heparin.** Splicing complexes were assembled in splicing extracts +/- 250 nM Pladienolide B and analyzed on native agarose gels after addition of increasing amounts of heparin. Bands were visulaized using a PhosphorImager and their identity is shown on the left. The comparison of the uninhibited and the Pladienolide B treated A complexes shows that the presence of Pladienolide B leads to a significant increase in the heparin sensitivity of the complex, starting at concentrations of 0.2 µg/µl.

The titration of heparin shows that the kinetic A complex is stable in the presence of heparin with a final concentration in excess of 0.6 µg/µl (Figure 5.4). The Pladienolide B containing A complex in contrast starts disassembling once the final

heparin concentration exceeds 0.1 µg/µl. Thus it can be concluded, that the A complex formed in the presence of Pladienolide B is less stable than the kinetic A complex assembled in HeLa nuclear extract. The failure to adopt a closed conformation of the HEAT-domain of SF3b1 could negatively impact the association of the U2 snRNP with the pre-mRNA. The presence of Pladienolide B could lead to changes in the binding of proteins needed for the stable A complex formation or could impair the structural rearrangement needed for a more stable A complex conformation. Taken together, these results suggest that while the presence of Pladienolide B does not impact the initial assembly it does have an influence on the A complex itself, since the complex is much less stable in the presence of heparin than its kinetic counterpart.

4.1.3 Affinity purification of Pladienolide B stalled A complexes

Stalling of spliceosome assembly by Pladienolide B can be the result of changes in the composition of the spliceosome or in its structure that prevent the tri-snRNP recruitment and stable integration. However, the exact consequences of the binding of Pladienolide B to SF3b1 on the structure and composition of spliceosomal complexes are currently unknown. I therefore affinity purified Pladienolide B stalled A complexes to check if their composition or structure was affected. For comparison I also purified kinetically stalled A complexes.

MS2-tagged MINX pre-mRNA was pre-incubated with an excess of MS2-MBP and then incubated under splicing conditions either without Pladienolide B or after adding 250 nM Pladienolide B (Figure 5.5a). The splicing reaction was loaded onto amylose beads and bound complexes were eluted with an excess of maltose (5 mM maltose). To separate the A complex from other complexes, the eluted complexes were separated by centrifugation on a 5-20% (w/v) sucrose gradient. Complexes in the A complex peak (Figure 5.5b) which contained only the MINX pre-mRNA plus U1 and U2 (Figure 5.5c), were pooled and used for subsequent analysis.

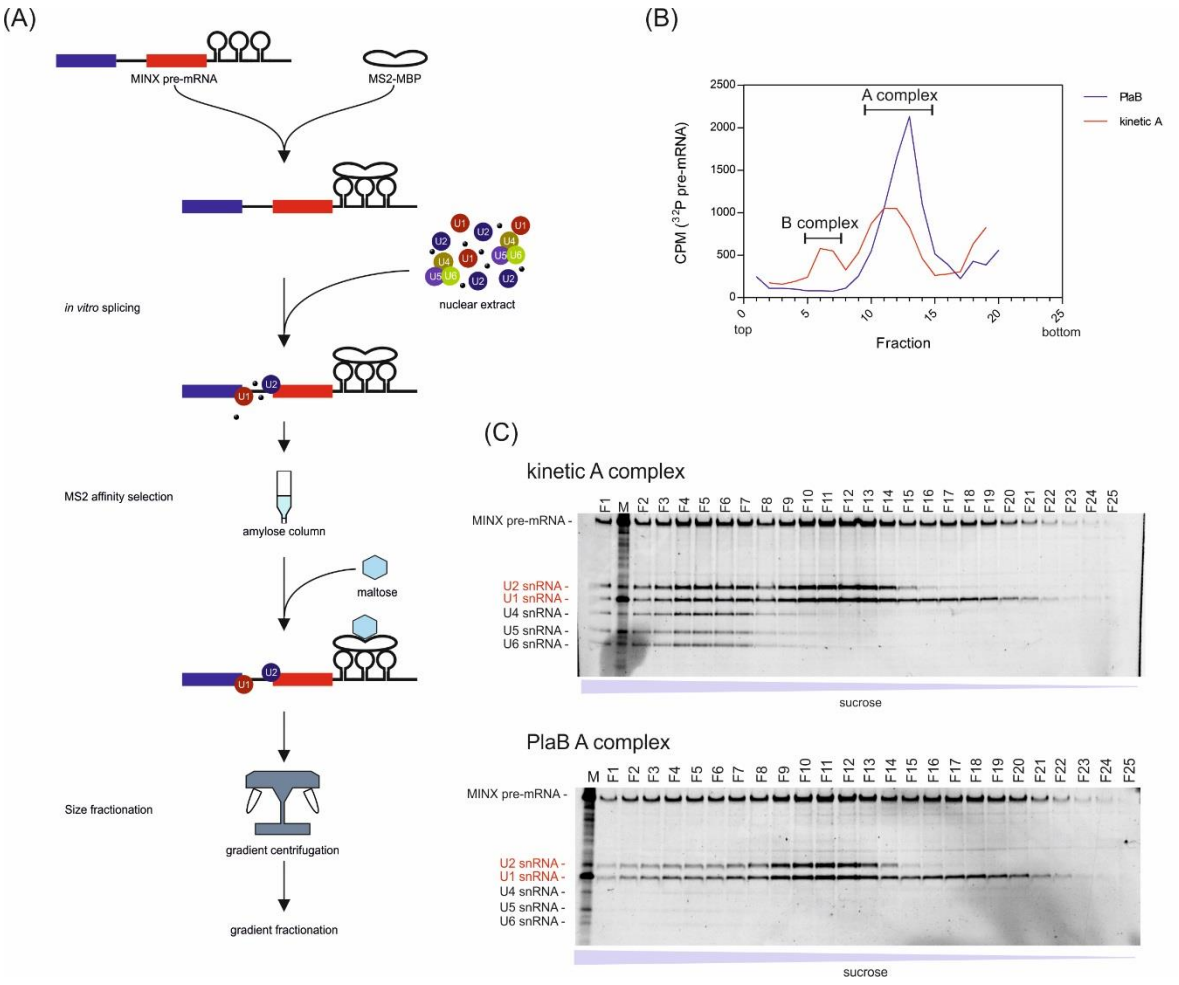


Figure 4.5 – **Purification of human A complexes under kinetically stalled conditions or stalled with Pladienolide B.** (A) Schematic of purification approach used for complex isolation from large *in vitro* splicing reactions. MS2-tagged MINX pre-mRNA substrate was pre-incubated with MS2-MBP fusion protein and then used for *in vitro* splicing. After splicing, the complexes were affinity-selected using amylose beads and the eluted complexes were separated by gradient centrifugation (5-20% (w/v) sucrose). The distribution of splicing complexes was detected by Cherenkov counting; peak fractions were pooled and used for further analysis. (B) Gradient sedimentation profile after bottom to top fractionation of kinetic and Pladienolide B stalled splicing reactions. (C) snRNA distribution across the gradients shown in B. RNA was analyzed by PAGE, visualized by SYBR gold and their identities are shown on the left.

4.1.4 Chase experiments with Pladienolide B stalled complexes

To check if Pladienolide B remains present within the spliceosome during the formation of spliceosomal complexes and their subsequent purification a chase experiment was performed. In a chase experiment spliceosomal complexes are assembled on radiolabeled MINX pre-mRNA in nuclear extract treated with Pladienolide B and then the A complexes are isolated using MS2-affinity selection. The isolated complexes can be added to nuclear extract to start the splicing process from a defined starting point. Initial experiments showed that stalled A complexes added to normal nuclear extract are converted into spliceosomes that can splice

pre-mRNA (data not shown). However, U2 particles present in the nuclear extract could replace the U2 particles containing SF3B bound to Pladienolide B if the stalled A complexes dissociate. Thus with this kind of experiment it is impossible to distinguish if splicing occurs because Pladienolide B is no longer bound to the complex or if there is an exchange of U2 particles in the A complex that re-enables splicing.

4.1.4.1 Depletion of SF3b from nuclear extract

For chase experiments involving the A complex the reassembly of the A complex has to be prevented, in order to determine if Pladienolide B is still present and inhibiting splicing. Since our department had an established protocol for the depletion of SF3b, a core component of the 17S U2 particle, I chose this approach to create a nuclear extract that cannot reform A complexes. Antibodies targeting an N-terminal epitope in SF3b1 were used to deplete HeLa S10 nuclear extract of the SF3b complex. This creates a nuclear extract without functional 17S U2 particles and it is therefore splicing deficient. To test for sufficient SF3B depletion the depleted extract was tested in splicing assays and the abundance of SF3b1 protein, as a marker for the abundance of SF3b complex, was analyzed by western-blots. To control for non-specific effects, mock-depleted extract (treated in the same way but with no antibody) was also prepared.

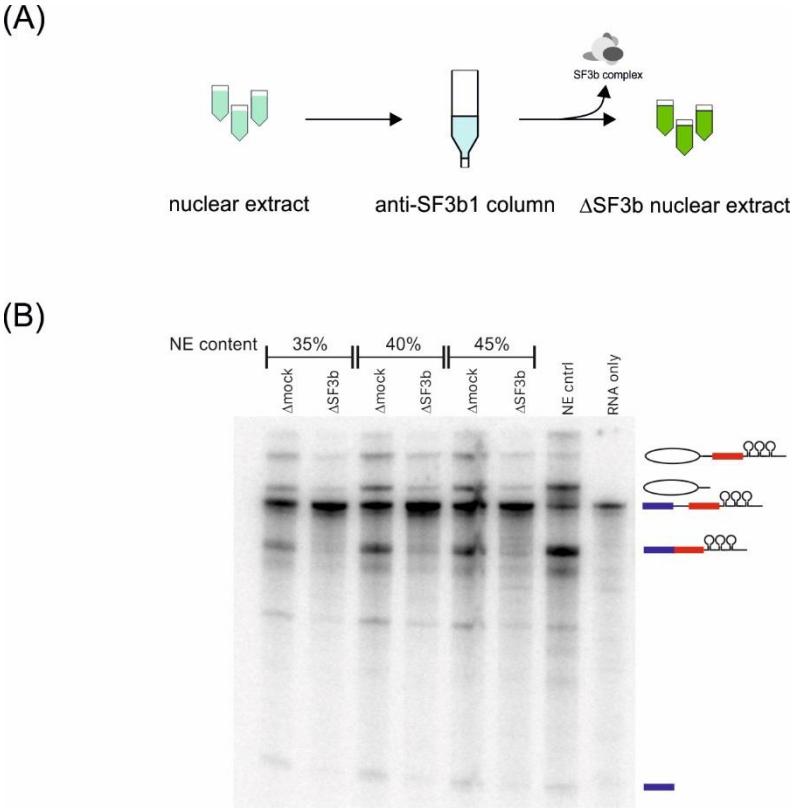


Figure 4.6 – **Immunodepletion of SF3b from HeLa nuclear extract.** (A) Scheme of the SF3b depletion process. (B) In vitro splicing of ^{32}P-MINX pre-mRNA using different concentrations of mock versus SF3b1 depleted extract. RNA was analyzed by denaturing PAGE and visualized with a PhosphorImager. The position of the splicing intermediates and products plus unspliced pre-mRNA is shown on the right. (NE cntrl = untreated nuclear extract).

Western blot analysis showed that greater than 95% of SF3b1 was removed from the nuclear extract (data not shown). The analysis of MINX pre-mRNA splicing confirmed that the depletion of SF3b inhibits splicing. A clear reduction in mRNA and spliced-out intron, and in the splicing intermediates, was seen in ΔSF3b extract compared to the mock-depleted extract at all extract concentrations tested. The mock-depletion compared to untreated extract shows that the experimental procedure itself causes only moderate reduction in the splicing efficiency of MINX pre-mRNA.

4.1.4.2 Complementation experiments

To investigate if Pladienolide B remains bound to the SF3b1 protein during A complex purification, I affinity purified A complexes stalled by Pladienolide B, or for comparison stalled with Anacardic Acid (AA), which disassembles tri-snRNPs but does not prevent A complexes from being chased into catalytically active spliceosomes (Kuhn et al., 2009).

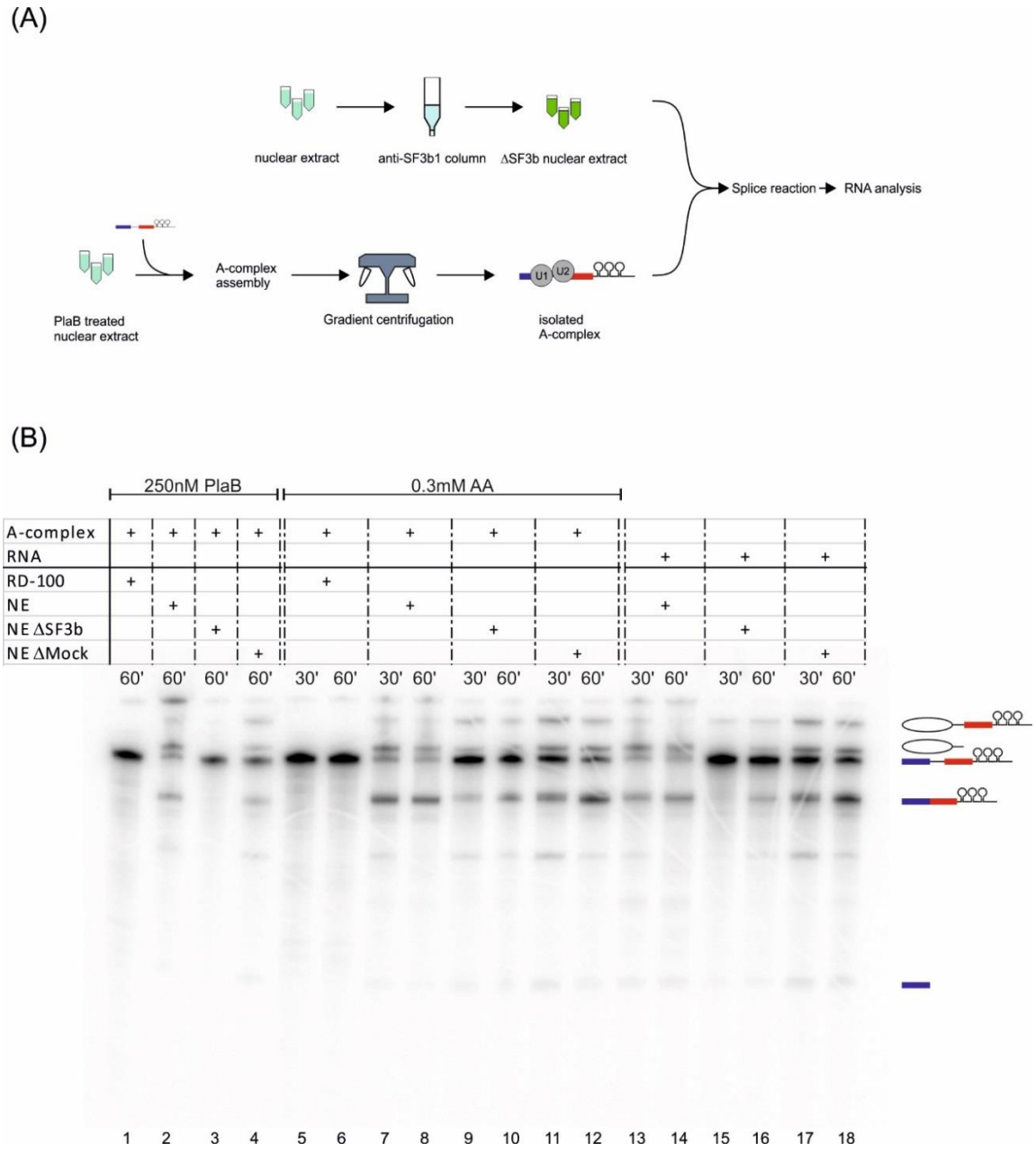


Figure 4.7- **Pladienolide B stalled A complexes are not chased into active spliceosomes.** (A) Scheme of the Pladienolide B chase experiment. (B) In vitro splicing of ^{32}P-labled MINX pre-mRNA in untreated nuclear extract (NE) or mock-depleted or ΔSF3b nuclear extract after addition of purified A complexes stalled with 250 nM Pladienolide B (Pladienolide B) or 0.3 mM anacardic acid (AA) as indicated. MINX pre-mRNA (RNA) was also added as a negative control. RNA was analyzed by denaturing PAGE and visualized with a PhosphorImager.

The gel in figure 5.7b shows the result of the chase experiment. Pladienolide B stalled A complexes can only be chased into splicing active spliceosomes, if they are incubated with nuclear extract containing intact 17S U2 particles, as is the case when the A complex is added to untreated or mock-depleted nuclear extract (lanes 2 and 4). If the isolated A complex is added to buffer (lane 1) splicing fails due to the absence of spliceosomal components in general. However, if the A complex is

added to nuclear extract that has been depleted of SF3b, and therefore of functional U2 particles, splicing does not take place (lane 3). In contrast, A complexes stalled by addition of AA, which only disrupts the tri-snRNP, can be chased into active complexes not only in untreated and mock treated extract, but also in the ΔSF3b extract (lanes 7-12). The second control, addition of RNA to the nuclear extracts, confirms the previous result that SF3b depleted nuclear extract is not capable of efficiently splicing MINX pre-mRNA, while the normal and the mock depleted nuclear extract are, again confirming the necessity of functional U2 snRNPs for splicing of pre-mRNA. These results indicate that the Pladienolide B compound likely remains present within the complex during the purification of the A complex. It also indicates that the presence of Pladienolide B within the A complex is sufficient to prevent the progression of pre-mRNA splicing.

4.1.5 Characterization of Pladienolide B stalled A complex

4.1.5.1 Protein composition of the Pladienolide B stalled A-complex

I first determined whether the presence of Pladienolide B causes changes in the protein composition of the stalled A complex. For this purpose affinity-purified A complexes, either kinetically stalled or stalled by Pladienolide B, were investigated by LC-MS/MS mass-spectrometry. Mass spectrometry was performed in collaboration with the Bioanalytical Mass spectrometry group at our institute.

Table 4.1- **Proteins identified by LC-MS/MS in kinetic and Pladienolide B stalled human A complexes.** Total peptide counts for each individual protein is shown for two independent preparations. The proteins are gouped according to function or association within the spliceosome. Common contaminants such as ribosomal proteins are omitted.

	kDa	human gene name	kinetic A #1	kinetic A #2	PlaB #1	PlaB #2
Sm proteins						
B/B'	24,6	SNRPB	7	8	11	9
D1	13,3	SNRPD1	3	7	6	6
D2	13,5	SNRPD2	7	13	14	8
D3	13,9	SNRPD3	14	24	23	19
E	10,8	SNRPE				
F	9,7	SNRPF				
G	8,5	SNRPG				
U1 snRNP						
U1-70K	51,6	SNRNP70	4	8	8	14
U1-A	31,3	SNRPA			2	2
U1-C	17,4	SNRPC				
U1 snRNP associated						
FBP11	108,8	PRPF40A				
S164 (fSAP94)	100,2	RBM25				
17S U2 snRNP						
U2A'	28,4	SNRPA1	16	28	27	26
U2B"	25,5	SNRPB2	2	3	1	2
SF3a120	88,9	SF3A1	72	106	98	122
SF3a66	49,3	SF3A2	9	15	13	17
SF3a60	58,9	SF3A3	59	60	73	117
SF3b155	145,8	SF3B1	259	224	284	364
SF3b145	100,2	SF3B2	88	99	123	158
SF3b130	135,6	SF3B3	292	283	314	426
SF3b49	44,4	SF3B4	12	15	14	16
SF3b14a/p14	14,6	SF3B6	6	13	10	16
SF3b14b	12,4	PHF5A	5	4		
SF3b10	10,1	SF3B5				
17S U2 related						
tat SF1	85,9	HTATSF1				
hPRP43	90,9	DHX15	6	19	2	21
SPF45	45,0	RBM17				
SPF30	26,7	SMNDC1	10	11		
U2AF65	53,5	U2AF2	2	3	7	12
U2AF35	27,9	U2AF1	1	1	1	2
SPF31	29,8	DNAJC8			1	0
hPRP5 (DDX46)	117,4	Prp5				
SR140	118,3	U2SURP	0	2		
CHERP	103,7	CHERP				
PUF60	59,9	PUF60				
SF3b125	103	DDX42			19	49
A proteins						
RBM10	103,5	RBM10	5	5	4	9
RBM5/LUCA15	92,2	RBM5	16	23	13	26

	kDa	human gene name	kinetic A #1	kinetic A #2	PlaB #1	PlaB #2
U5						
220K / hPrp8	273,6	PRPF8	32	39		
200K / hBrr2	244,5	SNRNP200	85	69	2	0
116K / hSnu114	109,4	EFTUD2	24	31	0	2
102K / hPrp6	106,9	PRPF6	8	14		
100K / hPrp28	95,6	DDX23	12	17		
A/B proteins						
RBM39 (RNPC2, CAPER)	58,5	RBM39	2	2	15	37
cap binding complex						
CBP20	18	NCBP2				
CBP80	91,8	NCBP1			14	35
hnRNP						
hnRNP H1	49,2	HNRNPH1	4			
hnRNP K	51	HNRNPK	2			
hnRNP L	64,1	HNRNPL	2			
hnRNP M	77,5	HNRNPM	9	0		
hnRNP U	90,5	HNRNPU	3	0		

In both A complexes all proteins commonly found at the A complex stage of spliceosome assembly could be identified by mass spectrometry, such as the U1 and U2 snRNP associated proteins, as well as various U2-related and A complex associated proteins (Table 5.1). In the case of the kinetic A complex, some proteins associated with the U5 snRNP could also be found. These proteins are observed in the spliceosome at later stages of splicing, after the recruitment of the tri-snRNP complex. The comparison of the kinetic human A complex and the Pladienolide B stalled human A complex reveals a similar protein composition for both complexes with only minor changes. Two very interesting differences between the kinetic A complex and the Pladienolide B stalled A complex are seen. The kinetic A complex lacks DDX42 compared to the Pladienolide B treated complex. DDX42 is a RNA-helicase that normally dissociates upon formation of the mature 17S U2 particle (Will and Lührmann, 2002). Another protein that appears to be absent in the Pladienolide B treated A complex, is SMNDC1 (SPF30). SMNDC1 is a factor required for the

docking of the tri-snRNP to the spliceosome (Meister et al., 2001; Rappsilber et al., 2001) and thus its absence could be significant for the lack of tri-snRNP binding seen after Pladienolide B treatment (see discussion).

4.1.6 Electron microscopic investigations of the Pladienolide B stalled A-complex

To determine if Pladienolide B affects the structure of the A complex, I set out to obtain high resolution structures of the spliceosomal A complex using cryo-electron microscopy, in collaboration with Dr. Norbert Rigo in our department. For this purpose, I first evaluated different purification conditions for both kinetic and Pladienolide B stalled A complexes to optimize yield and stability of the complexes, especially since the Pladienolide B stalled complexes are less stable. The complexes were assembled in HeLa S10 nuclear extract, and for the stalled complex in the presence of 250 nM Pladienolide B, and purified by MS2-selection. They were then subjected to gradient centrifugation and fixated using in-gradient glutaraldehyde crosslinking (GraFix) (Kastner et al., 2008). After gradient centrifugation, the A-complex peak was isolated and buffer-exchanged to reduce the sucrose concentration below 0.08% and then analyzed by negative stain EM and cyro-EM grids were prepared as well.

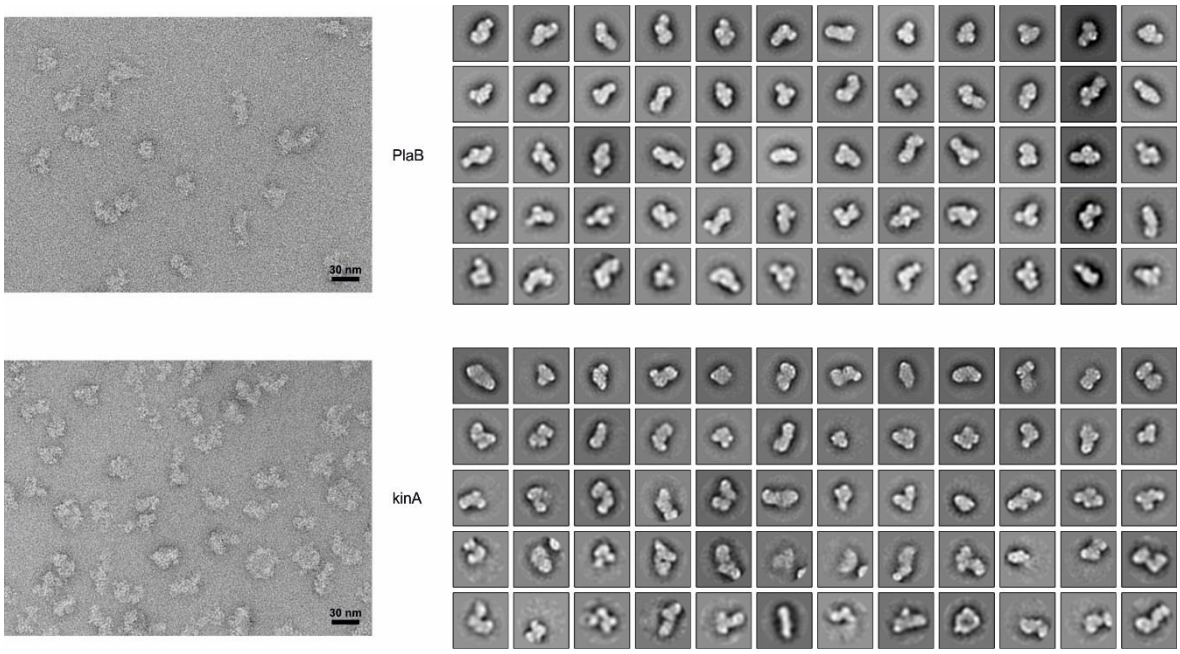


Figure 4.8 – **Electron microscopy of the Pladienolide B stalled human A complex and of the kinetic human A complex**. Complexes were formed on MINX-MS2 pre-mRNA in HeLa nuclear extract by MS2 affinity selection and then separated on a sucrose gradient (5-20% (w/v)) and cross-linked by GraFix. The peak fractions were pooled, buffer-exchanged and adsorbed onto a carbon film support for negative staining. Images were acquired in a CM200 microscope at 63000x magnification (A) Overview of raw images of the kinetic or Pladienolide B stalled A complex. (B) 2D class averages of the negative-stained complexes. Scale bar is 30nm.

Negative stain experiments with isolated human A complexes showed that intact particles could be isolated under the chosen conditions and that 2D classes could be calculated for both, the Pladienolide B stalled and the kinetic A complex (Figure 5.8). The A complex images have a maximum dimension of ca. 25nm and appear similar to previous A complex images from our lab (Behzadnia et al., 2007). The classes look similar for both A complexes, with elongated and more compact classes appearing in a roughly 1:1 ratio. Overall the presence of Pladienolide B does not cause major structural changes in the A complex, at least not any that would be visible in negative stain EM. The calculation of initial 3D models failed for both obtained datasets; however the manual selection of only compact classes from the Pladienolide B stalled data set led to an initial 3D model of the particle. The prepared cryo-grids however, did not allow us to calculate initial 2D classes. This was independent of the quenching method and time used or if the gold or carbon grids were prepared. Further investigations into the potential reasons for this behavior will be conducted in the future.

4.1.7 Pladienolide B and the cross-exon complex

In previous work, it was shown that when introns are very long the spliceosome first assembles across an exon. In this case, U2 binds to the branch site and U1 to the downstream 5’ splice site (Figure 5.9a), forming an A-like complex. The U5.U4/U6 tri-snRNP can also interact loosely with the A-like complex to form a 37S exon complex (Schneider et al., 2010b). This complex can be converted into a stable B-like complex by adding an oligonucleotide with a 5’SS consensus sequence (Schneider et al., 2010b). This dissociates the U1 snRNP and binds to the U6 snRNA. During these rearrangements the tri-snRNP is stably integrated into the complex. As it is not known if Pladienolide B inhibits the assembly of the cross-exon complex, I next investigated the effect of Pladienolide B on the assembly of these complexes formed on MINX exon RNA in HeLa nuclear extract. The MINX exon RNA contains an exon and part of the preceding intron with the BPS plus a downstream 5’SS (Figure 5.9b).

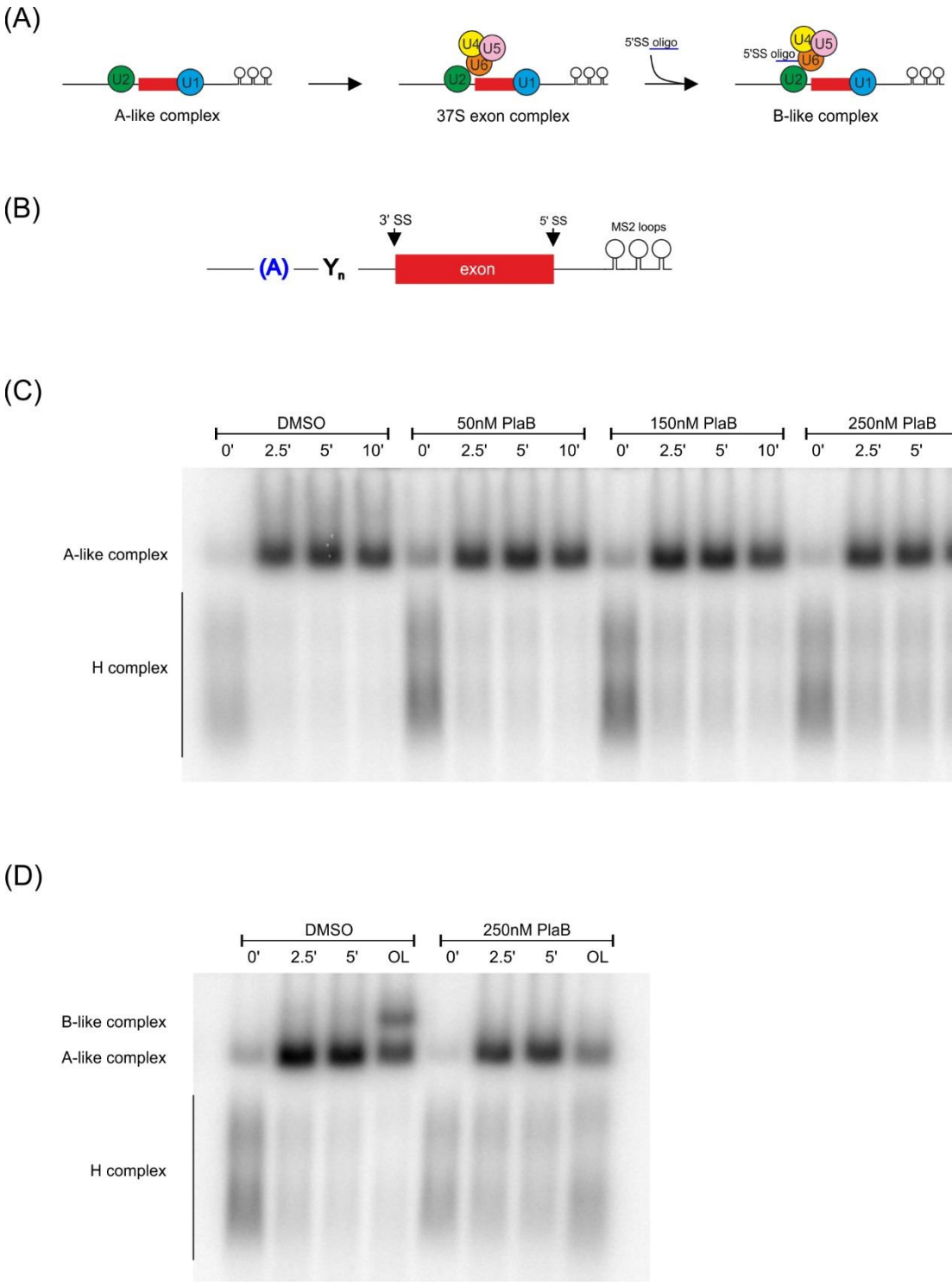


Figure 4.9- **Pladienolide B inhibits the formation of a cross-exon B-like complex.** (A) Complex transitions between the A-like and the B-like complex. (B) Scheme of the MINX exon RNA. (C) Cross-exon complexes were assembled in the presence of different concentrations of Pladienolide B or DMSO. Complex formation was analyzed on native agarose gels in the presence of 0.15 µg/µl heparin. Bands were visualized with a PhosphorImager and are indicated on the left. None of the tested concentrations had a negative impact on the assembly of the cross-exon A-like complex. (D) Cross-exon complexes were allowed to form for 2.5 minutes and then the reaction was incubated for an additional 2.5 minutes with or without an 100-fold molar excess of 5' splice site oligonucleotide (OL). Complex formation was analyzed on native agarose gels in the presence of 0.15 µg/µl heparin. Bands were visualized with a PhosphorImager and are indicated on the left.

Native gel electrophoresis of the assembled cross-exon complexes shows that the presence of Pladienolide B in the nuclear extract has no impact on the A-like cross-exon complex formation. At all Pladienolide B concentrations tested the A-like cross exon complex forms efficiently in less than 2.5 minutes and remains stable over the duration of the experiment (Figure 5.9c). Similar results are seen with the DMSO control. Thus, Pladienolide B does not hinder the binding of U2 to the branch site region of the exon RNA. Due to the presence of heparin, loosely associated tri-snRNP does not survive the native gel analysis.

Exon complexes with loosely associated tri-snRNP can be converted into a B-like complex upon addition of a small RNA oligomer that contains a 5' splice site sequence. This transition is characterized by the stable, heparin-resistant binding of the tri-snRNP, as described in previous publications (Schneider et al., 2010b). Since the Pladienolide B compound prevents the association of the tri-snRNP with cross-intron A complexes I next investigated if a similar effect is observed for the cross-exon complexes. To examine if this is the case, cross-exon complexes were assembled in the presence of Pladienolide B and then a 100-fold molar excess of the 5'SS oligo was added to the reaction, in order to trigger the transition into a stable B-like complex. The DMSO control in figure 5.9d shows that the addition of the 5'SS oligo after allowing A-like complexes to form for 2.5 minutes, leads to the formation of a B-like complex migrating higher than the A-like complex on a native agarose gel (Figure 5.9d). Pladienolide B does not affect the formation of the exon complex, but addition of the 5'SS oligonucleotide no longer induces the formation of a B-like complex. These results are very similar to the observations made in the context of the cross-intron A complex. In both cases, Pladienolide B prevents the association of the tri-snRNP with a pre-existing complex containing U1 and U2.

4.2 Cancer associated mutations in SF3b1

With the emergence of global transcriptomic analysis of cancer patient samples, spliceosome associated proteins have been identified as mutational targets in multiple different types of cancers (Yoshida and Ogawa, 2014). A number of cancer-linked mutations have been identified in the U2-associated, SF3b subunit, SF3b1 (Figure 2.14). The most common ones (i.e. hot-spot mutations) are clustered in the HEAT domain of human SF3b1 (Figure 2.14). The K700E mutation is found most frequently and is often found in myeloid malignancies (Boultwood et al., 2014; Yoshida and Ogawa, 2014). SF3b1 mutations lead to changes in alternative splicing patterns and appear to affect branch site selection (see Introduction). SF3b1 mutations have yet to be found as homozygous mutations in patient samples and the generation of homozygous mutated human cell lines in the laboratory setting has not been successful. This indicates that a change in both alleles of the SF3b1 gene is lethal for the effected cell. This circumstance greatly complicates the investigation of these mutations, both *in vitro* and *in vivo,* since the effects are always observed in the presence of the wild-type protein, which might mask or make the effects of the mutant protein milder. This makes it necessary to develop strategies that can be used to clearly distinguish the effects of the mutated and wild-type protein in all *in vitro* experiments.

4.2.1 Lentivirus experiments to generate mutant SF3b1 cell lines

During my thesis, I wanted to generate a cell line that expressed a tagged SF3b1 from an artificial gene that was introduced stably into the host-cell genome. Then I could use the affinity-tag to selectively purify my protein of interest or any potential complexes that it was a part of. Initially I chose a lentiviral transduction system to integrate the artificial gene into the host-cell genome without disrupting the endogenous SF3b1 expression. For this a 3 plasmid system was employed to produce lentiviral particles, which could infect the target cells. The plasmids used were obtained via Addgene (Trono Laboratory) and allowed for stable expression of the gene of interest (GOI) under control of a human EF1-alpha promoter.

4.2.1.1 Initial experiments

After generating the appropriate expression plasmids (see methods) I wanted to test if the lentivirus particles are capable of transducing my target cell line and if the

transduction leads to the creation of a stable cell line expressing my gene of interest (SF3b1 tagged with C-terminal His tag, or N-terminal SBP-tag and a C-terminal His-tag). To evaluate this, I created lentiviral particles containing only pWPI-GFP and transduced HEK293 and HeLa S3cells. The results of the initial transformation tests confirmed that the system of lentiviral transduction is capable of creating cells that stably express the gene located on the pWPI vector. The transduction efficiency for HEK293 cells was close to 100% under optimized conditions and in the range of 85% for the transduction of HeLa cells. As soon as the SF3b1 is inserted into the pWPI vector the transduction efficiency decreased to 90% for HEK293 cells and 75% for the harder to transduce HeLa S10 cells. This sudden decrease in the transduction efficiency raised the question whether the presence of the SF3b1 gene interferes with the expression of the GFP marker present on the vector, or if the SF3b1 gene itself cannot be expressed from the vector used. To evaluate this question the pWPI-SF3b1 vector was transfected transiently into HEK293 cells and the expression of GFP, as well as the GOI expression, was evaluated. The result of the transient expression experiments confirm that the pWPI vector containing the GOI is, capable of expressing both GFP and the GOI. The transduced cells were then sorted based on the presence of GFP expression, which is directly coupled to the expression of the GOI. After sorting, the cell population expressing GFP was expanded and retested for GOI expression after fixed time intervals. However, the expression of tagged SF3b1 could only be confirmed for short periods after transduction. Afterwards the sorted population of cells failed to express tagged protein, be it tagged wild-type SF3b1 or mutated SF3b1 protein. This loss of expression indicates a potential selective disadvantage for artificial SF3b1 expression, leading to the slow loss of the artificial gene inserted via lentiviral transduction.

4.2.1.2 Large scale transient transfections and SBP-purification

The establishment of lentiviral transduction protocols also created the alternative option of performing large scale transient transfection with the aforementioned expression constructs. These expression constructs were then used to transiently transfect large amounts of cells with a PEI-based transient transfection method. The cells were then used to produce nuclear extract for subsequent purification of tagged SF3b1.

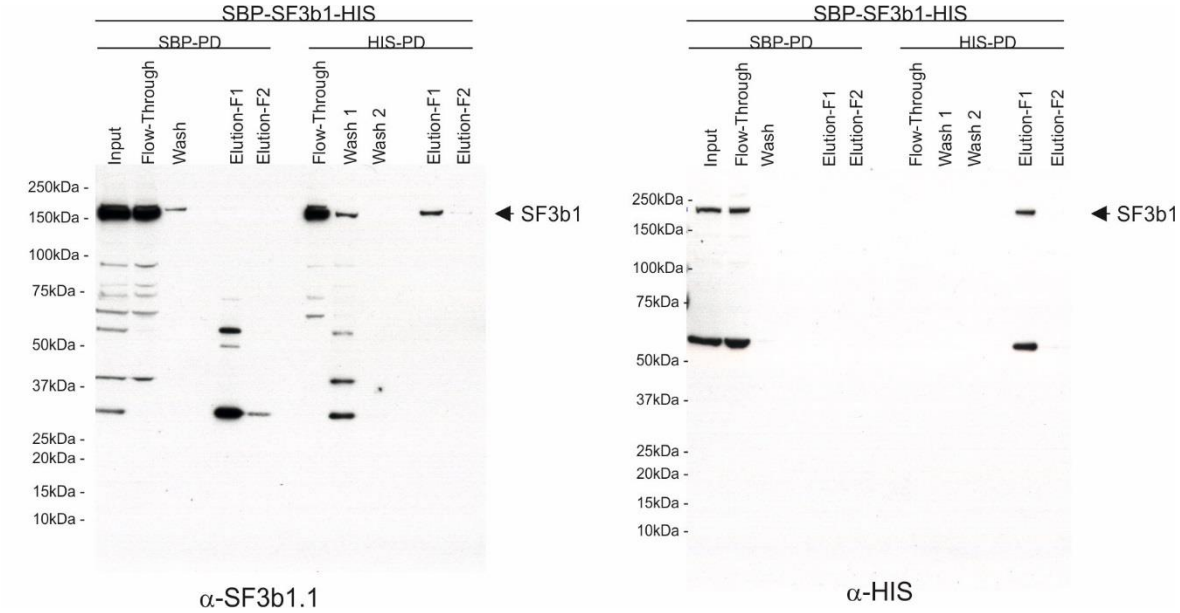


Figure 4.10 – **Tagged SF3b1 purification from transiently transfected HEK293T nuclear extract.** HEK293T cell derived nuclear containing N- and C-terminally tagged SF3b1 was used to purify SF3b1 using streptavidin beads (left) or Talon beads (right). The bound proteins were eluted using biotin for the streptavidin beads and with 0.1 M imidazole for the Talon beads. The Input, wash and eluate fractions were then analyzed on western blots. The blots were then developed with antibodies targeting the SBP- or the histidine-tag and visualized using HRP based chemiluminescence.

The construct expressing N- and C-terminally tagged SF3b1 (SBP-SF3b1-HIS) did lead to the expression of tagged protein. However, the input lane in figure 5.10 shows that at least some of the tagged protein is cleaved, evidenced by the degradation bands appearing on the western blot. Furthermore, the binding efficiency to streptavidin beads or Talon beads was very poor as evidenced by similar band intensities of the input and column flow-through. The purification of tagged SF3b1 using the N-terminal SBP-tag fails to purify full-length protein (a truncated version can be eluted). The N-terminal His-tag, however, can be used to purify low amounts of the tagged protein. However, significant amounts of unspecific contaminants are co-eluted (data not shown). In summary, the tested N- and C-terminally tagged SF3b1 protein is expressed in human cells, but is not suitable for the purification of SF3b1-containing complexes.

4.2.2 Characterization of K562 cell line

4.2.2.1 Introduction

To study the effect of mutations in SF3b1, I instead used a K562 cell line expressing a FLAG-tagged version of SF3b1 that was generated during my thesis by the laboratory of Dr. James Manley (Columbia University). The cell line was created by

using the CRISPR/Cas9 system to trigger a homologous recombination event that led to the replacement of one SF3b1 allele with an artificial SF3b1 gene encoding a protein with a 6xHIS-FLAG tag at the N-terminal end plus a substitution of lysine 700 with glutamate (K700E). As a control, a second cell line containing the same tag added to the N-terminal end of a wild-type SF3B1 allele was created.

4.2.2.2 Validation of FLAG-tagged SF3b1 expression

To validate expression of the target wild-type or mutant SF3b1, nuclear extracts of the K562 cell lines were prepared using the small scale nuclear extract protocols described in the method section. The nuclear extract of both cell lines was then tested for expression of FLAG-tagged proteins by western blot, using antibodies specifically recognizing SF3b1 or the FLAG tag.

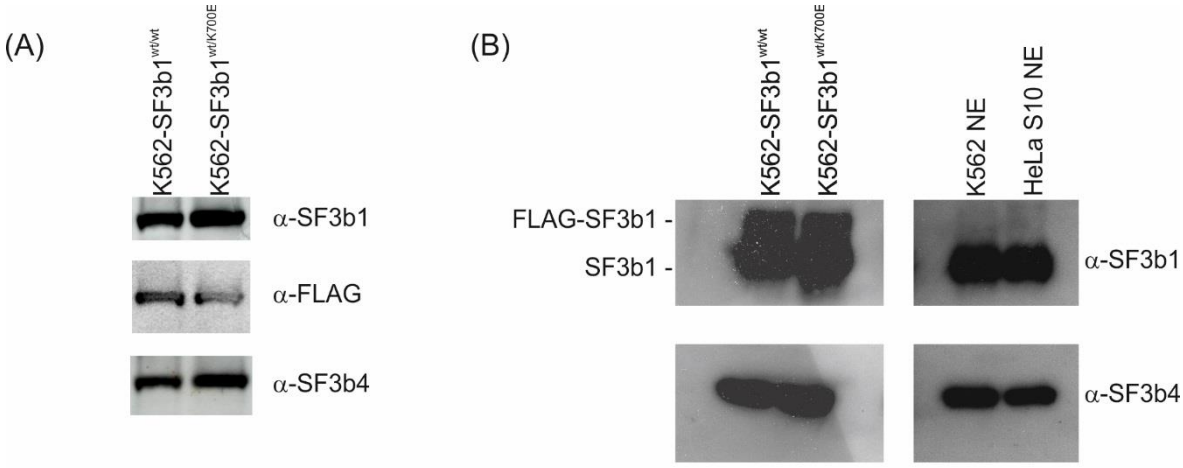
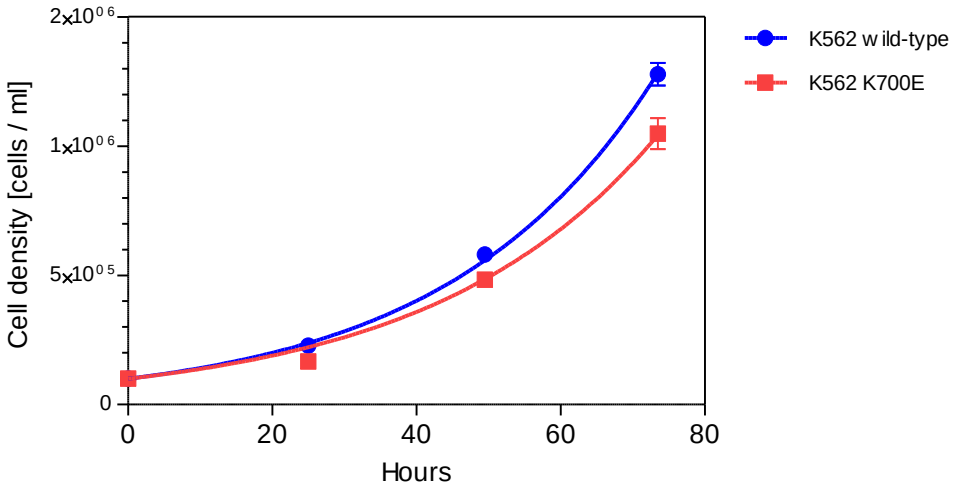


Figure 4.11 – **Both K562 cell lines express FLAG-tagged SF3b1 protein.** Western blot of nuclear extract obtained from K562 cells. Membranes was incubated in primary antibodies against SF3b1, FLAG-tag and SF3b4. Signal was detected by chemiluminescence using HRP-conjugated secondary antibodies.

As shown in figure 5.11, both cell lines express a FLAG-tagged protein with a size of approximately 155kDa and the relative expression of tagged protein is nearly identical in the two cell lines. This result confirms that the tagging of SF3b1 was successful and that both cell lines express the N-terminally tagged protein. A comparison of the western-blot signals, suggests that the ratio of endogenous SF3b1 to the tagged protein is almost 1:1.

4.2.2.3 Presence of the K700E mutation causes an increased doubling time in K562 cells

After confirming that both cell lines express tagged SF3b1, I characterized the cell line in more detail. The physical appearance of the K562 cells did not show any irregularities as a consequence of the introduction of the K700E mutation; cell size, granularity and visual appearance remained identical. To analyze the growth characteristics the number of cells generated with both cell lines was determined over a period of 72 h, and the average doubling time of both cell lines was calculated (Figure 5.12).



	K562 wild-type	K562 K700E
Doubling Time	19.96	21.72

Figure 4.12 – **Determination of the doubling times of K562 cells expressing a FLAG-tagged SF3b1^{wt} or SF3b1^{K700E} protein.** Cell density was determined by manual counting with six biological and two technical replicates for each time point and is plotted as a function of time.

Interestingly the introduction of the K700E amino acid substitution in SF3b1 cells causes an increase in the doubling time by roughly 2 h or 10%, from ~20 h to 22 h. Although the doubling-time is increased after the introduction of the mutation, a negative effect on population vitality could not be observed. The absence of an increased cell mortality could be an indicator that while the presence of the mutation decreases the rate of cell growth, it does not cause an increased incidence of apoptotic events. The observation that the K700E mutation also has an impact on the doubling time could indicate a widespread negative effect on cellular homeostasis.

4.2.2.4 Splicing of MINX pre-mRNA in the presence of SF3b1^{K700E}

I next checked if the K700E mutation had an impact on the efficiency of *in vitro* splicing. To test this radiolabeled MINX pre-mRNA was added to the nuclear extract under splicing conditions, and the intermediates and products of the splicing reaction were separated on a denaturing polyacrylamide gel.

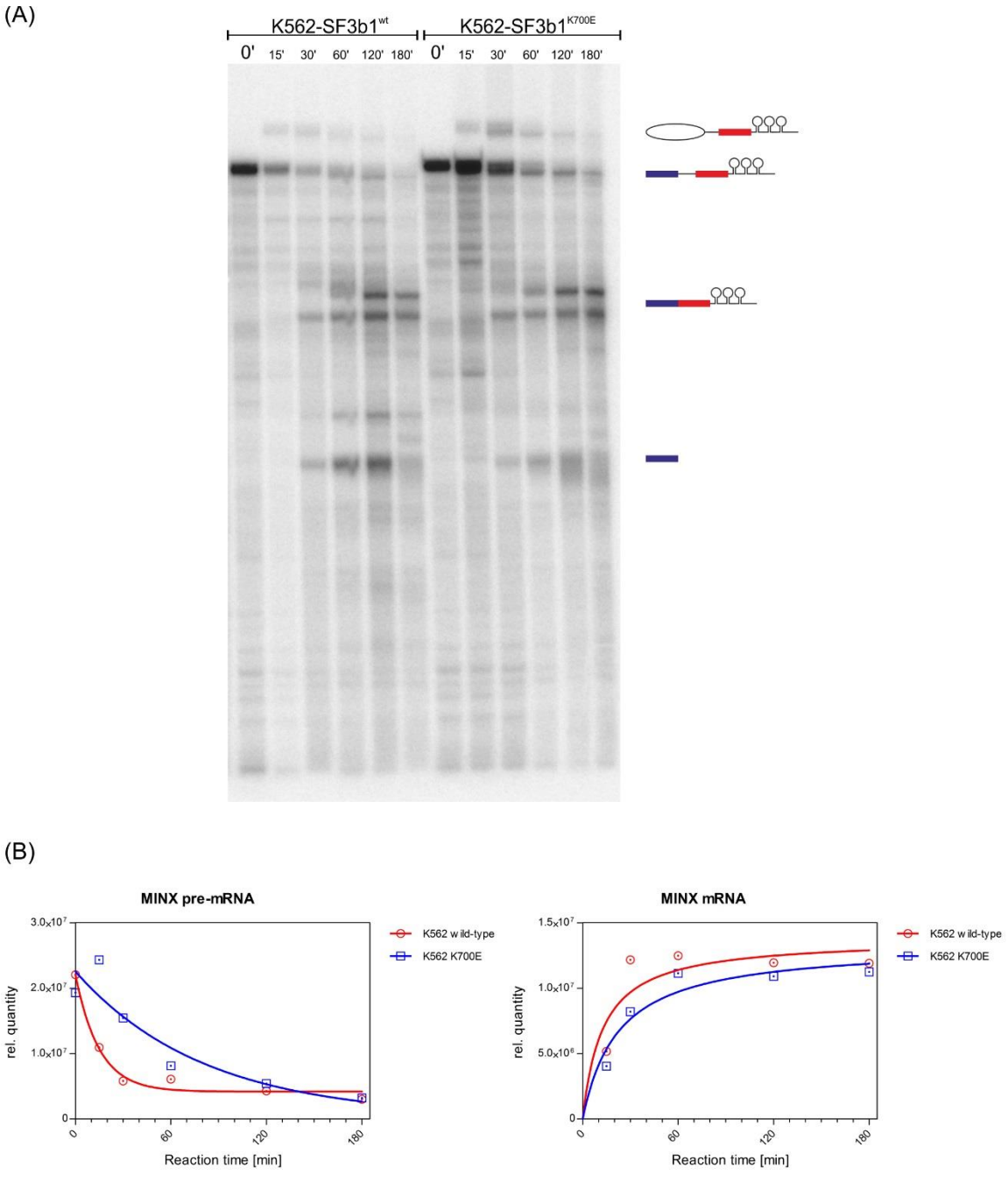


Figure 4.13 – **MINX pre-mRNA splicing in nuclear extract obtained from K562 cells expressing wild-type and tagged SF3b1.** Splicing reactions were incubated for the indicated times and RNA was isolated and analyzed on a 10% PA gel. The positions of the splicing intermediates, products and unspliced pre-mRNA are indicated on the right. In (B) quantification of mRNA formation is shown over time.

As shown in figure 5.13a MINX-MS2 pre-mRNA is efficiently spliced into mature mRNA in both nuclear extracts. However, quantification of the band intensities demonstrates that the rate at which the different steps of splicing occurs is different. The wild-type nuclear extract splices MINX pre-mRNA faster than the extract containing SF3b1^{K700E}. This is evident when comparing the rate of increase in the signal for the mature mRNA (Figure 5.13b). The splicing reaction in the wild-type nuclear extract reaches a plateau in the amount of mRNA formed after 30 minutes, while in the K700E nuclear extract a similar % of mRNA is reached first after 60 minutes. In addition, the splice reaction containing SF3b1^{K700E} accumulates more intron-lariat-3' exon intermediates at the early time points than the wild-type reaction. These results confirm that the nuclear extract obtained from K562 cells is capable of splicing MINX pre-mRNA. In addition, the experiments confirm that the expression of SF3b1^{K700E} does not substantially inhibit the splicing of MINX pre-mRNA. However, the quantification of the rate of mRNA formation shows that the presence of SF3b1^{K700E} leads to a slowdown of the splicing reaction.

4.2.2.5 Preparation and characterization of nuclear extract containing SF3b1^{K700E}

SF3b1 is a major component of the heptameric SF3b complex, which is a major subunit of the 17S U2 snRNP. I therefore next checked if the mutated protein affects the formation of the 17S U2 particle. To investigate this nuclear extract of both cell lines was prepared using the small scale nuclear extract protocol and then separated by 5-20% sucrose gradient centrifugation at 150 mM KCl.

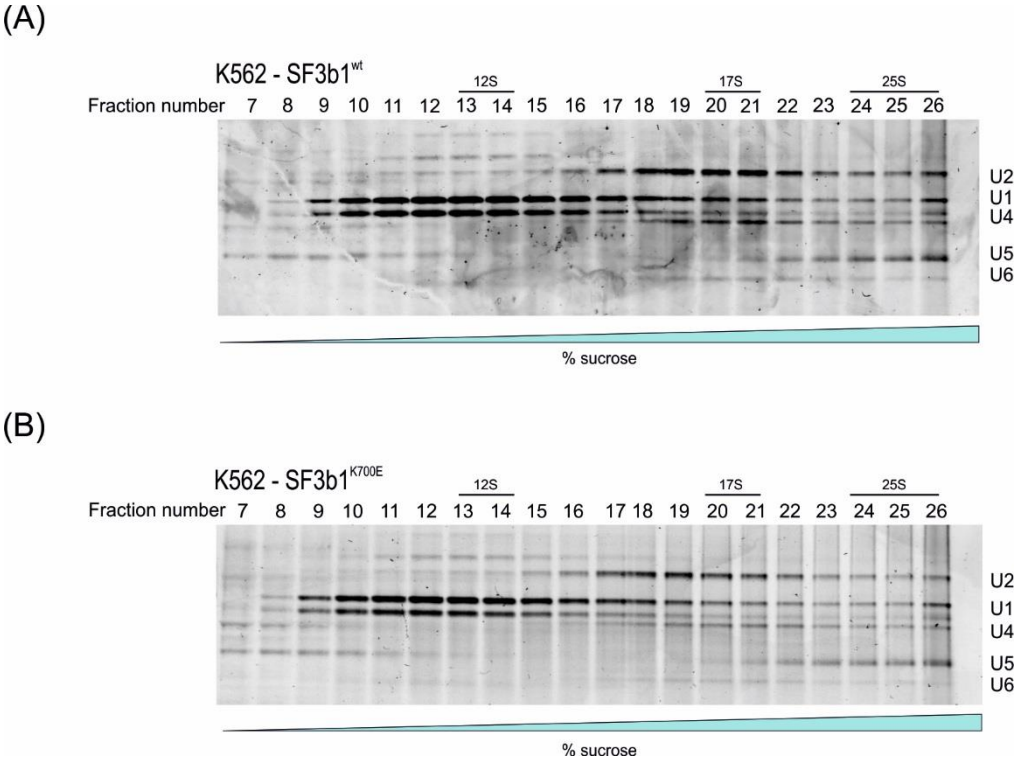


Figure 4.14 – **Sedimentation behavior of U1, U2, U5 and U4/U6 snRNPs in the presence of SF3b1^{K700E}.** Gradient centrifugation of nuclear extract obtained from K562 cells expressing (A) wild-type SF3b1 or (B) SF3b1^{K700E}. Gradients were fractionated by hand from the top. RNA was isolated, separated by PAGE and visualized by staining with SYBR gold. The position of the snRNAs is indicated on the right. The abundant small RNA directly below U1 is likely 5.8 rRNA.

The U snRNA distribution in the gradient is nearly identical in both nuclear extracts. The U1 snRNP peaks in fractions 11-14, corresponding to the 12S region of the gradient. The larger 17S U2 snRNP peaks in fractions 19-21. Most of the U5 and U4/U6 snRNPs, which form the tri-snRNP, can be found in the heavy fractions 22-26 of the gradients. Thus, the SF3b1^{K700E} mutation does not substantially alter the distribution of the U2 snRNP. This is an indication that the presence of SF3b1^{K700E} does not interfere with the formation of the 17S U2 snRNP to a degree that causes a significant change in particle size and/or geometry, since this would lead to a change in the sedimentation behavior of 17S U2.

4.2.3 Isolation and characterization of 17S U2 particles

A major goal of my studies is to determine what effect cancer related mutations in SF3b1 have on the structure and composition of the 17S U2 snRNP. To investigate these potential consequences, I isolated 17S U2 snRNPs and compared the particles assembled with SF3b1^{wt} with the particles assembled in presence of SF3b1^{K700E}. As mentioned in the introductory chapter SF3b1^{K700E} can only be

expressed in cells that still contain one allele expressing the wild-type protein. Thus, to isolate only the wild-type or mutant U2 snRNPs I performed affinity selection with antibodies against the N-terminal FLAG-tag of SF3b1.

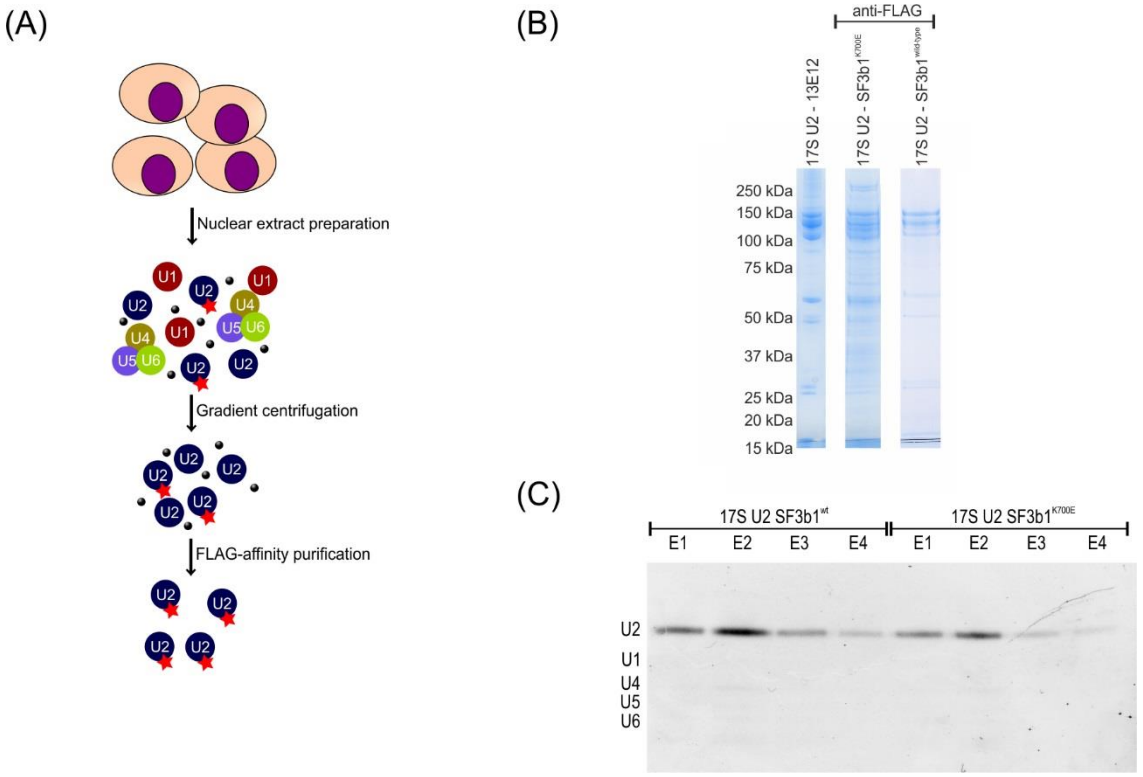


Figure 4.15- **Isolation of 17S U2 particles via the N-terminal FLAG-tag.** (A) Scheme of the experimental approach used for the purification of 17S U2 particles from K562 nuclear extract. (B) Protein composition of anti-FLAG purified complexes in comparison with 17S U2 snRNPs isolated from nuclear extract with an antibody against SF3b1. The RNA composition of the eluted complexes is shown in (C). RNA was isolated from the indicated anti-FLAG column eluates, separated by PAGE and visualized by SYBRgold staining.

The 17S U2 purification procedure is summarized in figure 5.15a. Nuclear extract from wild-type cells or from cells containing mutated SF3b1 was prepared using manual douncing (as described in chapter 4.3.5). The nuclear extract was then subjected to glycerol gradient centrifugation. Fractions corresponding to the 17S region of the gradient were then pooled, diluted with G-150 buffer and applied to beads containing covalently bound anti-FLAG antibodies. Using only the 17S region of the gradient ensures that little or no free SF3b complexes or SF3b1 are co-purified. After extensive washing, complexes were eluted in native form by competitive elution with a triple-FLAG peptide.

The eluted complexes were analyzed for their protein and RNA content. The RNA composition of the purified particles was analyzed by PAGE followed by SYBR gold

staining, which shows that solely that U2 snRNA is present in the eluates, with no contamination of other snRNAs. The protein composition is shown in panel C. The FLAG-tagged U2 snRNP contain nearly identical proteins as the purified endogenous 17S U2 snRNPs. Importantly, bands corresponding to high molecular weight SF3b proteins (SF3b1, SF3b2, SF3b3) and SF3a proteins are clearly visible, confirming the presence of 17S U2 snRNPs.

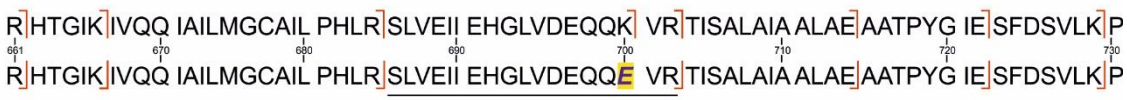
Thus, the purification approach outlined above is suitable to isolate native 17S U2 particles containing 6xHIS-FLAG-tagged SF3b1 protein.

4.2.3.1 Mass-spectrometry confirms the presence of solely SF3b1^{K700E} in the purified 17S U2 snRNPs

I next checked whether the 17S U2 snRNPs purified from the SF3b1^{K700E} K562 nuclear extract contained exclusively the mutated SF3b1 protein. The substitution of the lysine at position 700 with a glutamate changes the fragmentation pattern after trypsin digestion (which cleaves after lysines or argenines) of the protein, as is shown in figure 5.16a, making the identification of the mutated protein possible. The MS-analysis of the SF3b1 protein present in the purified 17S U2 particle shows that the peptide obtained only upon K700E substitution is found solely in particles isolated from cells expressing the SF3b1 mutant, and that the peptides generated by tryptic digestion of the wild-type SF3b1 protein are absent in this sample. This confirms that the purification strategy using an N-terminal FLAG tag on SF3b1 is capable of selectively isolating 17S U2 snRNPs containing the tagged wild-type or mutant SF3b1 protein.

(A)

Sequence of wild-type SF3b1 (aa 661 - 730)



Sequence of K700E SF3b1 (aa 661 - 730)

(B)

Peptide	SF3b1^{wt}	SF3b1^{K700E}
SLVEIIEHGLVDEQQK	12	0
SLVEIIEHGLVDEQQKVR	3	0
SLVEIIEHGLVDEQQEVR	0	3

Figure 4.16- **Fragmentation scheme of SF3b1^{K700E}**. The normal digestion pattern after trypsin-digestion is shown in (A). The introduction of the K700E amino-acid substitution leads to the removal of a cleavage site and changes the fragmentation pattern. (B) The MS-analysis shows that the sample isolated from the nuclear extract containing SF3b1^{K700E} contains only peptides associated with the changed fragmentation pattern. Numbers shown correspond to the number of peptides sequenced.

4.2.3.2 MS-analysis of the protein composition of the 17S U2 complex

To gain detailed information about their exact protein composition, affinity purified 17S U2 particles containing wild-type SF3b1 and particles containing SF3b1 with the K700E point mutation were analyzed by mass spectrometry.

Table 4.2 - **Proteins identified by LC-MS/MS in 17S U2 particles containing wild-type SF3b1 or SF3b1^{K700E}.** Total peptide counts for each individual protein are shown for two/ three independent preparations. The proteins are gouped according to function or association within the spliceosome.

	kDa	human gene name	17S U2 wild-type	17S U2 wild-type	17S U2 wild-type	17S U2 K700E	17S U2 K700E	17S U2 K700E
Sm proteins								
B/B'	24,6	SNRPB	42	34	40	24	21	19
D1	13,3	SNRPD1	9	11	8	3	3	6
D2	13,5	SNRPD2	26	24	28	15	11	23
D3	13,9	SNRPD3	30	40	29	12	9	16
E	10,8	SNRPE	16	18	13	4		13
F	9,7	SNRPF	7	3	6	5	2	8
G	8,5	SNRPG	5		6	3	3	2
17S U2 snRNP								
U2A'	28,4	SNRPA1	67	96	51	39	36	35
U2B"	25,5	SNRPB2	25	16	20	15	7	12
SF3a120	88,9	SF3A1	142	288	140	78	147	78
SF3a66	49,3	SF3A2	65	27	55	57	17	49
SF3a60	58,9	SF3A3	26	180	19	13	70	15
SF3b155	145,8	SF3B1	333	528	271	198	254	270
SF3b145	100,2	SF3B2	180	293	189	138	177	128
SF3b130	135,6	SF3B3	425	609	483	289	304	255
SF3b49	44,4	SF3B4	16	46	11	3	9	12
SF3b14a/p14	14,6	SF3B6	31	37	30	14	21	18
SF3b14b	12,4	PHF5A	24	18	19	13	12	13
SF3b10	10,1	SF3B5	4	6	10	5	2	8
17S U2 related								
tat SF1	85,9	HTATSF1	194	269	177	34	40	30
hPRP43	90,9	DHX15	113	298	114	123	154	148
SPF45	45,0	RBM17	34	80	33	32	47	35
SPF30	26,7	SMNDC1	3		2	1	3	2
U2AF65	53,5	U2AF2	4	36	3	0	10	8
U2AF35	27,9	U2AF1	1	3	2	1	3	2
SPF31	29,8	DNAJC8	7	11	1	4	2	2
hPRP5 / DDX46	117,4	DDX46	175	310	117	0	18	0
SR140	118,3	U2SURP	102	219	87	99	127	112
CHERP	103,7	CHERP	63	160	38	68	100	76
PUF60	59,9	PUF60		6			5	
SF3b125	103	DDX42		23				
A proteins								
RBM10	103,5	RBM10		152	29	40	62	
SF1	68,3	SF1	3					6
RBM5/LUCA15	92,2	RBM5	22	34	15	28	40	25
SF4 (F23858)	72,5	SUGP1	4		4	3	4	3

The comparison of the U2 snRNPs containing wild-type and SF3b1^{K700E} showed a similar protein composition for both complexes. All major 17S U2 associated proteins, including all SF3b and SF3a subunits, could be detected in both complexes with apparently similar stoichiometry based on the number of peptides sequenced. However, two significant differences are apparent. First, the complex containing wild-type SF3b1 appears to contain significantly more hTAT-SF1 protein than the 17S U2 snRNP containing SF3b1 with the K700E amino acid substitution. Secondly, the DExD/H-box protein Prp5 is almost completely absent in the complexes

containing SF3B1^{K700E}, while significant amounts of it can be found in the complexes containing wild-type SF3b1. These findings are of potential significance, because Prp5 is required for A complex formation and the stable recognition of the branch point by the 17S U2 snRNP (Will and Lührmann, 2002).

4.2.3.3 hPrp5 is underrepresented in 17S U2 particles containing SF3b1^{K700E}

The mass spectrometry data indicated that the presence of SF3b1^{K700E} in the 17S U2 particle leads to decreased levels of hPrp5 and hTAT-SF1 being found. To validate these data, the abundance of the two proteins was evaluated via western blot using polyclonal antibodies against the two proteins of interest, and SF3b4 as a loading control. The signal for Prp5 was substantially lower in the SF3b1^{K700E} containing particle versus the wild-type tagged SF3b1, confirming the mass spectrometry results. A much smaller difference was, however observed for hTAT-SF1, where quantification of the bands suggested only a 40% drop in signal intensity in the K700E versus wild-type particles.

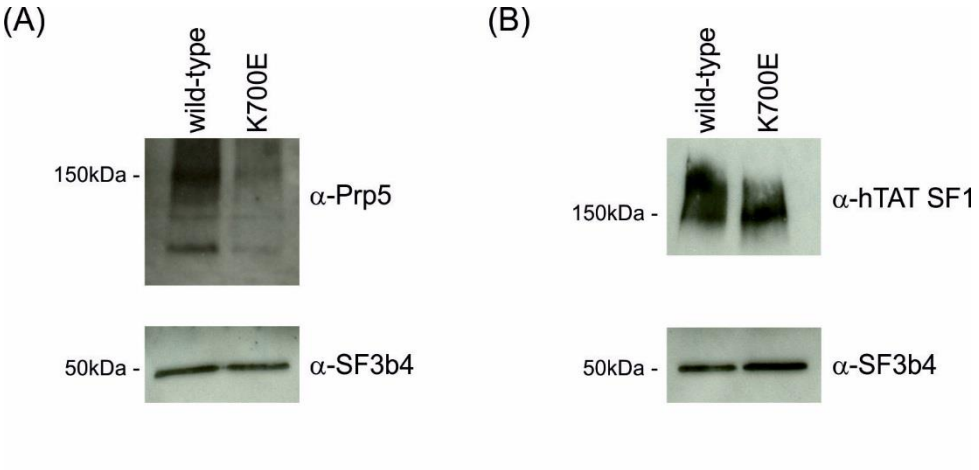


Figure 4.17- **Western blot analysis of hPrp5 and hTAT-SF1 levels in purified U2 snRNPs.** Purified 17S U2 snRNPs were analyzed by western blot after PAGE. Membranes were incubated with primary antibodies targeting hPrp5, hTAT-SF1 and SF3b4. Bound antibody was detected by chemiluminescence using HRP-conjugated secondary antibodies.

4.2.3.4 Cross-linking of 17S U2 particles

The lower amounts of hPrp5 and hTAT-SF1 observed in the 17S U2 particles containing SF3b1^{K700E} could be due to the K700E mutation directly affecting their interactions with U2 particles or due to changes in the SF3B1 HEAT domain structure. The SF3B1 HEAT domain is a scaffold for the binding of several other SF3b proteins (Cretu et al., 2016). Nevertheless, it is not known if human Prp5 and hTAT-SF1 proteins interact with SF3B1 and if they contact its HEAT domain. I

therefore carried out protein-protein crosslinking of 17S U2 particles in collaboration with Dr. Cindy Will and Dr. Olexandr Dybkov in our department. Large amounts of 17S U2 were affinity purified from HeLa nuclear extract using anti-peptide antibodies against SF3B1 (Will and Lührmann, 2002). After concentrating the purified particles, the crosslinking reagent BS3 (bis(sulfosuccinimidyl)suberate), which is an amine-to-amine crosslinker, reacting with primary amines as found on lysines, was added and a titration was performed to find the optimal concentration of BS3 where individual particles were still present. In the next step, crosslinked U2 particles were centrifuged on a gradient to separate individual crosslinked 17S particles from particles crosslinked to each other, which would have a higher S value. Crosslinked amino acids were afterwards identified by mass spectrometry.

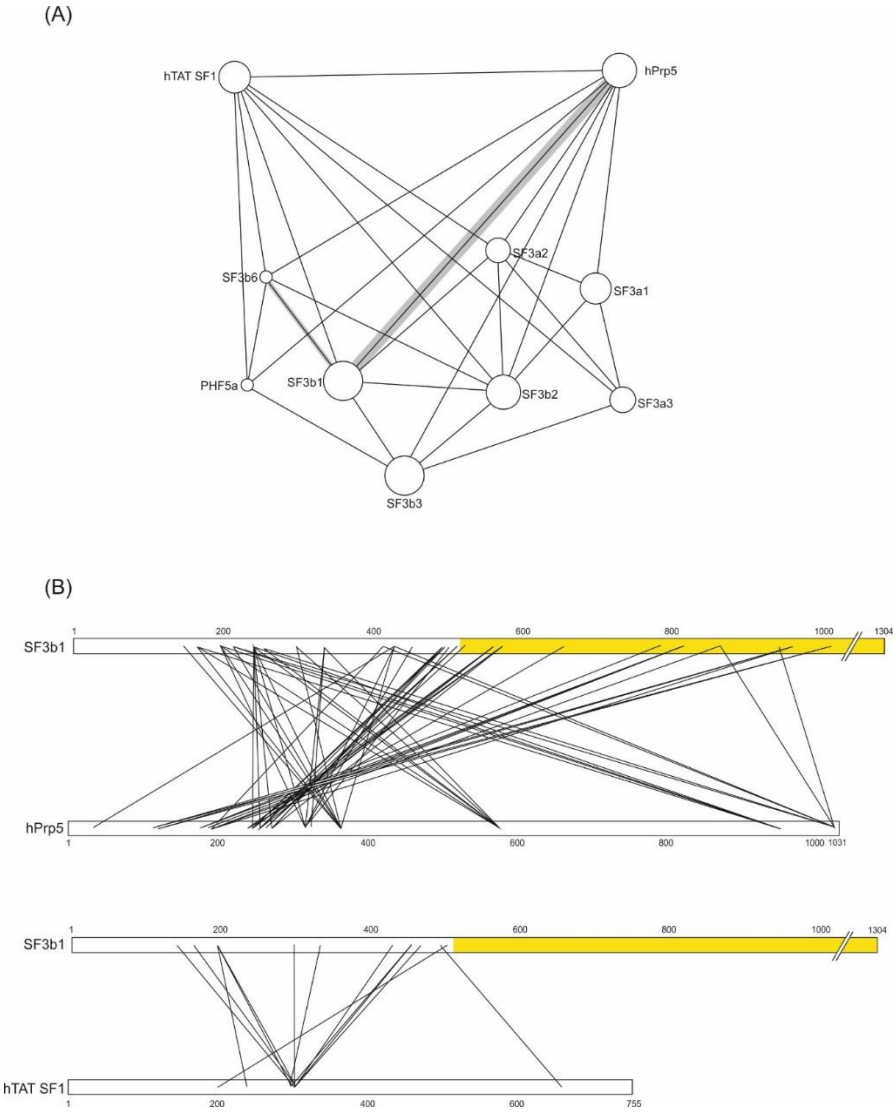


Figure 4.18 – **Cross-linking of 17S U2 snRNPs.** 17S U2 particles were affinity purified from HeLa nuclear extract and concentrated prior to cross-linking with BS3. (A) Protein-protein interaction network of the 17S U2

particle. Proteins are shown as nodes on the graph. Only proteins cross-linked hPrp5 or hTAT-SF1 are shown. Observed cross-links are shown as lines and high numbers of cross-links are indicated with grey shading. (B) Graphical representation of cross-links observed between SF3b1 and hPrp5 or hTAT-SF1. The HEAT-domain of SF3b1 is shown in yellow. (FDR = 1%, unfiltered).

hPrp5 and hTAT-SF1 could be cross-linked to multiple proteins of the 17S U2 particle (Figure 5.18a). The observed number of cross-links between SF3b1 and hPrp5 is especially high. The location of the cross-links found between SF3b1 and hPrp5, and between SF3b1 and hTAT-SF1 is shown in figure 5.18b. The cross-links between SF3b1 and hTAT-SF1 are mainly found in the N-terminal region of both proteins, with no cross-links involving residues of the HEAT-domain of SF3b1. hPrp5 and SF3b1 on the other hand form significantly more cross-links. Several lysines of SF3b1s N-terminal region could be cross-linked to hPrp5, but there were also crosslinks with lysines in the N-terminal HEAT repeats of the protein (but not K700). This is consistent with the idea that the K700E mutation might directly affect the binding of hPrp5 because it might change the structure of the HEAT domain. However, this structural change could also change the binding of proteins like SF3b2 and SF3b3 and indirectly alter the binding of hPrp5 or hTAT-SF1.

4.2.4 Does 17S U2 containing SF3b1^{K700E} function in A complex formation

I next checked if 17S U2 snRNP containing SF3b1^{K700E} can take part in spliceosomal A complex formation. In order to investigate this, I performed splicing in K562 nuclear extract containing either FLAG-tagged SF3b1^{wt} or FLAG-tagged SF3b1^{K700E}. Afterwards the reaction was applied to a linear glycerol gradient and the complexes in the A complex peak were isolated using MS2-affinity selection. Equal amounts of the peaks were loaded onto a 4-12% commercial protein gel and analyzed by western-blot. The blot was then developed using antibodies against the FLAG-tag and against SF3b4 as a loading control. The resulting signal could be correlated with the amount of tagged protein present in the A complex, which indicates how efficient the mutant 17S U2 binds.

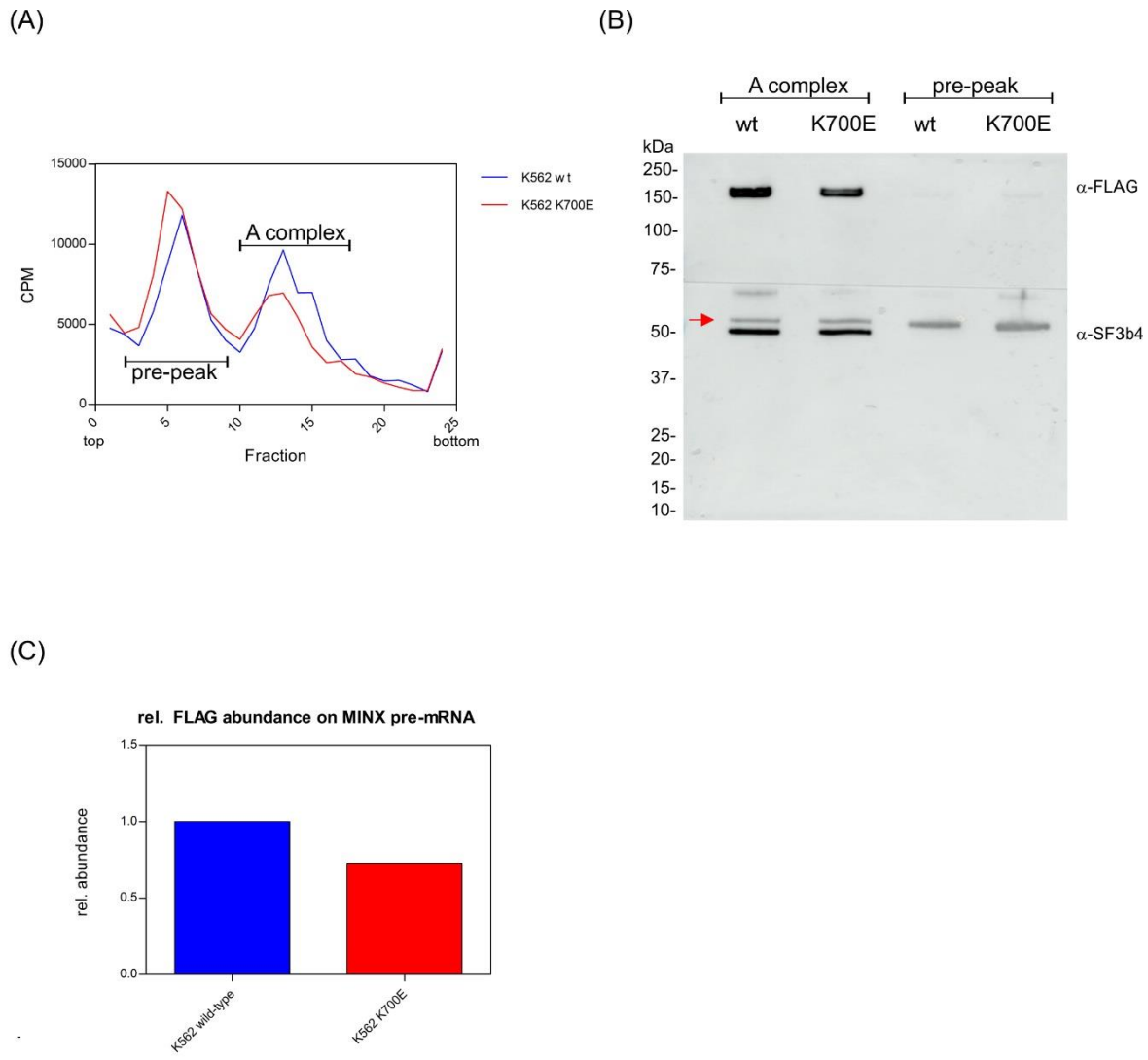


Figure 4.19 – **SF3b1^{K700E} is less abundant than SF3b1^{wt} in A complexes assembled on MINX pre-mRNA.** (A) Gradient centrifugation of A complexes assembled in K562 nuclear extract containing either FLAG-tagged SF3b1^{wt} or tagged SF3b1^{K700E}. Gradients were prepared with 10-30% (v/v) glycerol in G150 buffer and fractionated top to bottom. (B) Western blot of the pooled and precipitated peak fractions of the gradients. Proteins were detected with primary antibodies against the FLAG tag and SF3b4, which was used as loading control. (C) Relative abundance of FLAG-tagged SF3b1 in the A complex peaks.

As shown in figure 5.19a the extract containing a mixture of SF3b1^{wt} and SF3b1^{K700E} forms slightly less A complex than the extract containing only wild-type SF3b1 protein. However, both A complexes contain FLAG-tagged protein as shown by western blot with anti-FLAG antibodies, indicating that the K700E mutation in SF3b1 does not prevent 17S U2 from binding to the MINX pre-mRNA branch point region. The quantification of the western-blot (Figure 5.19b) shows that the SF3b1^{K700E} protein is less abundant (~75% of the wild-type FLAG-tagged protein) in the purified A complex relative to the amount of SF3b4 present, which serves as a loading control. Even though the total amount bound is not clear, these results show that

the FLAG-tag does not prevent the wild-type or the mutant U2 particle from binding the pre-mRNA.

These results also show that the 17S U2 containing the SF3b1 K700E mutation retains its ability to bind the branch point region of the pre-mRNA. However, the reduced occupancy when compared to the tagged wild-type protein suggests that the amino acid substitution might have a negative effect on the ability of the 17S U2 particle containing SF3b1^{K700E} to stably bind to the BPS. This could be a direct effect of the mutation on the SF3b1 structure or an indirect effect due to the changed, apparently less stable, interaction of Prp5 with 17S U2 containing the K700E mutation.

Initial attempts to purify A complexes containing solely the 17S U2 with mutated SF3b1 via anti-FLAG affinity purification allowed only small amounts of A complexes to be purified, suggesting that the FLAG-tag was not accessible for antibody binding. Purification of A complexes containing the mutation could reveal whether its stability, composition or even structure are altered by the K700E mutation. In the future, I will also examine if the mutant 17S U2 is incorporated into the spliceosomal B complex. This would mean that the A complex containing the mutant supports tri-snRNP binding and probably is not structurally or compositionally substantially different from the wild-type A complex.

4.2.5 Investigating the effect of SF3b1 mutations on alternative splice site selection *in vitro*

As mentioned in the introduction of this chapter, the mutations of SF3b1 in cancer patients and in cell lines are associated with a change in the cellular splicing patterns. RNAseq analysis of cells affected by the K700E mutation revealed significant changes in the 3'SS usage and the emergence of novel cryptic splice products. The exact cause for the emergence of cryptic splice products in the presence of SF3b1 mutations is currently unknown, but is thought to be due to a switch in branch site selection due to the less stable association of U2 containing mutated SF3b1 (Darman et al., 2015). Therefore, I wanted to investigate the effect of the SF3b1 K700E mutation using the controlled environment of an *in vitro* splicing system. Previous studies have identified certain introns, with which the splicing outcome *in vivo* is dependent on the presence or absence of SF3b1^{K700E} (Agrawal

et al., 2017; Darman et al., 2015). In addition, investigation of the K700E mutation in an *in* vitro setting would allow me to separate spliceosomes containing the mutation from their wild-type counterparts. Based on this information I designed constructs for *in vitro* splicing experiments.

4.2.5.1.1 Splicing complex formation on ZDHHC16 pre-mRNA

ZDHHC16 is one of the genes that has been identified utilize a cryptic 3'SS during splicing in the presence of SF3b1^{K700E} (Agrawal et al., 2017; Darman et al., 2015). In order to investigate the events leading to the emergence of cryptic splicing, I designed an *in vitro* splicing substrate based on the published mini-gene used for *in vivo* experiments. The artificial ZDHHC16 pre-mRNA contains two exons (exon 9 and 10) separated by an 860 nt intron derived from the ZDHHC16 intron 9 (Figure 5.20a). The 5'exon has a length of 58 nt and the 3'exon is 72 nt long. The ZDHHC16 pre-mRNA contains one 5' splice site and two branch points and 3' splice sites. The canonical branch point is located 30 nt upstream of the canonical 3'SS. The cryptic branch point (green A) is located 34 nt upstream of the intron-3'exon boundary. The cryptic 3'SS (3'cSS) can be found 20 nt upstream of the canonical splice site and 14 nt downstream of the cryptic branch point adenosine. In the ZDHHC16 -30 A>G pre-mRNA, the canonical branch point adenosine at position -30 (relative to the 3'SS) has been changed to a G, which should lead to usage of the cryptic branch point adenosine and cryptic 3'SS.

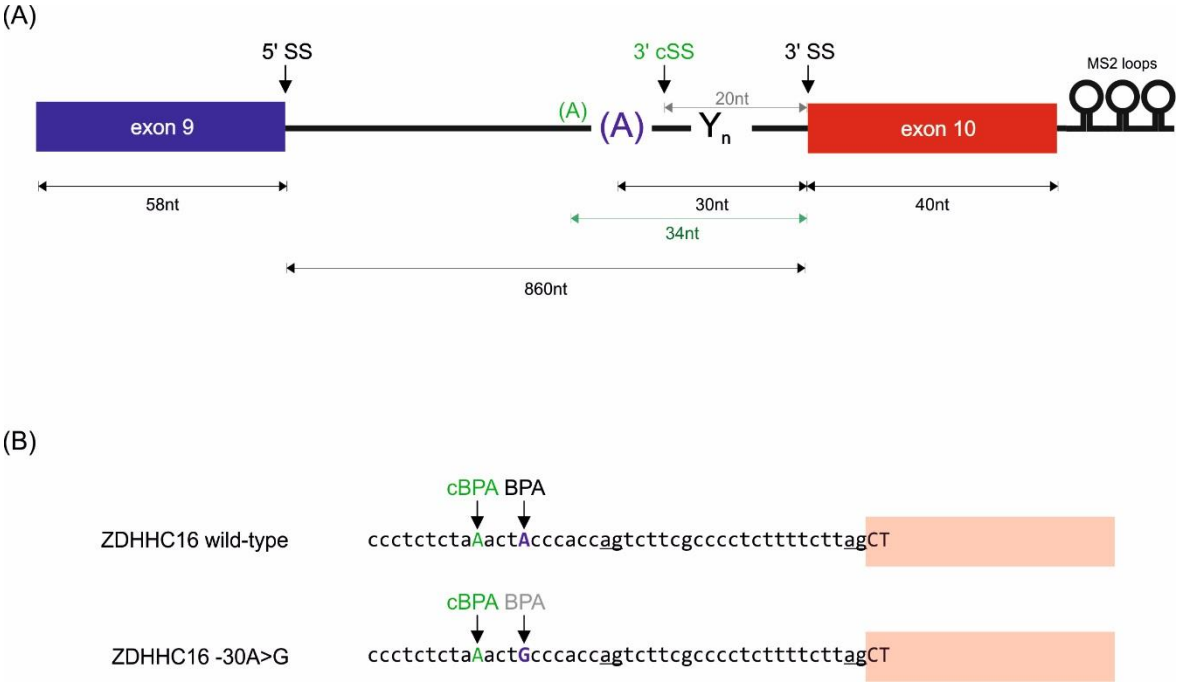


Figure 4.20 - **Schematic representation of the ZDHHC16 pre-mRNA transcript.** (A) The stretch spanning exon 9, intron 9 and exon 10 of the human ZDHHC16 gene was placed behind a T7 promoter and MS2-loops were added to the 3' end of the transcript. Exons are indicated as colored boxes and the connecting intron is depicted as thick black line. Branch point adenosine nucleotides are indicated as "A" (blue for canonical branch point and green for cryptic branch point) and the relative location of the 3' splice sites are indicated by arrows. The poly-pyrimidine tract is abbreviated as Y_n. All distances between relevant structures of the pre-mRNA template are indicated. (B) Exact sequence comparison of the 3' region of the intron 9 of ZDHHC16. The location of the cryptic and canonical BPA is indicated and the 3' splice sites are underlined. The exon 10 is shown as semi-transparent red box.

The ZDHHC16 construct was then transcribed into radiolabeled RNA and an *in vitro* splicing reaction using HeLa S10 nuclear extract was performed. For initial tests the ZDHHC16_wt pre-mRNA was used exclusively, since the RNA containing the A>G substitution at the canonical branch point was previously shown to be incompatible with splicing in the absence of SF3b1^{K700E} *in vivo*.

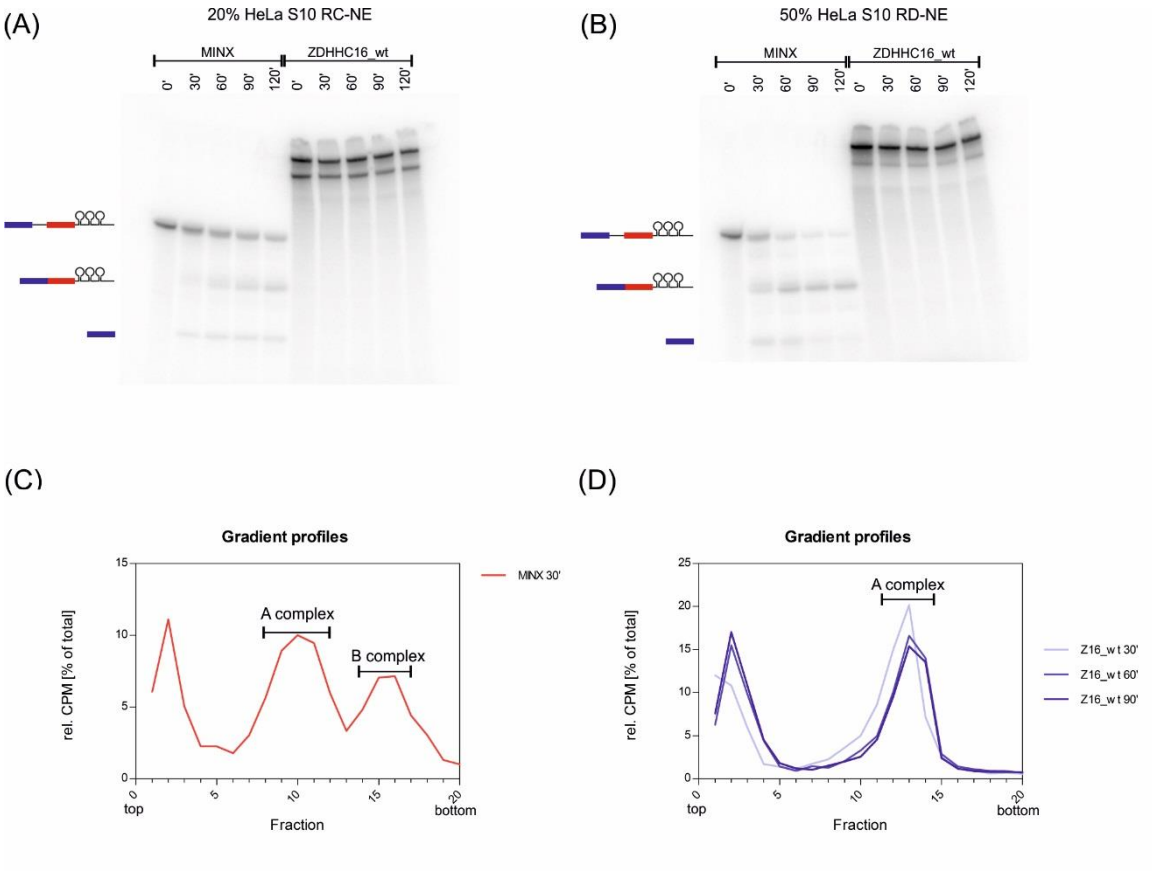


Figure 4.21- ***In vitro* splicing of the ZDHHC16_wt pre-mRNA.** MINX or ZDHHC16 pre-mRNA were incubated under splicing conditions using Röder C nuclear extract (A) or Röder D nuclear extract (B). RNA was isolated and analyzed on a 4-12% PA gel and visualized by PhosphorImager. The products of MINX splicing are indicated on the left. For ZDHHC16_wt only pre-mRNA can be found on the RNA gel. The additional smaller band observed with ZDHHC16_wt is most likely the result of RNA degradation. Gradient centrifugation of splicing complexes formed on MINX pre-mRNA (C) or ZDHHC16_wt pre-mRNA (D) after 30', 60' and 90' incubation at 30°C.

The results of the splicing tests are presented in figure 5.21. The RNA analysis shows that the ZDHHC16_wt pre-mRNA cannot be spliced in HeLa S10 RC or RD nuclear extract, while the MINX pre-mRNA is efficiently spliced to mRNA under both conditions. The faster migrating band in the ZDHHC16_wt samples is most likely the result of RNA degradation and cannot be attributed to splicing because it can already be seen at the zero time point. The analysis of splicing complex formation by gradient centrifugation shows (Figure 5.21c,d) that, while with MINX pre-mRNA a normal gradient profile consisting of A and B/B^{act} complexes is observed after 30 minutes, with the ZDHHC16_wt pre-mRNA only one peak is observed in fractions 11-14. The increased size of the ZDHHC16 pre-mRNA compared to MINX pre-mRNA (860bp vs 120bp) should result in a shift to heavier gradient fractions. Thus, it is not clear what this peak is. To clarify this, I affinity purified complexes formed on ZDHHC16 and ZDHHC16_-30 that peak in fractions 11-14. The comparison of

ZDHHC16_wt and ZDHHC16_-30 shows no difference in complex assembly and sedimentation behavior on gradients (data not shown).

4.2.5.1.2 Mass-Spectrometry of the stalled ZDHHC16 A complexes

Since no splicing of the ZDHHC16 pre-mRNA was observed, but a complex similar to the A complex was formed, I next determined the RNA and protein composition of this complex, to determine if it is indeed an A complex. For this purpose, I assembled complexes on three different pre-mRNAs (MINX, ZDHHC16_wild-type and ZDHHC16_-30) in HeLa S10 nuclear extract. Complexes assembled on the pre-mRNAs were isolated using MS2 affinity selection. Analysis of their RNA composition indicates that the complexes formed on ZDHHC16 pre-mRNA are indeed A complexes due to the presence of the pre-mRNA and mainly U1 and U2. However, with ZDHHC16_-30 there appears to be less U2 present, which may indicate that it is less stably bound and lost during affinity purification. The affinity purified complexes also seem to contain small amounts of U4 and U5 snRNA, and thus could contain some pre-B or B complexes.

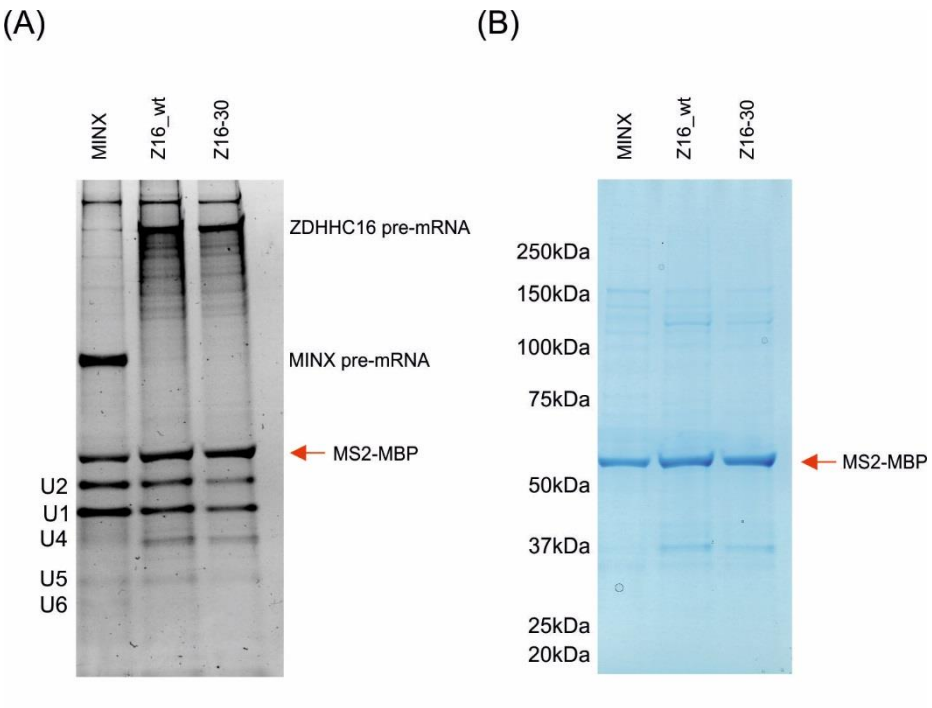


Figure 4.22 –**RNA** (A) **and Protein** (B) **composition of the MS2-selected particles.** The MS2-MBP used for binding to the MS2-loops and the amylose resin is marked (red arrow). The MS2-tagged pre-mRNA used for complex assembly is labeled in the RNA-gel.

The isolated complexes display major differences when analyzed by SDS-PAGE. The two ZDHHC16 complexes contain all protein bands observed in the MINX

control. However, two additional bands appear at an approximate size of 110 kDa and 37 kDa.

Table 4.3 - **Proteins identified by LC-MS/MS in human A complexes isolated on MINX or ZDHHC16 pre-mRNA.** Total peptide counts for each individual protein is shown. The proteins are gouped according to function or association within the spliceosome. Common contaminants such as ribosomal proteins are omitted.

	kDa	human gene name	hA MINX	hA Z16_wt	hA Z16_-30
Sm proteins					
B/B'	24,6	SNRPB	15	9	8
D1	13,3	SNRPD1	8	3	4
D2	13,5	SNRPD2	18	6	8
D3	13,9	SNRPD3	18	6	8
E	10,8	SNRPE	2	2	4
F	9,7	SNRPF	4	1	2
G	8,5	SNRPG	1	2	2
U1 snRNP					
U1-70K	51,6	SNRNP70	21	14	8
U1-A	31,3	SNRPA	16	5	6
U1-C	17,4	SNRPC	1	0	0
U1 snRNP associated					
FBP11	108,8	PRPF40A	7	10	4
S164 (fSAP94)	100,2	RBM25	14	7	10
17S U2 snRNP					
U2A'	28,4	SNRPA1	17	12	11
U2B"	25,5	SNRPB2	6	2	5
SF3a120	88,9	SF3A1	46	24	18
SF3a66	49,3	SF3A2	5	4	1
SF3a60	58,9	SF3A3	18	12	8
SF3b155	145,8	SF3B1	98	75	53
SF3b145	100,2	SF3B2	42	26	20
SF3b130	135,6	SF3B3	151	91	78
SF3b49	44,4	SF3B4	2	2	2
SF3b14a/p14	14,6	SF3B6	10	5	6
SF3b14b	12,4	PHF5A	9	5	4
SF3b10	10,1	SF3B5			
17S U2 related					
tat SF1	85,9	HTATSF1	11	5	2
hPRP43	90,9	DHX15	67	27	31
SPF45	45,0	RBM17	7	5	2
SPF30	26,7	SMNDC1	2	1	0
U2AF65	53,5	U2AF2	7	6	6
U2AF35	27,9	U2AF1	22	15	23
SPF31	29,8	DNAJC8	5	5	5
hPRP5 (DDX46)	117,4	Prp5	16	15	18
SR140	118,3	U2SURP	12	7	12
CHERP	103,7	CHERP	8	5	5
PUF60	59,9	PUF60	32	19	23
SF3b125	103	DDX42	35	13	11
A proteins					
THRAP3 (TRAP150)	108,7	THRAP3	13	9	8
BUB3	37,2	BUB3	19	7	11
FLJ10839, CCAR1	132,8	CCAR1	39	17	25
RBM10	103,5	RBM10	7	2	0
SF1	68,3	SF1	12	4	8
RBM5/LUCA15	92,2	RBM5	14	6	7
SF4 (F23858)	72,5	SUGP1	7	5	0
U5					
220K / hPrp8	273,6	PRPF8	50	19	13
200K / hBrr2	244,5	SNRNP200	88	44	25
116K / hSnu114	109,4	EFTUD2	30	20	18
40K	39,3	SNRNP40			
102K / hPrp6	106,9	PRPF6	25	6	6
15K / hDib1	16,8	TXNL4A	3	0	0
100K / hPrp28	95,6	DDX23	30	8	7
52K / hLin1	37,6	CD2BP2	4	0	0
U4/U6					
90K / hPrp3	77,6	PRPF3	8	4	1
60K / hPrp4	58,5	PRPF4	11	2	0
20K	19,2	PPIH	2	0	0
61K / hPrp31	55,5	PRPF31	7	4	2
15.5K (NHP2L1)	14,2	SNU13			
U4/U6.U5					
110K / hSnu66	90,3	SART1	15	4	1
65K / hSad1	65,4	USP39	6	1	4
B proteins					
hSmu-1 (fSAP57)	57,5	SMU1	20	4	5
RED	65,6	IK	9	2	3
hPRP19/CDC5L complex					
hPRP19	55,2	PRPF19	6	5	2
CDC5L	92,2	CDC5L	6	3	2
CCAP1 (hsp73)	70,4	HSPA8	41	45	44

	kDa	human gene name	hA MINX	hA Z16_wt	hA Z16_-30
hPRP19/CDC5L related					
hSYF1 (XAB2)	100	XAB2	0	16	24
KIAA0560 (fSAP164)	171,3	AQR	6	14	11
A/B proteins					
TCERG1 (CA150)	123,9	TCERG1			
RBM39 (RNPC2, CAPER)	58,5	RBM39	38	26	
TLS/FUS	53,4	FUS	6	4	4
ELAV (HuR)	36,1	ELAVL1	6	14	13
PPP1CA	37,5	PPP1CA	1	1	2
p72/DDX17	80,3	DDX17	16	22	24
DDX9, DHX9	141	DHX9	77	85	82
NFAR	95,4	ILF3	30	25	30
NF45	43,1	ILF2	12	21	14
LOC124245 (NHN1)	106,4	ZC3H18	8	0	0
ASR2B	100	SRRT	13	10	6
p68 (DDX5)	69,2	DDX5	18	26	31
Recruited prior to Bact					
hsp27	22,8	HSPB1	0	2	1
SKIV2L2 (KIAA0052, fSAP118)	117,8	SKIV2L2	8	18	19
Detected in Bact					
p30 DBC (KIAA1967)	102,9	CCAR2	0	22	20
Detected in C complex					
matrin 3	94,6	MATR3	14	122	96
EJC/mRNP					
Acinus (fSAP152)	151,9	ACIN1	17	0	4
UAP56	49	DDX39B	29	25	33
ELG	70,6	NCBP3	5	3	0
DDX3	73,2	DDX3X	15	17	15
mRNA binding proteins					
PABP1	70,7	PABPC1	8	8	9
cap binding complex					
CBP80	91,8	NCBP1	18	19	11
SR proteins					
SF2/ASF	27,8	SRSF1	39	18	24
9G8	27,4	SRSF7	17	7	12
SRp20	19,3	SFRS3	6	5	2
SRp30c	25,5	SRSF9	6	1	2
SRp40	31,3	SFRS5	16	5	2
SRp46	32,3	SRSF8	17	8	4
SRp55	39,6	SRSF6	31	17	12
SRp75	56,7	SRSF4	11	5	2
SC35	25,5	SFRS2	7	4	4
SFRS12	59,4	SREK1	13	7	5
SRSF11	53,5	SRSF11	5	4	4
hTra-2 beta/SFRS10	33,7	SRSF10	0	3	2
hnRNP					
hnRNP A0	30,9	HNRNPA0	6	15	19
hnRNP A1	38,7	HNRNPA1	34	115	145
hnRNP A3	39,6	HNRNPA3	8	28	32
hnRNP AB	36,2	HNRPAB	7	15	20
hnRNP A2/B1	37,4	HNRPA2B1	30	121	134
hnRNP C	33,7	HNRNPC	12	6	10
hnRNP D	38,4	HNRNPD	6	9	25
hnRNP F	45,7	HNRNPF	10	17	19
hnRNP G	42,3	RBMX	1	7	6
hnRNP H1	49,2	HNRNPH1	14	30	29
hnRNP H2	49,3	HNRNPH2	10	18	20
hnRNP H3	36,9	HNRNPH3	0	4	6
hnRNP K	51	HNRNPK	20	52	59
hnRNP L	64,1	HNRNPL	11	27	27
hnRNP L-like		HNRNPLL	0	6	8
hnRNP M	77,5	HNRNPM	14	9	10
hnRNP Q	69,6	SYNCRIP	14	30	34
hnRNP R	70,9	HNRNPR	13	13	20
hnRNP U	90,5	HNRNPU	6	36	40
hnRNP UL2 (SAF-A2)	85,1	HNRPUL2	5	15	17
PCBP1	37,5	PCBP1	17	21	30
PCBP2	38,1	PCBP2	9	11	13
RALY	32,4	RALY	5	0	1
E1B-AP5	95,7	HNRPUL1	5	10	8
Miscellaneous proteins					
IGF2BP3 (Koc1)	63,7	KOC1	7	7	17
Luc7-like 2	46,5	LUC7L2	14	5	4
PTBP1	57,2	PTBP1	5	39	32
FUBP2	73,1	KHSRP	17	5	12

Mass spectrometry of isolated spliceosomal complexes shows that with all three pre-mRNAs predominantly A complexes were isolated, with minor contaminations of B-complex and some B^{act} associated proteins (Table 5.3). The individual complex proteomes show only small differences in the A complex associated proteins, which include U1, U2 snRNP proteins. The most significant difference in the mass-spectrometry data is observed for the matrin-3 protein. The number of peptides for matrin-3 in the ZDHHC16 complexes vastly exceed the observed peptide count for the MINX pre-mRNA complexes (14 vs. 122/96). The protein CCAR2, a core component of the DBIRD complex, is also substantially more abundant in the complex formed on ZDHHC16 pre-mRNA. The DBIRD complex binds to RNA polymerase II and has been linked to alternative splicing (Close et al., 2012) and could be uniquely associated with the ZDHCC16 A complex. Another group of proteins where a significant difference can be detected are the hnRNP proteins. While the positive splice regulators, the SR proteins, can be found with a similar incidence on all three pre-mRNAs, higher peptide counts are seen for the hnRNPs for the ZDHHC16 pre-mRNAs. HnRNP A1, A2/B1 and K are much more abundant on the ZDHHC16 pre-mRNA and other hnRNP proteins are also found at elevated levels. This may be due to the much longer RNA sequence of the ZDHHC16 pre-mRNA compared to MINX. Potential differences between Z16_wt and Z16_-30 include the apparent absence of RBM39 and SF4 (SUGP1).

4.2.5.2 Cross exon constructs

After the initial experiments had shown that splicing of the complete ZDHHC16 construct is not possible under *in vitro* splicing conditions, I redesigned the constructs to be used to study the effects of SF3b1 mutations on BS/ 3' SS selection. The very long intron of ZDHHC16 makes these studies more difficult with cross-intron splicing complexes. Thus, one goal was to shorten the intron. Another was to simplify the analysis of BS/3'SS usage. For this reason, I switched to the cross-exon assembly pathway (see above). This enables the removal of most of the intronic sequence, since only the immediate surrounding area of the branch site and 3' SS are needed for complex assembly.

The cross exon RNA constructs were designed by inserting the 3' end of the respective intron into the MINX exon RNA scaffold. For this the region

encompassing 25nt upstream of the cryptic branch point to the canonical 3'SS AG dinucleotide was inserted into the similarly defined region of the original MINX exon RNA construct, after removing the MINX BS/3'SS nucleotides. Such hybrid constructs were not only designed for ZDHHC16, but also for TMEM14C and ENOSF1. Like ZDHHC16, TMEM14C intron 1 and ENOSF1 intron 7 have both a canonical and a cryptic 3'SS and branch site and *in vivo* it was shown that after mutation of the canonical branch site, splicing required the presence of the mutated SF3b1. These mutations theoretically would allow isolation of *in* vitro complexes containing only the mutated protein, even if the nuclear extract contains wild-type and mutated protein in a one to one ratio.

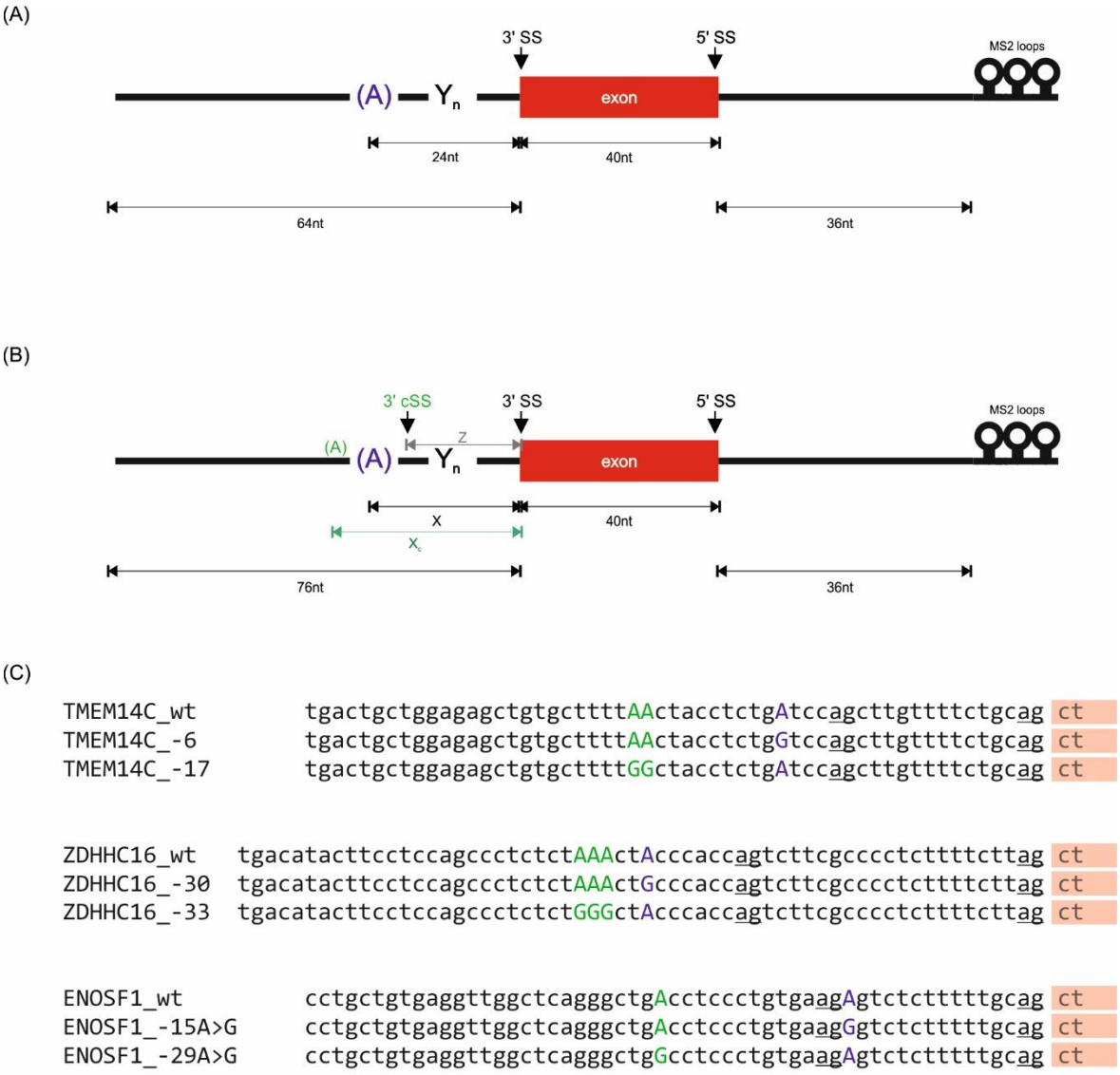


Figure 4.23- **Comparison of the TMEM14C, ZDHHC16 and ENOSF1 exon RNA constructs.** Scheme of the MINX exon RNA (A), the TMEM14C/ ZDHHC16/ ENOSF1 exon constructs (B). For the three new variants of the exon complex RNA the exact distances between 3'SS and branch point (X) and the distance between the

3' SS and the cryptic branch point (X_c) as well as the distance between the two 3' splice sites (Z) is indicated in the table Y. (C) Sequence alignment of the relevant areas on the MINX exon construct replaced by gene specific sequences. Green indicates the cryptic branch point nucleotides.

Table 4.4 – **Distances between different structural elements in the various exon RNA constructs.**

	X	**X_c**	**Z**
TMEM14C exon	20 nt	29 nt	14 nt
ZDHHC16 exon	30 nt	34 nt	21 nt
ENOSF1 exon	15 nt	29 nt	17 nt

4.2.5.2.1 Complex assembly on the ZDHHC16, ENOSF1 and TMEM14C cross exon constructs

The various exon RNAs were then used to analyze the initial assembly steps of cross exon complex formation. I first checked if the assembly of cross exon complexes is possible at all. For this, the RNAs were added to HeLa S10 nuclear extract and exon complex assembly was evaluated by native gel electrophoresis.

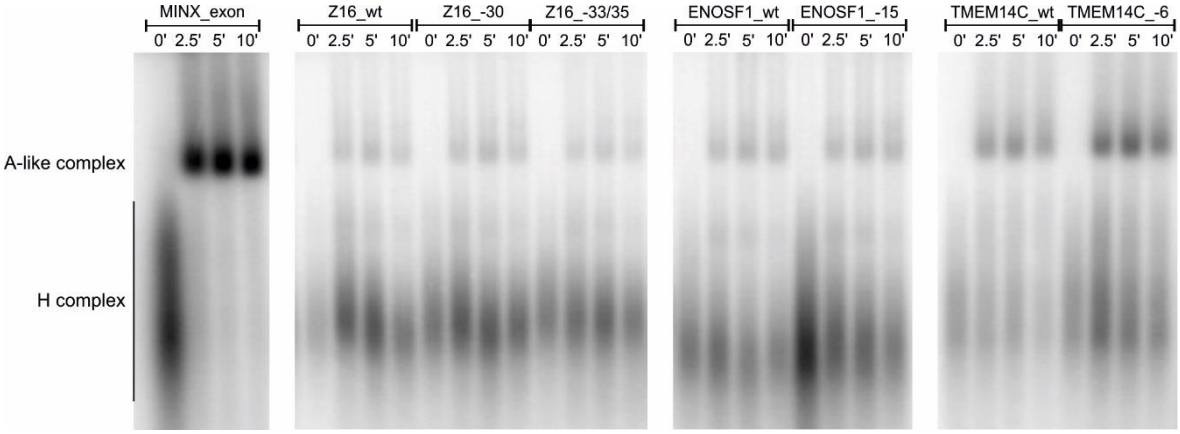


Figure 4.24 - **Native gel electrophoresis of cross exon complexes assembled on different RNAs in HeLa nuclear extract.** ^{32}P RNAs shown above were incubated in HeLa nuclear extract for the indicated times and, after addition of heparin to 0.15µg/µl, separated on a 1.5% agarose gel. Bands were visualized with a PhosphorImager and their identity is indicated on the left. The MINX exon RNA was used as a positive control.

The native gel electrophoresis shows that all exon RNAs are capable of forming exon complexes in HeLa nuclear extract (Figure 5.24). However, the efficiency of A-like complex formation on the ZDHHC16_wt, ENOSF1_wt and TMEM14C_wt RNAs is significantly reduced when compared to the MINX exon construct. In addition the gel also reveals that the substitution of the canonical branch point adenosine has no impact for ZDHHC16 (Z16_-30), or even a positive one for TMEM14C (TMEM14C_-6) on the amount of A-like complex formed. The observation that cross exon complex formation occurs despite removal of the

canonical branch point is unexpected. Published results, based on *in vivo* data, show that the removal of the canonical branch point abolishes splicing of the affected pre-mRNA in absence of the SF3b1^{K700E} mutation (Darman et al., 2015). The substitution of the cryptic branch point (ZDHHC_-33/35) appeared to cause a slight decrease in the amount of exon complex formed. These results are unexpected, since previously published work could show that the removal of the cryptic branch point has no impact on the formation of splice products *in vivo* in the absence of SF3b1^{K700E} (Darman et al., 2015).

4.2.6 Assembly of cross exon complexes in K562 nuclear extract

4.2.6.1 Comparison of MINX exon complexes assembled in K562 nuclear extract

I next wanted to investigate if the presence of SF3b1^{K700E} has an impact on the formation of cross exon complexes. To do this, I first analyzed MINX exon complexes that were assembled in K562 nuclear extract containing FLAG-tagged wild-type SF3b1 or tagged SF3b1^{K700E}.

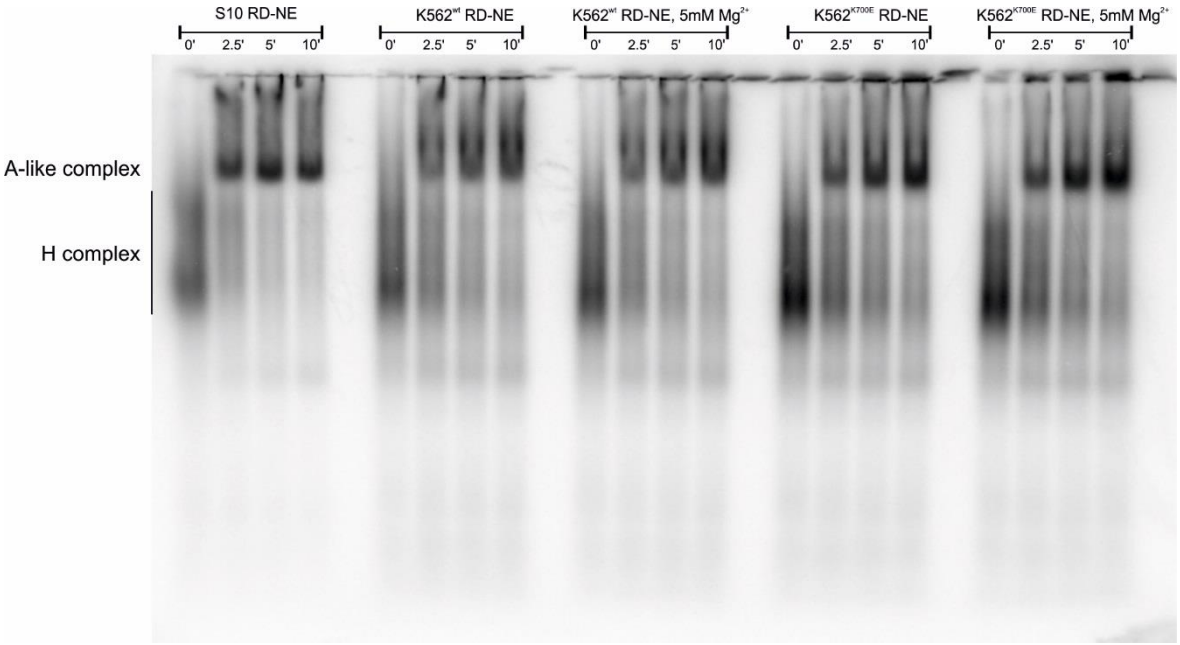


Figure 4.25 - **Native gel electrophoresis of MINX exon complexes assembled in different nuclear extracts.** ^{32}P-MINX exon RNA was incubated in HeLa S10 (S10) nuclear extract and K562 nuclear extract containing only wild-type SF3b1 (K562^{wt}) or wild-type and K700E SF3b1 (K562^{K700E}). For the indicated times and, after addition of heparin to 0.15µg/µl, separated on an 1.5% agarose gel. Bands were visualized with a PhosphorImager and their identity is indicated on the left.

The analysis of the complex formation shows that the MINX A-like exon complex forms in all three nuclear extracts tested. The efficiency of complex formation is highest in the HeLa nuclear extract, but complexes also form in wild-type K562 and

mutant K562 extract. To investigate if the presence of the SF3b1 mutation leads to the assembly of abnormal exon complexes on MINX exon RNA, I isolated MINX exon complexes containing only tagged SF3b1 and analyzed the complex by mass-spectrometry.

To separate the tagged from untagged complexes, the exon complexes formed on MS2-tagged exon RNA were purified by MS2-affinity selection, and then those exon complexes containing FLAG-tagged SF3b1 were separated from the remaining complexes by anti-FLAG affinity chromatography. The purification of MINX exon complexes using the aforementioned double-affinity purification approach allowed the purification of only very small amounts of A-like exon complex (<100 fmol from a 10 ml reaction) which equals ~1% of the input RNA. However, the total yield of the purification was very low for both nuclear extracts used, indicating limited availability of the FLAG-tag in the complex or a low incorporation of the FLAG-tagged SF3b1 protein.

Table 4.5- **Proteins identified by LC-MS/MS in human A-like exon complexes assembled on MINX pre-mRNA and isolated from different K562 nuclear extracts.** Total peptide counts for each individual protein are shown. The proteins are gouped according to function or association within the spliceosome. Common contaminants such as ribosomal proteins are omitted.

	kDa	human gene name	WT-NE	K700E-NE
Sm proteins				
B/B'	24,6	SNRPB	17	14
D1	13,3	SNRPD1	0	2
D2	13,5	SNRPD2	8	13
D3	13,9	SNRPD3	10	8
E	10,8	SNRPE	7	4
U1 snRNP				
U1-70K	51,6	SNRNP70	17	19
U1-A	31,3	SNRPA	4	4
U1-C	17,4	SNRPC	1	0
U1 snRNP associated				
FBP11	108,8	PRPF40A	28	25
S164 (fSAP94)	100,2	RBM25	8	9
17S U2 snRNP				
U2A'	28,4	SNRPA1	20	10
U2B"	25,5	SNRPB2	6	4
SF3a120	88,9	SF3A1	64	27
SF3a66	49,3	SF3A2	4	4
SF3a60	58,9	SF3A3	34	22
SF3b155	145,8	SF3B1	68	67
SF3b145	100,2	SF3B2	42	35
SF3b130	135,6	SF3B3	104	70
SF3b49	44,4	SF3B4	11	11
SF3b14a/p14	14,6	SF3B6	0	5
SF3b14b	12,4	PHF5A	0	4
SF3b10	10,1	SF3B5		
17S U2 related				
tat SF1	85,9	HTATSF1	0	0
hPRP43	90,9	DHX15	7	15
SPF45	45,0	RBM17	1	2
SPF30	26,7	SMNDC1	5	4
U2AF65	53,5	U2AF2	25	20
U2AF35	27,9	U2AF1	6	5
SPF31	29,8	DNAJC8	4	9
hPRP5 (DDX46)	117,4	Prp5	28	43
SR140	118,3	U2SURP	2	5
CHERP	103,7	CHERP	1	2
PUF60	59,9	PUF60	26	27
SF3b125	103	DDX42	12	4
A proteins				
THRAP3 (TRAP150)	108,7	THRAP3		
BUB3	37,2	BUB3	29	26
FLJ10839, CCAR1	132,8	CCAR1	0	0
RBM10	103,5	RBM10	13	0
SF1	68,3	SF1	11	11
RBM5/LUCA15	92,2	RBM5	0	0
ZNF207	50,8	ZNF207	19	13
SF4 (F23858)	72,5	SUGP1	1	1

	kDa	human gene name	WT-NE	K700E-NE
U5				
220K / hPrp8	273,6	PRPF8	26	15
200K / hBrr2	244,5	SNRNP200	20	25
116K / hSnu114	109,4	EFTUD2	35	11
102K / hPrp6	106,9	PRPF6	25	6
100K / hPrp28	95,6	DDX23	11	6
U4/U6				
90K / hPrp3	77,6	PRPF3	16	5
60K / hPrp4	58,5	PRPF4	13	6
20K	19,2	PPIH	0	1
U4/U6.U5				
110K / hSnu66	90,3	SART1	26	5
65K / hSad1	65,4	USP39	4	3
B proteins				
hSmu-1 (fSAP57)	57,5	SMU1	16	12
RED	65,6	IK	4	2
hPRP19/CDC5L complex				
CCAP1 (hsp73)	70,4	HSPA8	26	23
A/B proteins				
RBM39 (RNPC2, CAPER)	58,5	RBM39	34	35
p72/DDX17	80,3	DDX17	14	0
DDX9, DHX9	141	DHX9	11	6
p68 (DDX5)	69,2	DDX5	16	11
SR proteins				
SF2/ASF	27,8	SRSF1	26	33
9G8	27,4	SRSF7	16	12
SRp20	19,3	SFRS3	4	0
SC35	25,5	SFRS2	1	3
SRSF11	53,5	SRSF11	6	5
SR related proteins				
SRm300	300	SRRM2	11	17
hnRNP				
hnRNP A0	30,9	HNRNPA0	2	2
hnRNP A1	38,7	HNRNPA1	18	17
hnRNP A3	39,6	HNRNPA3	4	6
hnRNP AB	36,2	HNRPAB	0	1
hnRNP A2/B1	37,4	HNRPA2B1	18	16
hnRNP C	33,7	HNRNPC	4	5
hnRNP D	38,4	HNRNPD	1	2
hnRNP H1	49,2	HNRNPH1	1	0
hnRNP K	51	HNRNPK	1	3
hnRNP L	64,1	HNRNPL	0	1
hnRNP U	90,5	HNRNPU	2	3
PCBP1	37,5	PCBP1	22	23
Miscellaneous proteins				
Luc7-like 2	46,5	LUC7L2	4	3
FUBP2	73,1	KHSRP	18	24

The evaluation of the MS results for the MINX exon complexes assembled in the SF3b1^{wt} and SF3b1^{K700E} K562 nuclear extract shows that both complexes have a similar protein composition (Table 5.5). The amount of tri-snRNP associated proteins present in the exon complexes formed in the SF3b1^{K700E} extract versus the SF3b1^{wt} containing extract appears generally lower, suggesting less tri-snRNP association. One other interesting protein that is less abundant, based on the number of peptides sequenced, in the exon complexes assembled in SF3b1^{K700E} extract is DDX17, an RNA helicase shown to be involved in the process of splice site selection (Dardenne et al., 2014). Taken together these results suggest, that while the presence of SF3b1^{K700E} does not cause major changes in the proteome of

the exon complex minor it could potentially affect the association and binding of auxiliary factors that could play a role in the splice site selection during the assembly of the spliceosome. However, additional experiments are needed to show that these factors are reproducibly affected.

4.2.6.2 Formation of ZDHHC16, ENOSF1 and TMEM14C exon complexes in K562 nuclear extract

After showing that the K562 nuclear extract supports cross exon complex formation on the MINX exon RNA, I next tested for complex formation in K562 nuclear extract with ZDHHC16. ENOSF1 and TMEM14C exon RNAs, including versions where either the canonical or the cryptic branch point adenosines was mutated. As shown in figure 5.26, formation of an A-like complex was detected after 2.5 minutes in the wild-type and mutant K562 extract with the MINX exon RNA, but was much less efficient with the other RNAs.

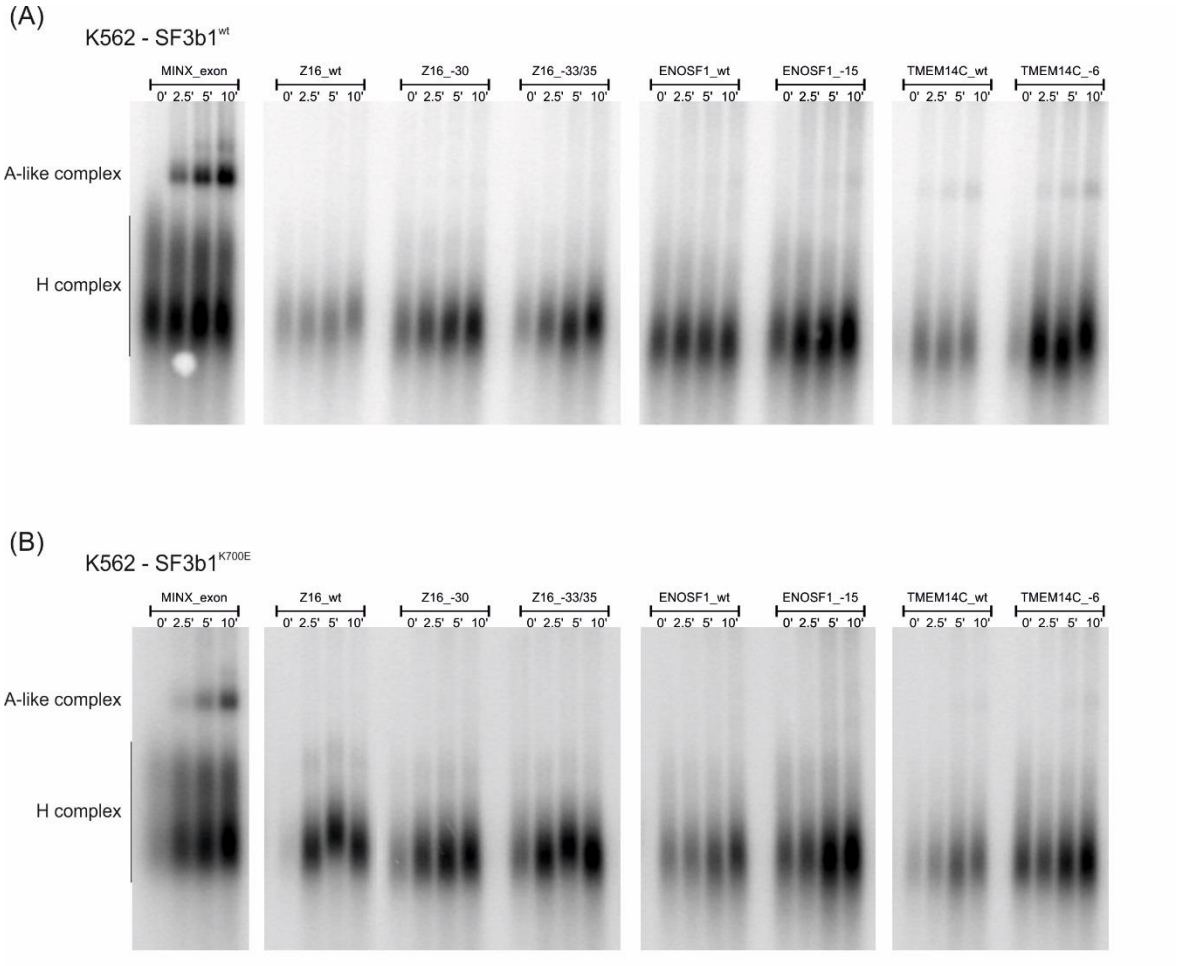


Figure 4.26 - **Native gel electrophoresis of cross exon complexes assembled on different RNAs.** ^{32}P RNAs shown above were incubated in K562 nuclear extract, either wild-type (A) or containing SF3b1^{K700E} (B), for the indicated times. After addition of heparin to 0.15µg/µl, complexes were separated on a 1.5% agarose gel. Bands were visualized with a PhosphorImager and their identity is indicated on the left. The MINX exon RNA was used as a positive control.

The native gel electrophoresis of cross exon complexes assembled in K562 nuclear extract obtained from cell lines expressing only wild-type SF3b1 (Figure 5.26a) or wild-type and SF3b1^{K700E} (Figure 5.26b) shows that the general efficiency of complex assembly decreased when compared to complex formation in HeLa S10 nuclear extract (compare with figure 5.24). The overall decrease in complex formation is even more pronounced in the extract containing mutated SF3b1 than it is in the wild-type K562 nuclear extract. For the three RNAs of interest, complex assembly does occur, but with a very low efficiency. The most efficient complex assembly occurs on TMEM14C_wt and TMEM14C_-6. This observation is in line with the previous result that in HeLa nuclear extract TMEM14C derived RNAs showed the largest amount of complex formation. This experiment shows that the presence of SF3b1^{K700E} in the nuclear extract impairs the formation of spliceosomal complexes in general.

4.2.7 TMEM14C complexes assembled in HeLa S10 nuclear extract

Due to the extremely low efficiency of complex assembly in K562 nuclear extract, I initially used HeLa cell nuclear extract to investigate the differences between complexes assembled on the different branch point mutations. The TMEM14C RNA was used as a substrate, because the complex assembly appeared to be slightly more efficient than with the ZDHHC16 and ENOSF1 RNA substrates. Analysis of complex formation on native agarose gels showed that the RNA substrate containing a mutation of the cryptic branch point adenosine (T14_-17) displays a decreased level of complex assembly when compared to the other to substrates, i.e. wild-type (T14_wt) and mutated canonical branch point (T14_-6) (Figure 5.27a). This result was unexpected, since the cryptic branch point of TMEM14C should be irrelevant for splicing in nuclear extract containing only wild-type SF3b1. In order to further investigate potential reasons for the decreased complex assembly, I isolated A-like complexes assembled on TMEM14C exon RNAs.

Five milliliter splice reactions were performed in the presence of 1nM radiolabeled exon RNA. Afterwards the assembled complexes were purified using MS2-selection and then separated by gradient centrifugation. The gradient peaks were then analyzed for their RNA content by SYBR gold staining and the protein composition was evaluated by MS.

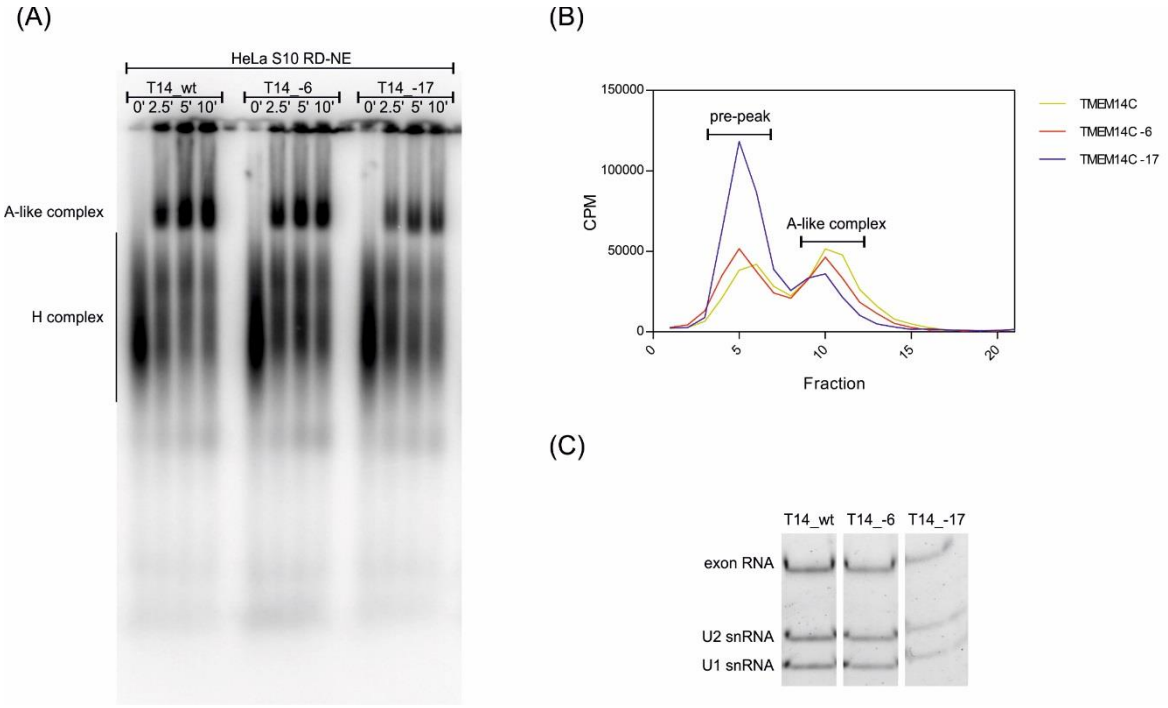


Figure 4.27- **Assembly of A-like exon complexes using TMEM14 variants in HeLa S10 nuclear extract.** (A) ^{32}P RNAs shown above were incubated in HeLa nuclear extract for the indicated times and, after addition of heparin to 0.15µg/µl, separated on a 1.5% agarose gel. Bands were visualized with a PhosphorImager and their identity is indicated on the left. (B) MS2-purified complexes were analyzed on a 10-30% (v/v) glycerol gradients prepared with 75mM KCl. Gradients were fractionated by hand and CPM values were determined by Cherenkow counting. (C) RNA composition of purified cross-exon A-like complexes.

The gradient centrifugation confirms the previous result from the native gel. That is, the substitution of the canonical branch point causes a smaller decrease in the amount of exon complex assembly than the substitution of the cryptic branch point nucleotide (Figure 5.27b). On the level of snRNA, no difference can be found. All purified cross exon complexes contained roughly equimolar amounts of U1 and U2 snRNA (Figure 5.27c). The RNA composition and the position in the gradient both indicate that the complex formed is indeed a cross exon A-like complex. The peak fractions were then pooled and analyzed by MS after SDS-PAGE.

Table 4.6 - **Proteins identified by LC-MS/MS in A-like exon complexes isolated on different TMEM14C RNA constructs from HeLa nuclear extract.** Total peptide counts for each individual protein is shown. The proteins are gouped according to function or association within the spliceosome. Common contaminants such as ribosomal proteins are omitted.

	kDa	human gene name	T14_wt	T14-6	T14_-17
Sm proteins					
B/B'	24,6	SNRPB	8	14	6
D1	13,3	SNRPD1	2	1	2
D2	13,5	SNRPD2	5	10	3
D3	13,9	SNRPD3	5	6	3
E	10,8	SNRPE	0	1	3
U1 snRNP					
U1-70K	51,6	SNRNP70	7	9	5
U1-A	31,3	SNRPA	5	3	6
U1-C	17,4	SNRPC	1	0	0
U1 snRNP associated					
FBP11	108,8	PRPF40A	14	23	7
S164 (fSAP94)	100,2	RBM25	10	10	9
17S U2 snRNP					
U2A'	28,4	SNRPA1	6	7	10
U2B"	25,5	SNRPB2	5	6	0
SF3a120	88,9	SF3A1	16	26	17
SF3a66	49,3	SF3A2	4	5	0
SF3a60	58,9	SF3A3	11	11	6
SF3b155	145,8	SF3B1	39	51	25
SF3b145	100,2	SF3B2	17	30	13
SF3b130	135,6	SF3B3	53	69	46
SF3b49	44,4	SF3B4	3	10	2
SF3b14a/p14	14,6	SF3B6	5	9	10
SF3b14b	12,4	PHF5A	1	2	3
SF3b10	10,1	SF3B5			
17S U2 related					
tat SF1	85,9	HTATSF1	5	8	0
hPRP43	90,9	DHX15	5	5	2
SPF45	45,0	RBM17	2	0	0
SPF30	26,7	SMNDC1	1	1	0
U2AF65	53,5	U2AF2	7	10	3
U2AF35	27,9	U2AF1	3	5	3
SPF31	29,8	DNAJC8	7	6	2
hPRP5 (DDX46)	117,4	Prp5	20	43	22
SR140	118,3	U2SURP	3	1	0
CHERP	103,7	CHERP	3	2	0
PUF60	59,9	PUF60	20	20	11
SF3b125	103	DDX42	0	0	0

	kDa	human gene name	T14_wt	T14-6	T14_-17
A proteins					
BUB3	37,2	BUB3	8	11	3
SF1	68,3	SF1	15	21	0
RBM5/LUCA15	92,2	RBM5	2	3	0
ZNF207	50,8	ZNF207	5	9	2
SF4 (F23858)	72,5	SUGP1	3	6	2
U5					
200K / hBrr2	244,5	SNRNP200	5	2	5
hPRP19/CDC5L complex					
CCAP1 (hsp73)	70,4	HSPA8	9	20	32
A/B proteins					
RBM39 (RNPC2, CAPER)	58,5	RBM39	11	18	10
p72/DDX17	80,3	DDX17	0	0	3
DDX9, DHX9	141	DHX9	8	16	0
NFAR	95,4	ILF3	2	1	0
p68 (DDX5)	69,2	DDX5	5	6	0
SR proteins					
SF2/ASF	27,8	SRSF1	17	21	11
9G8	27,4	SRSF7	8	5	0
SC35	25,5	SFRS2	2	1	0
SRSF11	53,5	SRSF11	2	0	0
SR related proteins					
SRm300	300	SRRM2	13	15	14
hnRNP					
hnRNP A0	30,9	HNRNPA0	2	1	0
hnRNP A1	38,7	HNRNPA1	5	10	2
hnRNP AB	36,2	HNRPAB	0	1	0
hnRNP A2/B1	37,4	HNRPA2B1	5	8	16
hnRNP D	38,4	HNRNPD	0	1	0
hnRNP K	51	HNRNPK	1	0	0
hnRNP Q	69,6	SYNCRIP	0	1	0
hnRNP R	70,9	HNRNPR	2	1	0
hnRNP U	90,5	HNRNPU	2	2	2
Miscellaneous proteins					
Luc7-like 2	46,5	LUC7L2	4	7	0

The MS analysis of the three isolated TMEM14C exon complexes shows only minor differences between the three complexes. All three complexes contain the U1 snRNP and U2 snRNP associated proteins, which one would expect to find in a cross-exon A-like complex. The only apparent differences are the absence of SF1, DDX9 and several additional A complex proteins in the exon complex assembled on the RNA containing a mutated cryptic branch point (T14_-17). However due to the low peptide counts, in some cases it is not clear if there is a difference in the amount of a certain protein present.

4.2.8 Transition of MINX exon complexes into B-like complexes in the presence of SF3b1^{K700E}

Next, I wanted to investigate if the presence of SF3b1^{K700E} has a negative impact on the ability of the exon complexes to transition into a B-like complex. For this, I performed complex assembly with the MINX exon RNA in K562-derived nuclear extracts containing either only SF3b1^{wt} or SF3b1^{wt} and SF3b1^{K700E}. Formation of the A-like and B-like complexes was evaluated using native gel electrophoresis. After

allowing A-like complexes to form, I added the 5’ SS oligonucleotide to investigate if the transition from the A-like complex into a B-like complex is hindered. For this, a 100-fold molar excess of 5’SS oligonucleotide over the radiolabeled MINX exon RNA was added to the reaction after 5 minutes and incubated at 30°C for an additional 5 minutes.

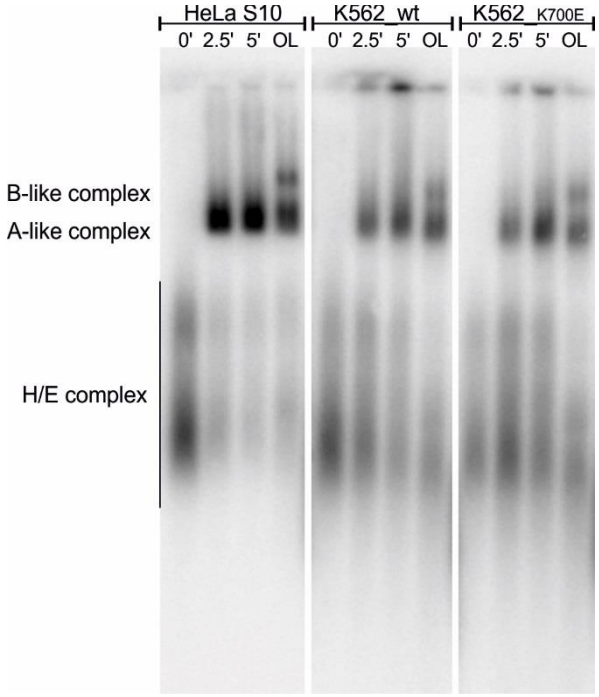


Figure 4.28- **A-like to B-like transition of MINX exon complexes after oligo treatment.** Exon complexes were assembled in three different nuclear extracts and incubated for the time indicated above the gel. Lanes labeled with “OL” were treated with the 5’ splice site oligonucleotide. Assembled complexes were visualized using PhosphorImager screens and their identity is indicated on the left side.

The assembly of cross exon A-like complexes is observed in all nuclear extracts that were tested (Figure 5.28). After addition of a 5’SS oligonucleotide the formation of a B-like complex can be observed in all three nuclear extracts. The transition to the B-like complex occurs with a similar efficiency in both K562 derived nuclear extracts. The HeLa S10 control extract shows a higher amount of A-like complex formation and subsequently also an increased total amount of B-like complex. However, the relative amount of exon B-like complex formed remains similar independent of the nuclear extract used. This suggests that the K700E mutation does not inhibit tri-snRNP recruitment. However, it is presently not clear whether the A-like complexes generated contain substantial amounts of the mutant SF3b1 protein. In the future, additional experiments characterizing the protein composition of the exon

complexes formed, will give insight into whether the SF3b1^{K700E} protein is abundantly present within the complex.

4.2.9 Consequences of the K700E mutation on the proteome of the K562 cell

In parallel to the investigation of direct impacts of the SF3b1 amino acid substitution on the splicing of specific pre-mRNAs I also wanted to check, if the mutation leads to changes in the proteome of the cell. Several studies could show that SF3b1 and its misregulation leads to widespread changes within the cell, that affect multiple, not directly splicing related, processes. In order to evaluate changes in the proteome of the K562 cells expressing the mutated protein, TMT-labeling experiments were performed in cooperation with Dr. Kuan-Ting Pan in the department of Bioanalytical Mass Spectrometry at our Institute. The samples used for the analysis were labeled with isobaric molecules and quantitatively analyzed by mass spectrometry.

A comparison of the cells containing the SF3b1^{K700E} with the wild-type K562 cells revealed quantitative differences in the amounts of many proteins. The presence of SF3b1^{K700E} affects the expression of over 2000 proteins that are involved in a variety of cellular processes (Figure 5.29).

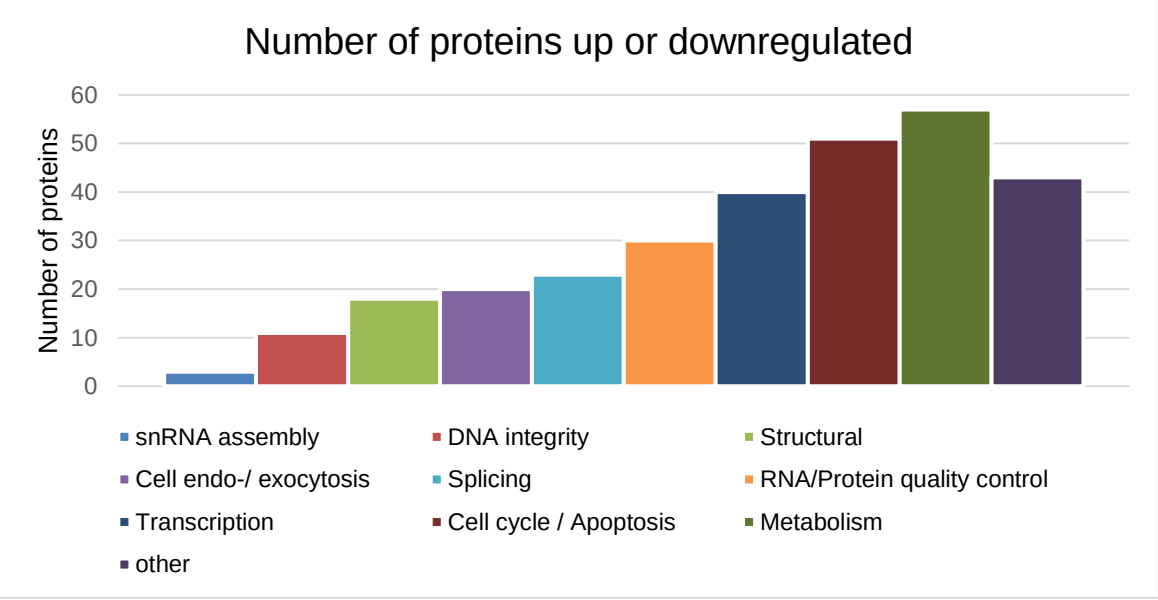


Figure 4.29 **– Expression of multiple proteins changes as a result of SF3b1^{K700E} expression.** Many proteins are up or downregulated in their abundance as a direct or indirect consequence of SF3b1^{K700E} expression. Significantly changed proteins were grouped according to their biological function.

Amongst others, notable changes were seen in the down-regulation of some spliceosomal proteins. The highest reduction in protein levels was observed for the

tri-snRNP associated protein USP39. The amount of USP39 detected in the SF3b1^{K700E} cells is reduced by approximately 50%. The second most down regulated protein is the alternative splicing factor TRA2B (-45%). Several other spliceosome associated proteins are also found among those proteins downregulated in the cells expressing mutated SF3b1 in addition to the wild-type isoform. These include GPKOW (-40%), SUGP2 (-35%), SF3A2 (-30%) and U2SURP (-25%). Splicing associated proteins were also upregulated, including CLNS1A (+70%), RBM10 (+50%), AQR (+35%) as well as the hnRNPs A1 and UL1 (+47% and +41%). Taken together the proteome analysis showed widespread changes in the cellular proteome as a consequence of the amino acid substitution in SF3b1, including changes in the levels of multiple splicing and splicing regulation associated proteins. However, most changes were moderate and might reflect the presence of large amounts of the wild-type protein that is still present in the K562^{K700E} cells.

5 Discussion

5.1 Pladienolide B stalled A complexes

Pladienolide B is a natural small molecule macrolide compound that inhibits cancer cell growth and stalls splicing by binding to SF3b1 (Kotake et al., 2007). The binding of Pladienolide B is thought to block the HEAT domain of SF3b1 in an open conformation and prevents the correct binding of the branch point adenosine (Cretu et al., 2018). Pladienolide B and the bound branch point adenosine appears to compete for accommodation by the same region of the SF3b1 protein. In the current model, if the branch point adenosine is bound correctly, the HEAT domain switches to a closed conformation, stabilizing the formed U2/branch point helix during A complex formation (Cretu et al., 2018). However, the exact mechanism by which Pladienolide B stalls the splicing of pre-mRNA still remains to be elucidated.

5.1.1 Pladienolide B blocks docking of the tri-snRNP to the spliceosomal A complex

The interactions of Pladienolide B with the HEAT domain of SF3b1 were proposed to arrest the protein in an open conformation (Cretu et al., 2018) and thereby might interfere with the stable interaction between U2 and the branch point sequence. In a first set of experiments, I set out to dissect the consequences that Pladienolide B has on the composition and structure of the spliceosome. To do so, I initially investigated at which stage the presence of Pladienolide B stalls the spliceosome assembly by analyzing cross-intron spliceosome complexes. First, using native agarose gel electrophoresis I checked at which stage the spliceosome assembly is stalled in the presence of different Pladienolide B concentrations and also how stable the formed complexes are by performing a heparin titration. Heparin disrupts weak and non-specific interactions between proteins in a concentration-dependent manner. At a Pladienolide B concentration of 250nM splicing catalysis was inhibited completely and only A complexes were detected on native agarose gels. This observation is consistent with previous investigations of the influence of Pladienolide B on splicing (Effenberger et al., 2014b; Kotake et al., 2007). By performing a heparin titration, I could show that the A complex formed in the presence of Pladienolide B is less stable compared to the wild-type A complex. The A complex containing Pladienolide B dissociates at a heparin concentration of 0.2 µg/µl, while

the uninhibited A complex remains stable up to a heparin concentration of 0.6 µg/µl. The decreased stability of Pladienolide B stalled A complexes could indicate destabilizing effect of Pladienolide B on the U2/BPS interaction. This could directly affect splice site selection processes, by favoring strong branch point sequences over weaker ones, leading to changes in cellular splicing patterns in the presence of the inhibitor. This idea is supported by *in vivo* observations, which show that Pladienolide B modulates pre-mRNA splicing rather than inhibiting it completely. The resulting changes in the alternative splicing patterns, in particular those affecting cell proliferation, could also explain its anti-tumor activity (Effenberger et al., 2014a).

When the assembly of spliceosomal complexes in the presence and absence of Pladienolide B was analyzed by gradient centrifugation, a less stringent method to evaluate complex formation, I could show that the presence of Pladienolide B prevents even the initial docking of the tri-snRNP. This means that the presence of Pladienolide B prevents the formation of a pre-B complex, where the tri-snRNP is only loosely bound (Boesler et al., 2016). Similar observations were made when I evaluated the influence of Pladienolide B on the formation of cross exon complexes. The A-like cross exon complex is formed even in the presence of high concentrations of Pladienolide B (250 nM), but it could not be converted into a B-like cross exon complex with the tri-snRNP stably bound. Thus, like other SF3b1 inhibitors, including SSA and the pladienolide B derivate E7107, splicing is inhibited at a very early assembly stage, after U2 binding. The initial docking of the tri-snRNP during formation of the pre-B complex involves the formation of the U2/U6 helix II (Boesler et al., 2016), which consists of base pairing between the 5’ end of U2 and 3’ end of U6. The recent cryo-EM structures of the human pre-B complex show that the main interface between U2 and the tri-snRNP forms via interactions between SF3b and the U6 LSm-ring at the very end of U6 and the U4 Sm-ring (Charenton et al., 2019; Zhan et al., 2018). During the B complex formation, SF3b proteins, especially SF3b3, also play an important role in the bridging of U2 and U5. If now the SF3b1 protein, which is a scaffolding protein upon which most other SF3b proteins and various other spliceosomal proteins bind, is prevented from obtaining its correct structure due to the binding of Pladienolide B, then this could prevent the interactions important for the stable integration of the tri-snRNP.

5.1.2 Pladienolide B remains bound to SF3b1 throughout A complex formation

Even though Pladienolide B stalled A complexes were less stable, it was still possible to purify them using the established MS2 affinity purification of the AG Lührmann. To evaluate if Pladienolide B remains present in the spliceosomal complexes throughout the MS2 purification procedure I performed chase experiments with SF3b depleted nuclear extracts. The results of these experiments were consistent with the idea that Pladienolide B remains bound within the formed A complexes, since purified A complexes assembled in the presence of Pladienolide B remained inhibited after adding nuclear extract lacking the inhibitor, if reassembly of spliceosomes was prevented. However, it could also be that the presence of Pladienolide B in the early stages irreversibly damages the building blocks of the A complex making progression of splicing impossible even if Pladienolide B itself is not present anymore. Thus, the observations made in the chase experiments do not definitively prove that Pladienolide B remains present in the spliceosomal A complex. To confirm these experiments with radiolabeled or fluorescence labeled Pladienolide B would have to be conducted.

5.1.3 Changes in the protein composition of purified Pladienolide B stalled A complexes

Mass spectrometry of the wild-type and the Pladienolide B stalled A complexes showed that the presence of Pladienolide B does not affect the abundance of proteins associated with the stable binding of the U2 snRNP to the branch point region, such as the SF3a and SF3b complexes. This observation is consistent with the data indicating that the presence of Pladienolide B does not impair the formation of the 17S U2 particle itself, which means the U2 particles are not dissociated upon addition of the inhibitor. If Pladienolide B had an effect during the formation of the 17S U2 particle itself, one would expect the formation of the A complex to be inhibited as well, which is not the case. Previous studies showed that the Pladienolide derivative E7107, also had no effect on the protein composition of the 17S U2 snRNP (Folco et al., 2011). Unexpectedly, hPrp5, which is important for the stable U2 binding during A complex formation (see below), was not detected in the wild-type A complexes; normally, low amounts of hPrp5 are detected (see for example table 5.1- MINX vs. ZDHHC16 A complex) in the A complex. hPrp5 was also not found in Pladienolide B stalled A complexes, but as hPrp5 was not detected

in the wild-type complexes either, it remains unclear whether it is specifically absent in the Pladienolide B A complexes or was just not detected in this particular experiment. Experiments with other inhibitors stalling spliceosome assembly at the stage of an A complex, such as the previously mentioned E7107, did not cause a reduction in the amount of hPrp5 detected in the 17S U2 snRNP (Folco et al., 2011).

On the other hand, mass spectrometry revealed for the first time that there are differences in certain proteins associated with the wild-type versus Pladienolide B A complexes. One intriguing difference is the accumulation of DDX42 (SF3b125) in Pladienolide B stalled A complexes. DDX42 is a DEAD-box, ATP-dependent RNA helicase, with a currently unclear role in pre-mRNA splicing. It is found in submolar amounts in purified SF3b but very little is found in purified 17S U2 particles (Will and Lührmann, 2002) and it is normally also not found in the human A complex (Agafonov et al., 2011). While the exact function of DDX42 in the process of splicing remains unclear, it was originally proposed to play a role in the formation of the 17S U2 snRNP (Will and Lührmann, 2002). DDX42 has been shown to be involved in both the unwinding of double-stranded RNA and also the annealing of single stranded RNA molecules, even if the involved RNA molecules are partially obstructed by RNA binding proteins (Uhlmann-Schiffler, 2006). In addition DDX42 was recently implicated to be involved in the process of apoptosis (Sohn and Chay, 2019) and was also proposed as a cellular sensor for double stranded RNA molecules (Lin et al., 2008).

The presence of DDX42 in the stalled A complex could indicate that certain maturation steps or conformational changes that under normal circumstances lead to the loss of DDX42 prior to the formation of the A complex do not occur in the presence of Pladienolide B. Although the binding of Pladienolide B was shown not to change the three dimensional structure of the SF3b complex (Cretu et al., 2016), it is possible that its binding prevents subsequent structural changes, during the formation of the 17S U2 particle or the A complex from occurring and that these normally lead to the release of DDX42. In the future, it will be interesting to see if 17S U2 particles purified in the presence of Pladienolide B bind more DDX42 than particles isolated in the absence of Pladienolide B. In summary, the accumulation of DDX42 in the Pladienolide B stalled A complexes is most likely caused by an

inhibitor related impairment of the DDX42 release, but it can only be speculated if the persistence of the helicase has any functional implications for the spliceosome itself, or if it is just a side effect of the stalling brought about by the presence of Pladienolide B.

A second interesting difference between the two analyzed A complexes is the absence of SPF30 in the Pladienolide B stalled A complex. SPF30 is a Tudor-domain containing protein and is essential for the formation of B complexes – Immunodepletion of SPF30 shows stalling of spliceosome assembly at the stage of an A complex, meaning that it is required for the integration of the tri-snRNP into the complex (Meister et al., 2001; Rappsilber et al., 2001). SPF30 is present in the 17S U2 snRNP (Rappsilber et al., 2001; Will et al., 2001) and also interacts with hPrp3 (a U4/U6 and tri-snRNP associated protein) (Little and Jurica, 2008; Meister et al., 2001). This means the SPF30 protein most likely bridges the U2 snRNP and the tri-snRNP in the spliceosome and thereby helps the latter to be recruited to the A complex. While SPF30 is found in the A complex and small amounts can still be detected in the pre-B complex, it is no longer present in the spliceosomal B complex (Agafonov et al., 2011; Boesler et al., 2016). This would indicate that SPF30 aids in the recruitment of the tri-snRNP but is then released upon binding and stable integration. One currently unanswered question is if the presence of Pladienolide B leads to the dissociation of SPF30 from the 17S U2 particle already, or if this takes place only during A complex formation.

Not only Pladienolide B, but also many other natural inhibitors of splicing stall the assembly of complexes at the stage of an A complex (Effenberger et al., 2017). While little is known about the protein composition of the complexes that assemble in the presence of Spliceostatin A, E7107 and other SF3b1 inhibitors (Corrionero et al., 2011; Folco et al., 2011), all of these compounds appear to interfere with the recruitment and integration of the tri-snRNP. In the future it would be interesting to compare the protein composition of A complexes stalled with different compounds, because if the changes observed are similar, it would indicate a common mechanism of action.

5.1.4 Pladienolide B and the structure of the human A complex

Pladienolide B binds to SF3b1 and could prevent structural changes in the protein, which are required for the stabilization of the U2/ BPS interactions. Currently it is assumed that the stable integration of the U2 particle that has to occur during the transition from the E to the A complex, revolves around the proposed clamping mechanism of the SF3b1 HEAT domain. However, due to the absence of a published high resolution structure of the human A complex and thus the poor resolution of the U2 snRNP, and SF3b1 in particular, this proposed mechanism has yet to be confirmed. Very recently, a cryo-EM structure of the yeast A complex was published (Plaschka et al., 2018); however, the resolution in the region of the U2 snRNP was limited to 6-10 Å and so the precise structure of the SF3b1 HEAT domain was unclear. However, the high resolution structure of the B^{act} complex shows that the HEAT domain adopts a closed conformation, compared to the structure in the SF3b complex (Cretu et al., 2018). The HEAT domain forms a superhelical solenoid structure and the U2/ BPS helix is clamped between its N- and C-terminal HEAT repeat (Haselbach et al., 2018; Zhan et al., 2018). Although the resolution is not as high as in B^{act}, the SF3b1 HEAT domain seems to adopt a closed conformation in the human B complex as well (Bertram et al., 2017a; Zhan et al., 2018).

One aim of my work was therefore to perform a structural analysis of the human A complex in the presence and absence of Pladienolide B. I chose to perform cryo-EM with the hope to obtain structures of the A complex at a resolution that would allow me to see conformational changes due to the presence of Pladienolide B. For this, the standard MS2-purification protocol of the department was optimized for the purification of A complexes. During these investigations, multiple salt concentrations were tested for both initial spliceosome assembly conditions, as well as for the subsequent gradient centrifugations. Interestingly, neither changes in the salt concentration of the reaction (65 mM KCl to 37.5 mM KCl) nor changes in the salt concentration used during the subsequent gradient centrifugation (150 mM to 37.5 mM KCl) affected the ability of the nuclear extract to form A complexes, nor did it change the protein composition of the purified complex (data not shown). The salt conditions did change the sedimentation pattern of the complex, with lower salt concentrations leading to complexes with higher sedimentation values. This change

in the sedimentation value would indicate a changed complex morphology. However, the analysis of the purified spliceosomes by negative stain EM, revealed massive aggregation of the individual spliceosomes. The aggregated spliceosomes themselves were not visibly changed from the entities purified at higher salt concentrations. Using different quenching conditions after cross-linking particles with glutaraldehyde also did not improve the EM images. All of our cryo-EM preparations, however, failed to produce datasets that yielded usable 2D classes. The failure to obtain datasets that ultimately produce a three dimensional structure of the human A complex can be due to various reasons. One possible explanation are errors in the sample preparations that, while not visible on the biochemical level, might destroy the particle prior to data acquisition. Initial experiments could show that the A complex is a fairly sensitive complex and that incorrect blotting conditions lead to a flattened particle that does not yield 2D classes. Alternatively it is possible that the A complex itself is a highly flexible particle that naturally exists in many different conformations at the same time. In this case, much larger data sets would be required to compute a three dimensional structure of sufficient quality. As an alternative to cryo-EM, biochemical methods might in the future also reveal structural differences between the wild-type and the Pladienolide B stalled A complex. For example, cross-liking studies could give further insight into potential changes in the interactome of the complexes due to the presence of the inhibitor.

5.2 Cancer related mutations of SF3b1

The cancer-linked hotspot mutations found in SF3b1 predominantly cluster within the HEAT domain of the protein with most found in HEAT repeats 2-9 (Bonnal et al., 2012), SF3b1 amino acids mutated in cancer cells are important for the tertiary structure of SF3b1 (Cretu et al., 2016); thus these mutations have been proposed to alter the structure of the HEAT domain. The SF3b1 HEAT domain is a crucial interaction platform within the SF3b subcomplex and the larger 17S U2 snRNP (Cretu et al., 2016). For example, PHF5A contacts the concave surface of the HEAT domain both at the N- and C-terminus of the superhelix and it together with SF3b1 forms the binding pocket for the BPS adenosine. Given that mutations of SF3b1 are directly connected with alterations in branch point and 3’ splice site usage (Darman

et al., 2015), I wanted to investigate the consequences of these mutations on the structure and function of the 17S U2 snRNP. To do so, I focused on the K700E amino acid substitution, which is the one of the most common cancer-linked mutations found in SF3b1 (Yoshida and Ogawa, 2014). The K700 amino acid is located in HEAT repeat seven, in close proximity to where the pre-mRNA emerges from the inner area of the HEAT domain (Kastner et al., 2019). In addition the HEAT repeat seven is a part of the interaction surface between SF3b1 and U2AF65 (Cretu et al., 2016).

Multiple models have been proposed to explain the changes in splice site selection observed upon emergence of the K700E mutation; they are largely based on the transcriptomic data obtained from patients and *in vivo* experiments. One possible mechanism is based on the observation that the mutation of the K700E amino acid is located right at the “hinge” between two HEAT-repeats of SF3b1. This substitution could lead to a change in the three-dimensional structure of the protein, forcing it into a more open conformation, as compared to the wild-type SF3b1. This open conformation could impair SF3b1 interactions with the canonical branch sites on the pre-mRNA and instead promote the recognition of the cryptic branch site, which on average has a slightly higher affinity for the U2 snRNA than its canonical counterpart. Alternatively, the presence of the mutation could change the positioning of the pre-mRNA within the spliceosome thus changing its interaction with the spliceosomal components. Here the spliceosome assembly on cryptic splice sites is possible in the presence of the mutation, whereas in case of the wild-type SF3b1 protein steric obstacles prevent the assembly of splicing-competent spliceosomes. The discussed theories provide possible explanations for the mechanism of action of SF3b1 mutations; however, because the currently available data represents only endpoint observations, the actual changes in splicing leading to the K700E-connected phenotypes still remain to be elucidated.

5.2.1 The SF3b1 K700E mutation alters the composition of the 17S U2 snRNP

A better understanding of how cancer-related mutations affect the normal function of SF3b and the U2 snRNP is important for improving our understanding of the onset of MDS and cancer. It may also be useful for developing better therapeutic approaches. To investigate the consequences of the K700E mutation, I utilized a

K562 cell line expressing SF3b1^{K700E} from one allele. The nuclear extract obtained from this cell line was active in splicing and contained wild-type amounts of 17S U2 particles suggesting that the presence of mutated SF3b1 protein does not interfere with the formation of 17S U2 particles. The sedimentation behavior on a sucrose or glycerol gradient was unchanged. This suggests that the incorporation of the mutation into the 17S U2 particle does not cause major compositional or structural changes in the U2. After confirming that the 17S U2 snRNPs forms normally, I affinity purified particles containing the wild-type or mutated SF3b1 via the N-terminal FLAG tag of SF3b1. The purification of FLAG-tagged 17S U2 particles was efficient (~80% of the bound particles) and mass spectrometry could confirm that the particle was not contaminated with significant amounts of untagged particles. In addition, the mass spectrometry analysis revealed that the majority of the U2-associated proteins are equally abundant in the wild-type and K700E particles. All commonly found proteins, such as the SF3a and SF3b proteins as well as others, like SPF45, hPrp43, SPF30 and U2 SURP, were found in equal amounts based on the number of peptides sequenced. Only two proteins were less abundant in the mutant 17S U2 snRNPs, hPrp5 and hTAT-SF1. A substantial reduction in hPrp5 and only a moderate reduction in hTAT-SF1 was also confirmed by western-blotting.

hTAT-SF1 is a transcription elongation factor that is also involved in pre-mRNA splicing (Miller et al., 2011). RNAi knockdown of hTAT-SF1 changes alternative splicing patterns, but the mechanism by which splicing is affected is not clear. In humans hTAT-SF1 interacts with the 17S U2 snRNP (Will and Lührmann, 2002). A recent study showed that it interacts with the ULM motives of SF3b1 in its N-terminal region (Loerch et al., 2019). This is consistent with my protein-protein crosslinking studies and indicates that the reduced association of hTAT-SF1 is not a direct effect of a change in the structure of the HEAT domain due to the SF3b1 mutation. In yeast, the hTAT-SF1 homolog Cus2 plays an important role in determining the structure of the U2 snRNA, which can have competing conformations, and has been implicated in the surveillance of U2 structure (Perriman and Ares, 2000). Mutagenesis and genetic depletion studies indicate that ATP hydrolysis by Prp5 is needed to displace Cus2 and thereby ensure the proper U2 structure for stable interaction with the BPS (Perriman et al., 2003). Binding of hTAT-SF1 to the human

U2 snRNP was only mildly affected and it is not clear if this change has a functional consequence for the U2/BPS interaction when SF3b1 is mutated.

Prp5 is an ATP-dependent DEAD-box RNA helicase, essential for splicing of pre-mRNAs in both human and yeast spliceosomes. For the human Prp5 (DDX46) protein very little is known, but it is necessary for the stable integration of the BPS helix and the formation of the spliceosomal A complex and (Will and Lührmann, 2002). The exact function of Prp5 in the splicing process in higher eukaryotes is currently unknown. Knock-down experiments in human cell lines could show that the reduction of Prp5 abundance causes a nearly complete cell cycle arrest and the depletion of Prp5 from nuclear extracts stalls pre-mRNA splicing *in vitro* (Will and Lührmann, 2002). Early studies in yeast indicate that Prp5 facilitates structural changes in U2 needed for the stable binding to the BPS. The absence of hPrp5 in the purified 17S U2 particles could be explained by a change in its interaction, either due to the binding of hPrp5 being less stable or completely prohibited. Both mechanisms would point towards structural changes as an underlying reason for the observed phenotype. Currently, it is not known exactly where hPrp5 binds to the human 17S U2 snRNP. My cross-linking experiments indicate that hPrp5 might contact multiple proteins within the particle, although the presence of cross-links only indicates physical proximity and not necessarily direct protein-protein contacts. Experiments performed in yeast showed that yPrp5 binds to the HEAT domains 1-6 and 9-12 (Tang et al., 2016). This observation most likely translates to the human 17S U2 particle because the cross-links between human SF3b1 and hPrp5 were found in similar regions. However, the yeast Prp5 protein lacks an N-terminal RS-domain found in humans, and thus the human protein might form additional interaction networks not identified in yeast. Nonetheless, the reduced binding of hPrp5 to the 17S U2 snRNP could result from changes in the conformation of the HEAT domain due to the K700E mutation. This can be due to both direct hPrp5-SF3b1 interactions or also indirectly through other protein-protein interactions affected by the K700E mutation.

5.2.2 Mechanisms whereby reduced Prp5 binding due to SF3b1 mutations can affect BPS selection

Studies with the yeast Prp5 (yPrp5) gave further insight into the function of Prp5 during spliceosome assembly and potential mechanisms for the selection of cryptic splice sites after mutations in the protein. Interestingly, the presence of the SF3b1 myelodysplastic syndrome associated mutations in yeast spliceosomes also leads to the emergence of splicing on cryptic branch point sequences, without causing splicing of divergent 5' or 3' splice sites (Carrocci et al., 2017).

The observations made in yeast mirror some of the observations I could make in the human system. The introduction of the cancer associated mutations in the homologous residues of the *S. cerevisiae* SF3b1 protein (denoted Hsh155) leads to reduced interactions between yPrp5 and ySF3b1 (Carrocci et al., 2017; Tang et al., 2016). Currently, there is no known mechanism how the presence of these mutations leads to the emergence of cryptic splicing, however, multiple models exist.

One model connecting the reduced interactions between SF3b1 and Prp5 with the usage of cryptic branch sites, explains the influence of the mutations on branch point usage by separating the actions of Prp5 and SF3b1 during spliceosome assembly. In this model, the mutations of SF3b1 decrease the stability of the interaction with the consensus branch site sequences, while Prp5 acts during a second proofreading step at later stages (Carrocci et al., 2017). This Prp5 independent mechanism mainly relies on structural changes in SF3b1 and resulting less stable binding of U2 as the cause for the changes in the selection of the branch site (Carrocci et al., 2017). Prp5 catalyzes a conformational change in U2 that allows stable U2 binding. The reduced interaction with Prp5 could hinder this structural change. Another function of Prp5 in the spliceosome is the bridging of the U1 and the U2 particle during the formation of early spliceosomal complexes (Shao et al., 2012; Xu et al., 2004). U1 binding has been shown to stabilize the U2/BPS interaction. Potentially the reduced interactions between SF3b1 and Prp5 would also destabilize the interaction of U2 during A complex assembly by disrupting the bridge to the U1 snRNP and thereby lead to the usage of alternative branch sites.

An alternative proposed mechanism describes the selection of cryptic splice sites as a consequence of altered proof-reading of the initial branch site selection due to the changed interaction between ySF3b1 and yPrp5 (Tang et al., 2016). In this model, the fidelity of for the branch point selection is directly proportional to the strength of the ySF3b1-yPrp5 interaction. Studies could also show that the dissociation speed and ATPase activity of the helicase directly correlates with the "quality" of the BPS used, with better branch point sequences, causing faster dissociation of yPrp5. If the U2/BPS interaction passes the quality control step, yPrp5 would dissociate and thereby allow the association of the tri-snRNP particle and progression of splicing (Tang et al., 2016). This would mean that the impaired interaction between ySF3b1 containing the MDS associated mutations and yPrp5 would lead to a reduced rate of U2/ BPS interaction, and conversion to the B complex. Thus, the slowed reaction would increase the time for finding cryptic splice sites (Carrocci et al., 2017; Tang et al., 2016). This observation is in line with the observations made *in vivo* in human cells expressing the MDS mutations of SF3b1 that the branch points used in the presence of the cancer related mutations of SF3b1 are not the preferred BPS used in the absence of SF3b1 mutations. The cryptic apparently are not optimal although they often have a higher complementary to the U2 snRNA (Darman et al., 2015).

My experiments demonstrate that the presence of the K700E mutation does negatively affects the interaction between SF3b1 and Prp5 in the human U2 snRNP. In addition, previous studies of the consequences of the mutations on the transcriptome of the affected cells showed that the mutations lead to the utilization of cryptic branch sites (Darman et al., 2015). However, the branch point sequences used in human spliceosomes are very degenerate compared to the yeast consensus sequence. This means that proofreading of the correct U2 RNA-BPS interactions by Prp5 may be absent in higher eukaryotes. Nonetheless, Prp5 may control for correct RNA-protein interactions. However, a change in proof reading is unlikely to be the only factor causing the selection of cryptic splice sites.

Future experiments, which will require substantial upscaling of K562 nuclear extract preparation, will be targeted at understanding how the consequences of the K700E mutation could affect the structure and function of the 17S U2 particle and pre-

mRNA splicing. A first goal would be to perform cryo-EM to resolve the three dimensional structure of both the wild-type and the K700E U2 snRNP at a high enough resolution to see the consequences of the amino acid substitution on the three-dimensional structure. However, it has been difficult to date to generate a high resolution 17S U2 structure, due to conformational flexibility. An alternative to this approach would be the cross-linking of proteins in the wild-type and SF3b1^{K700E} 17S U2 snRNP and a detailed comparison of the cross-links to identify potential differences. Secondly, to gain further insight into the function of U2 containing the SF3b1^{K700E} mutation, it will be interesting to see how *in vitro* splicing is affected using U2-depleted nuclear extract. It will also be of great importance to investigate *in* vitro which factors exactly impact the selection of branch points in the human system. However, because the K700E mutation also affects the Prp5 interaction with U2, it may be difficult to separate effects caused by Prp5 and those directly due to the SF3b1^{K700E} mutation.

5.2.3 Pre-mRNA splicing and splicing complex formation in the presence of SF3b1^{K700E}

The expression of SF3b1^{K700E} causes widespread changes in the splicing of pre-mRNA in the cell (Darman et al., 2015). To investigate the consequences of this mutation, I investigated the *in vitro* splicing pre-mRNAs that are alternatively spliced *in vivo* using nuclear extracts obtained from a SF3b1^{K700E} expressing K562 cell line. These pre-mRNAs, which were used as mini-genes for *in vivo* studies included ZDHHC16, ENOSF1 and TMEM14C (Darman et al., 2015). Unfortunately, the tested pre-mRNAs failed to splice under *in vitro* conditions; only formation of an A complex was observed in HeLa nuclear extract. Interestingly, no significant changes in the proteins commonly found in the A complex formed on MINX pre-mRNA where detected with the alternatively spliced pre-mRNA constructs. Only lower amounts of tri-snRNP associated proteins were found. This indicates that while normal formation of an A complex appears to take place; the transition towards later spliceosomal complexes (pre-B/ B complex) is impaired on the tested pre-mRNA constructs. Interestingly two other proteins, not directly associated with the A complex, show a significant difference between the MINX pre-mRNA control and the ZDHHC16 derived pre-mRNA constructs, namely Matrin-3 and CCAR2; both

are significantly more abundant in A complexes formed on ZDHHC16 and ZDHHC16_-30 pre-mRNA than on MINX.

Matrin-3 is a nuclear matrix protein that binds RNA (Castello et al., 2012; Uemura et al., 2017). It is a known regulator of alternative splicing and interacts with multiple splicing factors; however its exact function in the splicing process remains poorly understood (Coelho et al., 2015). CCAR2 is a component of the DBIRD-complex, which binds to RNA polymerase II and has been linked to alternative splicing and pre-mRNA transcription regulation (Close et al., 2012). In addition to these two proteins, the ZDHHC16 pre-mRNAs also bind more hnRNPs than the MINX pre-mRNA, while the amount of SR-proteins recruited remains at a similar level. The former is likely the result of the much longer RNA sequence of the ZDHHC16 versus MINX pre-mRNA (312 nt vs. 1073 nt).

It is not clear why the spliceosome assembly on the ZDHHC16 pre-mRNA stalls at the A complex stage under *in vitro* conditions. The *in vitro* system could lack essential factors required for later assembly steps and for the splicing of these specific constructs. Although an A complex forms it is not clear if it is functionally competent for subsequent splicing steps. Long introns are typically poorly spliced in HeLa nuclear extract and may reflect the need for cross-exon assembly instead of cross-intron assembly.

5.2.4 Cross-exon complex formation in the presence of SF3b1^{K700E}

To circumvent this problem, I thus switched to investigating cross-exon complexes on the alternatively spliced pre-mRNAs. All genes have a long intron preceding the cryptic 3’ splice site, making the cross-exon assembly the most likely path of spliceosome assembly *in vivo*. It is important to note, that on average, the cryptic branch point sequence is stronger (have a more complementarity to the U2 snRNA) compared to the canonical sequence (Deboever et al., 2015).

The investigation of A-like exon complex assembly revealed multiple interesting observations. First, the presence of SF3b1^{K700E} does not appear to interfere with the formation of A-like cross exon complexes on MINX exon RNA. Second, the constructs designed with the 3’ end of introns containing cryptic splice sites, form significantly less or less stable exon complex than the MINX exon RNA. This

observation held true independent of the nuclear extract used. Mutation of canonical branch point adenosines did not appear to impact the formation of the cross-exon A complexes indicating that alternative or cryptic branch sites are used instead. The observation that mutations of single nucleotides within the branch point sequence do not interfere with the assembly of early cross-intron spliceosomal complexes has been made before (Reed and Maniatis, 1988), but it is interesting to observe that cross-exon complexes also still assemble. In some cases, the mutation of the canonical branch point even appears to increase the amount of complex formed. This would suggest that the binding of U2 to the cryptic site is more stable or that simultaneous binding to both sites may interfere with or destabilize U2 snRNP binding.

Studies *in vivo* assaying for splicing and alternative 3' splice site usage of the ZDHHC16, ENOSF1 and TMEM14C introns that I investigated here, showed that mutation of the canonical branch site inhibited splicing. However, splicing and the utilization of a cryptic BS and 3'SS was observed if the cells expressed a mutated SF3b1 protein (Darman et al., 2015). Thus, one reason for using version of these pre-mRNAs or exon RNAs with a mutated canonical or cryptic branch site was to specifically purify splicing complexes containing solely the SF3b1^{K700E}. Cross-exon A-like formation *in vitro* was not dependent on the presence of SF3b1^{K700E} when the canonical BPS was mutated. This discrepancy with the *in vivo* results with ZDHHC16, ENOSF1 and TMEM14C substrates indicates that the cross-exon complex form on the cryptic BPS even when SF3b1^{K700E} is absent.

Complex assembly most likely takes place on both potential branch point sequences; independent of the presence of SF3b1^{K700E}, and the actual selection of the 3' SS must therefore take place at a later stage. One way to explain this would be that both the canonical and the cryptic splice site fulfill the requirements for the initial association of the U2 snRNP, but then at a later stage the A complex formed on the cryptic splice site fails to convert into a later spliceosomal complex, perhaps because Prp5 proof-reading leads to the discard of these complexes or simply because their structure or composition are not optimal for later steps of splicing.

5.2.5 Future studies

One aim of my studies was to determine if the presence of the SF3b1^{K700E} mutant affects the structure and composition of early splicing complexes. Immunoblotting of MS2-affinity purified cross-intron A complexes formed in K562 nuclear extract containing both wild-type and mutated SF3b1 showed that the tagged, mutant protein is indeed present in the assembled A complexes. Future studies will involve scaling up the production of the K562 nuclear wild-type and mutant cells to generate large amounts of nuclear extract. This will potentially allow the purification of sufficient amounts of both 17S U2 snRNPs and splicing complexes containing either the wild-type or solely the mutant SF3b1 protein. Initial attempts at using the N-terminal FLAG-tag of SF3b1 for the purification of spliceosomal components have shown that while the efficient purification of the 17S U2 particle is possible, the yield of purified A complex is significantly lower. This makes the upscaling of the reactions used for affinity purifications absolutely required to investigate the function of SF3b1^{K700E} in later spliceosomal complexes. Similar studies, including protein-protein and RNA-protein cross-linking and mass spectrometry analysis, can potentially be performed with the cross-exon complex formed on the various exon RNAs with cryptic 3’ SS and branch point sequences, although it is not presently clear if sufficient amounts of the FLAG-tagged wild-type or mutated SF3b1 is present in the A-like exon complexes. These studies could provide novel insights into the effects of the cancer-related mutations in SF3b1 on SF3b function and may improve our understanding of how they lead to cancer.

6 References

Absmeier, E., Santos, K.F., and Wahl, M.C. (2016). Functions and regulation of the Brr2 RNA helicase during splicing. Cell cycle (Georgetown, Tex) *15*, 3362-3377.

Achsel, T., Brahms, H., Kastner, B., Bachi, A., Wilm, M., and Luhrmann, R. (1999). A doughnut-shaped heteromer of human Sm-like proteins binds to the 3'-end of U6 snRNA, thereby facilitating U4/U6 duplex formation in vitro. Embo j *18*, 5789-5802.

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

Agrawal, A.A., Seiler, M., Brinton, L.T., Mantel, R., Lapalombella, R., Woyach, J.A., Johnson, A.J., Zhu, P., Warmuth, M., Yu, L.*, et al.* (2017). Novel SF3B1 in-frame deletions result in aberrant RNA splicing in CLL patients. Blood Adv *1*, 995-1000.

Akhtar, J., Kreim, N., Marini, F., Mohana, G., Brüne, D., Binder, H., and Roignant, J.-Y. (2019). Promoter-proximal pausing mediated by the exon junction complex regulates splicing. Nature Communications *10*.

Aronova, A., Bacikova, D., Crotti, L.B., Horowitz, D.S., and Schwer, B. (2007). Functional interactions between Prp8, Prp18, Slu7, and U5 snRNA during the second step of pre-mRNA splicing. Rna-a Publication of the Rna Society *13*, 1437-1444.

Bach, M., Winkelmann, G., and Luhrmann, R. (1989). 20S small nuclear ribonucleoprotein U5 shows a surprisingly complex protein composition. *86*, 6038-6042.

Bao, P., Hobartner, C., Hartmuth, K., and Luhrmann, R. (2017). Yeast Prp2 liberates the 5' splice site and the branch site adenosine for catalysis of pre-mRNA splicing. Rna *23*, 1770-1779.

Behrens, S.E., and Luhrmann, R. (1991). Immunoaffinity purification of a [U4/U6.U5] tri-snRNP from human cells. Genes Dev *5*, 1439-1452.

Behzadnia, N., Golas, M.M., Hartmuth, K., Sander, B., Kastner, B., Deckert, J., Dube, P., Will, C.L., Urlaub, H., Stark, H.*, et al.* (2007). Composition and three-dimensional EM structure of double affinity-purified, human prespliceosomal A complexes. *26*, 1737-1748.

Berget, S.M. (1995). Exon Recognition in Vertebrate Splicing. *270*, 2411-2414.

Bertram, K., Agafonov, D.E., Dybkov, O., Haselbach, D., Leelaram, M.N., Will, C.L., Urlaub, H., Kastner, B., Luhrmann, R., and Stark, H. (2017a). Cryo-EM Structure of a Pre-catalytic Human Spliceosome Primed for Activation. Cell *170*, 701-713.e711.

Bertram, K., Agafonov, D.E., Liu, W.-T., Dybkov, O., Will, C.L., Hartmuth, K., Urlaub, H., Kastner, B., Stark, H., and LüHrmann, R. (2017b). Cryo-EM structure of a human spliceosome activated for step 2 of splicing.

References

Bessonov, S., Anokhina, M., Krasauskas, A., Golas, M.M., Sander, B., Will, C.L., Urlaub, H., Stark, H., and 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.

Boesler, C., Rigo, N., Anokhina, M.M., Tauchert, M.J., Agafonov, D.E., Kastner, B., Urlaub, H., Ficner, R., Will, C.L., and Lührmann, R. (2016). A spliceosome intermediate with loosely associated tri-snRNP accumulates in the absence of Prp28 ATPase activity. *7*, 11997.

Bonnal, S., Vigevani, L., and Valcarcel, J. (2012). The spliceosome as a target of novel antitumour drugs. Nature reviews Drug discovery *11*, 847-859.

Boultwood, J., Dolatshad, H., Varanasi, S.S., Yip, B.H., and Pellagatti, A. (2014). The role of splicing factor mutations in the pathogenesis of the myelodysplastic syndromes. Adv Biol Regul *54*, 153-161.

Brow, D.A., and Guthrie, C. (1988). Spliceosomal RNA U6 is remarkably conserved from yeast to mammals. *334*, 213-218.

Burge, C.B., Sharp, P.A., and Tuschl, T. (1999). The RNA world. Cold Spring Harbor Press, 525-560.

Burgess, S.M., and Guthrie, C. (1993). A MECHANISM TO ENHANCE MESSENGER-RNA SPLICING FIDELITY - THE RNA-DEPENDENT ATPASE PRP16 GOVERNS USAGE OF A DISCARD PATHWAY FOR ABERRANT LARIAT INTERMEDIATES. Cell *73*, 1377-1391.

Bush, S.J., Chen, L., Tovar-Corona, J.M., and Urrutia, A.O. (2017). Alternative splicing and the evolution of phenotypic novelty. *372*, 20150474.

Carrocci, T.J., Zoerner, D.M., Paulson, J.C., and Hoskins, A.A. (2017). SF3b1 mutations associated with myelodysplastic syndromes alter the fidelity of branchsite selection in yeast. Nucleic Acids Res *45*, 4837-4852.

Cartegni, L., Chew, S.L., and Krainer, A.R. (2002). Listening to silence and understanding nonsense: exonic mutations that affect splicing. Nat Rev Genet *3*, 285-298.

Castello, A., Fischer, B., Eichelbaum, K., Horos, R., Beckmann, B.M., Strein, C., Davey, N.E., Humphreys, D.T., Preiss, T., Steinmetz, L.M.*, et al.* (2012). Insights into RNA biology from an atlas of mammalian mRNA-binding proteins. Cell *149*, 1393-1406.

Chan, S.P., Kao, D.I., Tsai, W.Y., and Cheng, S.C. (2003). The Prp19p-associated complex in spliceosome activation. Science *302*, 279-282.

Chanarat, S., and Strasser, K. (2013). Splicing and beyond: the many faces of the Prp19 complex. Biochim Biophys Acta *1833*, 2126-2134.

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

Chen, J.Y.-F., Stands, L., Staley, J.P., Jackups, R.R., Latus, L.J., and Chang, T.-H. (2001). Specific Alterations of U1-C Protein or U1 Small Nuclear RNA Can Eliminate the Requirement of Prp28p, an Essential DEAD Box Splicing Factor. *7*, 227-232.

Close, P., East, P., Dirac-Svejstrup, A.B., Hartmann, H., Heron, M., Maslen, S., Chariot, A., Soding, J., Skehel, M., and Svejstrup, J.Q. (2012). DBIRD complex integrates alternative mRNA splicing with RNA polymerase II transcript elongation. Nature *484*, 386-389.

Coelho, M.B., Attig, J., Bellora, N., König, J., Hallegger, M., Kayikci, M., Eyras, E., Ule, J., and Smith, C.W.J. (2015). Nuclear matrix protein Matrin3 regulates alternative splicing and forms overlapping regulatory networks with PTB. The EMBO journal *34*, 653-668.

Company, M., Arenas, J., and Abelson, J. (1991). REQUIREMENT OF THE RNA HELICASE-LIKE PROTEIN PRP22 FOR RELEASE OF MESSENGER-RNA FROM SPLICEOSOMES. Nature *349*, 487-493.

Cordin, O., Hahn, D., and Beggs, J.D. (2012). Structure, function and regulation of spliceosomal RNA helicases. *24*, 431-438.

Corrionero, A., Minana, B., and 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., Agrawal, A.A., Cook, A., Will, C.L., Fekkes, P., Smith, P.G., Lührmann, R., Larsen, N., Buonamici, S., and Pena, V. (2018). Structural Basis of Splicing Modulation by Antitumor Macrolide Compounds. Molecular Cell *70*, 265-273.e268.

Cretu, C., Schmitzová, J., Ponce-Salvatierra, A., Dybkov, O., Evelina, Sharma, K., Cindy, Urlaub, H., Lührmann, R., and Pena, V. (2016). Molecular Architecture of SF3b and Structural Consequences of Its Cancer-Related Mutations. *64*, 307-319.

Dardenne, E., Micaela, Fattet, L., Germann, S., Lambert, M.-P., Neil, H., Zonta, E., Mortada, H., Gratadou, L., Deygas, M.*, et al.* (2014). RNA Helicases DDX5 and DDX17 Dynamically Orchestrate Transcription, miRNA, and Splicing Programs in Cell Differentiation. *7*, 1900-1913.

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.*, 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.

David, C.J., and Manley, J.L. (2010). Alternative pre-mRNA splicing regulation in cancer: pathways and programs unhinged. *24*, 2343-2364.

De, I., Bessonov, S., Hofele, R., dos Santos, K., Will, C.L., Urlaub, H., Luhrmann, R., and Pena, V. (2015). The RNA helicase Aquarius exhibits structural adaptations mediating its recruitment to spliceosomes. Nat Struct Mol Biol *22*, 138-144.

Deboever, C., Ghia, E.M., Shepard, P.J., Rassenti, L., Barrett, C.L., Jepsen, K., Jamieson, C.H.M., Carson, D., Kipps, T.J., and Frazer, K.A. (2015). Transcriptome Sequencing Reveals Potential Mechanism of Cryptic 3' Splice Site Selection in SF3B1-mutated Cancers. *11*, e1004105.

Dignam, J.D., Lebovitz, R.M., and Roeder, R.G. (1983). Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. Nucleic acids research *11*, 1475-1489.

Dziembowski, A., Ventura, A.P., Rutz, B., Caspary, F., Faux, C., Halgand, F., Laprevote, O., and Seraphin, B. (2004). Proteomic analysis identifies a new complex required for nuclear pre-mRNA retention and splicing. Embo j *23*, 4847-4856.

Effenberger, K.A., Anderson, D.D., Bray, W.M., Prichard, B.E., Ma, N., Adams, M.S., Ghosh, A.K., and Jurica, M.S. (2014a). Coherence between Cellular Responses and in Vitro Splicing Inhibition for the Anti-tumor Drug Pladienolide B and Its Analogs. *289*, 1938-1947.

Effenberger, K.A., Anderson, D.D., Bray, W.M., Prichard, B.E., Ma, N.C., Adams, M.S., Ghosh, A.K., and Jurica, M.S. (2014b). Coherence between Cellular Responses and in Vitro Splicing Inhibition for the Anti-tumor Drug Pladienolide B and Its Analogs. Journal of Biological Chemistry *289*, 1938-1947.

Effenberger, K.A., Perriman, R.J., Bray, W.M., Lokey, R.S., Ares, M., and Jurica, M.S. (2013). A High-Throughput Splicing Assay Identifies New Classes of Inhibitors of Human and Yeast Spliceosomes. Journal of Biomolecular Screening *18*, 1110-1120.

Effenberger, K.A., Urabe, V.K., and Jurica, M.S. (2017). Modulating splicing with small molecular inhibitors of the spliceosome. Wiley Interdiscip Rev RNA *8*.

Egecioglu, D.E., and Chanfreau, G. (2011). Proofreading and spellchecking: A two-tier strategy for pre-mRNA splicing quality control. Rna *17*, 383-389.

Fabrizio, P., Dannenberg, J., Dube, P., Kastner, B., Stark, H., Urlaub, H., and Luehrmann, R. (2009). The Evolutionarily Conserved Core Design of the Catalytic Activation Step of the Yeast Spliceosome. Molecular Cell *36*, 593-608.

Fairman-Williams, M.E., Guenther, U.-P., and Jankowsky, E. (2010). SF1 and SF2 helicases: family matters. Current Opinion in Structural Biology *20*, 313-324.

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

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

Fica, S.M., Tuttle, N., Novak, T., Li, N.-S., Lu, J., Koodathingal, P., Dai, Q., Staley, J.P., and Piccirilli, J.A. (2013). RNA catalyses nuclear pre-mRNA splicing.

Folco, E.G., Coil, K.E., and Reed, R. (2011). The anti-tumor drug E7107 reveals an essential role for SF3b in remodeling U2 snRNP to expose the branch point-binding region. Genes Dev *25*, 440-444.

Galej, W.P., Wilkinson, M.E., Fica, S.M., Oubridge, C., Newman, A.J., and Nagai, K. (2016). Cryo-EM structure of the spliceosome immediately after branching. Nature *537*, 197-201.

Girard, C., Will, C.L., Peng, J., Makarov, E.M., Kastner, B., Lemm, I., Urlaub, H., Hartmuth, K., and Luhrmann, R. (2012). Post-transcriptional spliceosomes are retained in nuclear speckles until splicing completion. Nat Commun *3*, 994.

Gozani, O., Feld, R., and Reed, R. (1996). Evidence that sequence-independent binding of highly conserved U2 snRNP proteins upstream of the branch site is required for assembly of spliceosomal complex A. Genes Dev *10*, 233-243.

Gozani, O., Potashkin, J., and Reed, R. (1998). A potential role for U2AF-SAP 155 interactions in recruiting U2 snRNP to the branch site. Molecular and Cellular Biology *18*, 4752-4760.

Grinthal, A., Adamovic, I., Weiner, B., Karplus, M., and Kleckner, N. (2010). PR65, the HEAT-repeat scaffold of phosphatase PP2A, is an elastic connector that links force and catalysis. Proc Natl Acad Sci U S A *107*, 2467-2472.

Hall, S.L., and Padgett, R.A. (1994). Conserved sequences in a class of rare eukaryotic nuclear introns with non-consensus splice sites. Journal of molecular biology *239*, 357-365.

Hall, S.L., and Padgett, R.A. (1996). Requirement of U12 snRNA for in vivo splicing of a minor class of eukaryotic nuclear pre-mRNA introns. Science *271*, 1716-1718.

Hasegawa, M., Miura, T., Kuzuya, K., Inoue, A., Ki, S.W., Horinouchi, S., Yoshida, T., Kunoh, T., Koseki, K., Mino, K.*, et al.* (2011). Identification of SAP155 as the Target of GEX1A (Herboxidiene), an Antitumor Natural Product. Acs Chemical Biology *6*, 229-233.

Haselbach, D., Komarov, I., Agafonov, D.E., Hartmuth, K., Graf, B., Dybkov, O., Urlaub, H., Kastner, B., Luhrmann, R., and Stark, H. (2018). Structure and Conformational Dynamics of the Human Spliceosomal B(act) Complex. Cell *172*, 454-464.e411.

Heinrichs, V., Bach, M., Winkelmann, G., and Luhrmann, R. (1990). U1-specific protein C needed for efficient complex formation of U1 snRNP with a 5' splice site. *247*, 69-72.

Jackson, R.J., and Standart, N. (1990). Do the poly(A) tail and 3′ untranslated region control mRNA translation? *62*, 15-24.

Jenkins, J.L., and Kielkopf, C.L. (2017). Splicing Factor Mutations in Myelodysplasias: Insights from Spliceosome Structures. Trends Genet *33*, 336-348.

Juneau, K., Nislow, C., and Davis, R.W. (2009). Alternative splicing of PTC7 in Saccharomyces cerevisiae determines protein localization. Genetics *183*, 185-194.

Kaida, D., Motoyoshi, H., Tashiro, E., Nojima, T., Hagiwara, M., Ishigami, K., Watanabe, H., Kitahara, T., Yoshida, T., Nakajima, H.*, et al.* (2007). Spliceostatin A targets SF3b and inhibits both splicing and nuclear retention of pre-mRNA. Nat Chem Biol *3*, 576-583.

Kastner, B., Fischer, N., Golas, M.M., Sander, B., Dube, P., Boehringer, D., Hartmuth, K., Deckert, J., Hauer, F., Wolf, E.*, et al.* (2008). GraFix: sample preparation for single-particle electron cryomicroscopy. Nat Methods *5*, 53-55.

Kastner, B., Will, C.L., Stark, H., and Luhrmann, R. (2019). Structural Insights into Nuclear pre-mRNA Splicing in Higher Eukaryotes. Cold Spring Harb Perspect Biol.

Kelly, S.M., and Corbett, A.H. (2009). Messenger RNA export from the nucleus: a series of molecular wardrobe changes. Traffic (Copenhagen, Denmark) *10*, 1199-1208.

Keren, H., Lev-Maor, G., and Ast, G. (2010). Alternative splicing and evolution: diversification, exon definition and function. Nature Reviews Genetics *11*, 345-355.

Kohtz, J.D., Jamison, S.F., Will, C.L., Zuo, P., Lührmann, R., Garcia-Blanco, M.A., and Manley, J.L. (1994). Protein–protein interactions and 5'-splice-site recognition in mammalian mRNA precursors. *368*, 119-124.

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

Koonin, E.V. (1991). Similarities in RNA helicases. *352*, 290-290.

Kotake, Y., Sagane, K., Owa, T., Mimori-Kiyosue, Y., Shimizu, H., Uesugi, M., Ishihama, Y., Iwata, M., and Mizui, Y. (2007). Splicing factor SF3b as a target of the antitumor natural product pladienolide. *3*, 570-575.

Kuhn, A.N., van Santen, M.A., Schwienhorst, A., Urlaub, H., and Luhrmann, R. (2009). Stalling of spliceosome assembly at distinct stages by small-molecule inhibitors of protein acetylation and deacetylation. RNA *15*, 153-175.

Lagisetti, C., Pourpak, A., Jiang, Q., Cui, X.L., Goronga, T., Morris, S.W., and Webb, T.R. (2008). Antitumor compounds based on a natural product consensus pharmacophore. Journal of Medicinal Chemistry *51*, 6220-6224.

Lauber, J., Plessel, G., Prehn, S., Will, C.L., Fabrizio, P., Groning, K., Lane, W.S., and Luhrmann, R. (1997). The human U4/U6 snRNP contains 60 and 90kD proteins that are structurally homologous to the yeast splicing factors Prp4p and Prp3p. Rna *3*, 926-941.

Leitner, A., Joachimiak, L.A., Unverdorben, P., Walzthoeni, T., Frydman, J., Forster, F., and Aebersold, R. (2014). Chemical cross-linking/mass spectrometry targeting acidic residues in proteins and protein complexes. Proc Natl Acad Sci U S A *111*, 9455-9460.

Lin, C.W., Cheng, C.W., Yang, T.C., Li, S.W., Cheng, M.H., Wan, L., Lin, Y.J., Lai, C.H., Lin, W.Y., and Kao, M.C. (2008). Interferon antagonist function of Japanese encephalitis

virus NS4A and its interaction with DEAD-box RNA helicase DDX42. Virus research *137*, 49-55.

Lin, S., and Fu, X.D. (2007). SR proteins and related factors in alternative splicing. Adv Exp Med Biol *623*, 107-122.

Little, J.T., and Jurica, M.S. (2008). Splicing factor SPF30 bridges an interaction between the prespliceosome protein U2AF35 and tri-small nuclear ribonucleoprotein protein hPrp3. The Journal of biological chemistry *283*, 8145-8152.

Loerch, S., Leach, J.R., Horner, S.W., Maji, D., Jenkins, J.L., Pulvino, M.J., and Kielkopf, C.L. (2019). The pre-mRNA splicing and transcription factor Tat-SF1 is a functional partner of the spliceosome SF3b1 subunit via a U2AF homology motif interface. The Journal of biological chemistry *294*, 2892-2902.

Lührmann, R., Kastner, B., and Bach, M. (1990). Structure of spliceosomal snRNPs and their role in pre-mRNA splicing. *1087*, 265-292.

Makarova, O.V., Makarov, E.M., Urlaub, H., Will, C.L., Gentzel, M., Wilm, M., and Lührmann, R. (2004). A subset of human 35S U5 proteins, including Prp19, function prior to catalytic step 1 of splicing. *23*, 2381-2391.

Martin, A., Schneider, S., and Schwer, B. (2002). Prp43 is an essential RNA-dependent ATPase required for release of lariat-intron from the spliceosome. Journal of Biological Chemistry *277*, 17743-17750.

Martinez-Contreras, R., Cloutier, P., Shkreta, L., Fisette, J.F., Revil, T., and Chabot, B. (2007). hnRNP proteins and splicing control. Adv Exp Med Biol *623*, 123-147.

Mathew, R., Hartmuth, K., Mohlmann, S., Urlaub, H., Ficner, R., and Luhrmann, R. (2008). Phosphorylation of human PRP28 by SRPK2 is required for integration of the U4/U6-U5 tri-snRNP into the spliceosome. Nat Struct Mol Biol *15*, 435-443.

Mattaj, I.W. (1986). Cap trimethylation of U snRNA is cytoplasmic and dependent on U snRNP protein binding. Cell *46*, 905-911.

Meister, G., Hannus, S., Plottner, O., Baars, T., Hartmann, E., Fakan, S., Laggerbauer, B., and Fischer, U. (2001). SMNrp is an essential pre-mRNA splicing factor required for the formation of the mature spliceosome. Embo j *20*, 2304-2314.

Miller, H.B., Robinson, T.J., Gordân, R., Hartemink, A.J., and Garcia-Blanco, M.A. (2011). Identification of Tat-SF1 cellular targets by exon array analysis reveals dual roles in transcription and splicing. RNA (New York, NY) *17*, 665-674.

Möhlmann, S., Mathew, R., Neumann, P., Schmitt, A., Lührmann, R., and Ficner, R. (2014). Structural and functional analysis of the human spliceosomal DEAD-box helicase Prp28. *70*, 1622-1630.

Moore, M.J., and Sharp, P.A. (1993). Evidence for two active sites in the spliceosome provided by stereochemistry of pre-mRNA splicing. Nature *365*, 364-368.

Mozaffari-Jovin, S., Santos, K.F., Hsiao, H.H., Will, C.L., Urlaub, H., Wahl, M.C., and Luhrmann, R. (2012). The Prp8 RNase H-like domain inhibits Brr2-mediated U4/U6 snRNA unwinding by blocking Brr2 loading onto the U4 snRNA. *26*, 2422-2434.

Mozaffari-Jovin, S., Wandersleben, T., Santos, K.F., Will, C.L., Lührmann, R., and Wahl, M.C. (2014). Novel regulatory principles of the spliceosomal Brr2 RNA helicase and links to retinal disease in humans. *11*, 298-312.

Mroczek, S., and Dziembowski, A. (2013). U6 RNA biogenesis and disease association. *4*, 581-592.

Nesic, D., and Kramer, A. (2001). Domains in Human Splicing Factors SF3a60 and SF3a66 Required for Binding to SF3a120, Assembly of the 17S U2 snRNP, and Prespliceosome Formation. Molecular and Cellular Biology *21*, 6406-6417.

Newman, A.J. (1997). The role of U5 snRNP in pre-mRNA splicing. *16*, 5797-5800.

Nguyen, T.H., Galej, W.P., Bai, X.C., Savva, C.G., Newman, A.J., Scheres, S.H., and Nagai, K. (2015). The architecture of the spliceosomal U4/U6.U5 tri-snRNP. Nature *523*, 47-52.

Nguyen, T.H.D., Galej, W.P., Bai, X.C., Oubridge, C., Newman, A.J., Scheres, S.H.W., and Nagai, K. (2016). Cryo-EM structure of the yeast U4/U6.U5 tri-snRNP at 3.7 A resolution. Nature *530*, 298-302.

Nilsen, T.W. (1994). RNA-RNA interactions in the spliceosome: Unraveling the ties that bind. *78*, 1-4.

O'Day, C.L., Dalbadie-Mcfarland, G., and Abelson, J. (1996). The Saccharomyces cerevisiae Prp5 Protein Has RNA-dependent ATPase Activity with Specificity for U2 Small Nuclear RNA. *271*, 33261-33267.

Patel, A.A., and Steitz, J.A. (2003). Splicing double: insights from the second spliceosome. *4*, 960-970.

Pawellek, A., McElroy, S., Samatov, T., Mitchell, L., Woodland, A., Ryder, U., Gray, D., Luhrmann, R., and Lamond, A.I. (2014). Identification of Small Molecule Inhibitors of Pre-mRNA Splicing. Journal of Biological Chemistry *289*, 34683-34698.

Perriman, R., and Ares, M., Jr. (2000). ATP can be dispensable for prespliceosome formation in yeast. Genes & development *14*, 97-107.

Perriman, R., Barta, I., Voeltz, G.K., Abelson, J., and Ares, M. (2003). ATP requirement for Prp5p function is determined by Cus2p and the structure of U2 small nuclear RNA. *100*, 13857-13862.

Perriman, R.J., and Ares, M. (2007). Rearrangement of competing U2 RNA helices within the spliceosome promotes multiple steps in splicing. *21*, 811-820.

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

Price, A.M., Gornemann, J., Guthrie, C., and Brow, D.A. (2014). An unanticipated early function of DEAD-box ATPase Prp28 during commitment to splicing is modulated by U5 snRNP protein Prp8. *20*, 46-60.

Rappsilber, J., Ajuh, P., Lamond, A.I., and Mann, M. (2001). SPF30 is an essential human splicing factor required for assembly of the U4/U5/U6 tri-small nuclear ribonucleoprotein into the spliceosome. The Journal of biological chemistry *276*, 31142-31150.

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

Reed, R., and Maniatis, T. (1988). The role of the mammalian branchpoint sequence in pre-mRNA splicing. Genes Dev *2*, 1268-1276.

Ruby, S.W., Chang, T.H., and Abelson, J. (1993). Four yeast spliceosomal proteins (PRP5, PRP9, PRP11, and PRP21) interact to promote U2 snRNP binding to pre-mRNA. Genes Dev *7*, 1909-1925.

Sakharkar, M.K., Chow, V.T., and Kangueane, P. (2004). Distributions of exons and introns in the human genome. In silico biology *4*, 387-393.

Schaffert, N., Hossbach, M., Heintzmann, R., Achsel, T., and Lührmann, R. (2004). RNAi knockdown of hPrp31 leads to an accumulation of U4/U6 di-snRNPs in Cajal bodies. *23*, 3000-3009.

Schneider, M., Hsiao, H.H., Will, C.L., Giet, R., Urlaub, H., and Luhrmann, R. (2010a). Human PRP4 kinase is required for stable tri-snRNP association during spliceosomal B complex formation. Nature Structural & Molecular Biology *17*, 216-U212.

Schneider, M., Will, C.L., Anokhina, M., Tazi, J., Urlaub, H., and Luhrmann, R. (2010b). Exon definition complexes contain the tri-snRNP and can be directly converted into B-like precatalytic splicing complexes. Mol Cell *38*, 223-235.

Schneider, S., Hotz, H.R., and Schwer, B. (2002). Characterization of dominant-negative mutants of the DEAH-box splicing factors Prp22 and Prp16. Journal of Biological Chemistry *277*, 15452-15458.

Shao, W., Kim, H.-S., Cao, Y., Xu, Y.-Z., and Query, C.C. (2012). A U1-U2 snRNP Interaction Network during Intron Definition. Molecular and Cellular Biology *32*, 470-478.

Shen, H., Zheng, X., Shen, J., Zhang, L., Zhao, R., and Green, M.R. (2008). Distinct activities of the DExD/H-box splicing factor hUAP56 facilitate stepwise assembly of the spliceosome. *22*, 1796-1803.

Shevchenko, A., Wilm, M., Vorm, O., and Mann, M. (1996). Mass spectrometric sequencing of proteins silver-stained polyacrylamide gels. Analytical chemistry *68*, 850-858.

Silverman, E.J., Maeda, A., Wei, J., Smith, P., Beggs, J.D., and Lin, R.J. (2004). Interaction between a G-patch protein and a spliceosomal DEXD/H-box ATPase that is critical for splicing. Molecular and Cellular Biology *24*, 10101-10110.

Small, E.C., Leggett, S.R., Winans, A.A., and Staley, J.P. (2006). The EF-G-like GTPase Snu114p Regulates Spliceosome Dynamics Mediated by Brr2p, a DExD/H Box ATPase. Molecular Cell *23*, 389-399.

Sohn, S.O., and Chay, K.O. (2019). The ATP-dependent RNA helicase, DDX42 interacts with paxillin and regulates apoptosis and polarization of Ba/F3 cells. Animal cells and systems *23*, 1-9.

Sontheimer, E., and Steitz, J. (1993). The U5 and U6 small nuclear RNAs as active site components of the spliceosome. *262*, 1989-1996.

Staley, J.P., and Guthrie, C. (1999). An RNA Switch at the 5′ Splice Site Requires ATP and the DEAD Box Protein Prp28p. *3*, 55-64.

Sterner, D.A., Carlo, T., and Berget, S.M. (1996). Architectural limits on split genes. Proc Natl Acad Sci U S A *93*, 15081-15085.

Tang, Q., Rodriguez-Santiago, S., Wang, J., Pu, J., Yuste, A., Gupta, V., Moldon, A., Xu, Y.Z., and Query, C.C. (2016). SF3B1/Hsh155 HEAT motif mutations affect interaction with the spliceosomal ATPase Prp5, resulting in altered branch site selectivity in pre-mRNA splicing. Genes Dev *30*, 2710-2723.

Thompson, A., Schafer, J., Kuhn, K., Kienle, S., Schwarz, J., Schmidt, G., Neumann, T., Johnstone, R., Mohammed, A.K., and Hamon, C. (2003). Tandem mass tags: a novel quantification strategy for comparative analysis of complex protein mixtures by MS/MS. Analytical chemistry *75*, 1895-1904.

Toor, N., Rajashankar, K., Keating, K.S., and Pyle, A.M. (2008). Structural basis for exon recognition by a group II intron. *15*, 1221-1222.

Tsai, R.-T., Tseng, C.-K., Lee, P.-J., Chen, H.-C., Fu, R.-H., Chang, K.-j., Yeh, F.-L., and Cheng, S.-C. (2007). Dynamic interactions of Ntr1-Ntr2 with Prp43 and with U5 govern the recruitment of Prp43 to mediate spliceosome disassembly. Molecular and Cellular Biology *27*, 8027-8037.

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

Uemura, Y., Oshima, T., Yamamoto, M., Reyes, C.J., Costa Cruz, P.H., Shibuya, T., and Kawahara, Y. (2017). Matrin3 binds directly to intronic pyrimidine-rich sequences and controls alternative splicing. Genes to cells : devoted to molecular & cellular mechanisms *22*, 785-798.

Uhlmann-Schiffler, H. (2006). Ddx42p--a human DEAD box protein with RNA chaperone activities. *34*, 10-22.

Vigevani, L., and Valcárcel, J. A splicing magic bullet Drugs that modulate RNA splicing are potential therapeutics for spinal muscular atrophy.

Wahl, M.C., Will, C.L., and Lührmann, R. (2009). The Spliceosome: Design Principles of a Dynamic RNP Machine. Cell *136*, 701-718.

Wang, E.T., Sandberg, R., Luo, S., Khrebtukova, I., Zhang, L., Mayr, C., Kingsmore, S.F., Schroth, G.P., and Burge, C.B. (2008). Alternative isoform regulation in human tissue transcriptomes. Nature *456*, 470.

Warkocki, Z., Odenwaelder, P., Schmitzova, J., Platzmann, F., Stark, H., Urlaub, H., Ficner, R., Fabrizio, P., and Luehrmann, R. (2009). Reconstitution of both steps of Saccharomyces cerevisiae splicing with purified spliceosomal components. Nature Structural & Molecular Biology *16*, 1237-U1250.

Will, C.L., and Luhrmann, R. (2011). Spliceosome structure and function. Cold Spring Harb Perspect Biol *3*.

Will, C.L., and Lührmann, R. (2002). Characterization of novel SF3b and 17S U2 snRNP proteins, including a human Prp5p homologue and an SF3b DEAD-box protein. The EMBO Journal *21*, 4978-4988.

Will, C.L., Schneider, C., MacMillan, A.M., Katopodis, N.F., Neubauer, G., Wilm, M., Luhrmann, R., and Query, C.C. (2001). A novel U2 and U11/U12 snRNP protein that associates with the pre-mRNA branch site. *20*, 4536-4546.

Xu, Y.-Z., and Query, C.C. (2007). Competition between the ATPase prp5 and branch region-U2 snRNA pairing modulates the fidelity of spliceosome assembly. Molecular Cell *28*, 838-849.

Xu, Y.Z., Newnham, C.M., Kameoka, S., Huang, T., Konarska, M.M., and Query, C.C. (2004). Prp5 bridges U1 and U2 snRNPs and enables stable U2 snRNP association with intron RNA. Embo j *23*, 376-385.

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

Yean, S.L., Wuenschell, G., Termini, J., and Lin, R.J. (2000). Metal-ion coordination by U6 small nuclear RNA contributes to catalysis in the spliceosome. Nature *408*, 881-884.

Yokoi, A., Kotake, Y., Takahashi, K., Kadowaki, T., Matsumoto, Y., Minoshima, Y., Sugi, N.H., Sagane, K., Hamaguchi, M., Iwata, M.*, et al.* (2011). Biological validation that SF3b is a target of the antitumor macrolide pladienolide. *278*, 4870-4880.

Yoshida, K., and Ogawa, S. (2014). Splicing factor mutations and cancer. Wiley Interdiscip Rev RNA *5*, 445-459.

Zachariae, U., and Grubmuller, H. (2008). Importin-beta: structural and dynamic determinants of a molecular spring. Structure *16*, 906-915.

Zhan, X., Yan, C., Zhang, X., Lei, J., and Shi, Y. (2018). Structures of the human pre-catalytic spliceosome and its precursor spliceosome. Cell Res *28*, 1129-1140.

Zhang, X., Yan, C., Hang, J., Finci, L.I., Lei, J., and Shi, Y. (2017). An Atomic Structure of the Human Spliceosome. Cell *169*, 918-929.e914.

Zhang, X., Yan, C., Zhan, X., Li, L., Lei, J., and Shi, Y. (2018). Structure of the human activated spliceosome in three conformational states. Cell Res *28*, 307-322.

Zillmann, M., Zapp, M.L., and Berget, S.M. (1988). Gel electrophoretic isolation of splicing complexes containing U1 small nuclear ribonucleoprotein particles. Mol Cell Biol *8*, 814-821.

7 Appendix

7.1 List of Figures

Appendix

Figure 5.3 – Glycerol sedimentation profiles of splicing complexes formed in the presence of 250nM Pladienolide B or DMSO.

Figure 5.3 – Glycerol sedimentation profiles of splicing complexes formed in the presence of 250nM Pladienolide B or DMSO.

Figure 5.4- Stability of A complexes in the presence of increasing concentrations of heparin.

Figure 5.5 – Purification of human A complexes under kinetically-stalled conditions or stalled with Pladienolide B.

Figure 5.6 – Immunodepletion of SF3b from HeLa nuclear extract.

Figure 5.7- Pladienolide B stalled A complexes are not chased into active spliceosomes.

Figure 5.8 – Electron microscopy of the Pladienolide B stalled human A complex and of the kinetic human A complex.

Figure 5.9- Pladienolide B inhibits the formation of a cross-exon B-like complex.

Figure 5.10 – Tagged SF3b1 purification from transiently transfected HEK293T nuclear extract.

Figure 5.11 – Both K562 cell lines express FLAG-tagged SF3b1 protein.

Figure 5.12 – Determination of the doubling times of K562 cells expressing a FLAG-tagged SF3b1wt or SF3b1K700E protein.

Figure 5.13 – MINX pre-mRNA splicing in nuclear extract obtained from K562 cells expressing wild-type and tagged SF3b1.

Figure 5.14 – Sedimentation behavior of U1, U2, U5 and U4/U6 snRNPs in the presence of SF3b1K700E.

Figure 5.15- Isolation of 17S U2 particles via the N-terminal FLAG-tag.

Figure 5.16- Fragmentation scheme of SF3b1K700E.

7.2 List of Tables

Abbreviations

°C	degrees Celsius
µl	microliter
µm	micrometer
µM	micromolar
2D	two dimensional
^{32}P	radioactive isotope of phosphorus
3D	three dimensional
3'SS	3 prime splice site
5'SS	5 prime splice site
A	adenosine
Å	Angström
AA	acrylamide
AA	anacardic acid
AML	acute myeloid leukemia
APS	ammonium persulfate
ATP	adenosine tri-phosphate
BBP	branch point binding protein
bp	base pair/ base pairs
BP	branch point
BPA	branch point adenosine
BPS	branch point sequence
C	cytosine
CMML	chronic myelomonocytic leukemia
cntrl	control
Da	Dalton
ddH2O	bidistilled water
DMEM	Dulbecco's Modified Eagle Medium
DMSO	Dimethyl sulfoxide
DNA	desoxyribonucleic acid
DTE	dithioerythritol
DTT	dithiothreitol
E.coli	Escherichia coli
ECL	enhanced chemiluminescence
EDTA	ethylenediaminetetraacetic acid
EJC	exon junction complex
EM	electron microscopy
ENOSF1	Enolase Superfamily Member 1
ESE	exonic splice enhancers
ESS	exonic splice silencers
EtOH	ethanol
G	guanine
h	hour
HEK293T	Human embryonic kidney 293 cells
HEPES	N-2-Hydroxyethylpiperazin-N-2-ethansulfonic acid

hnRNP	heterogeneous ribonucleoprotein particle
IMEM	Improved Minimum Essential Medium
ISE	intronic splice silencers
ISS	intronic splice enhancers
kDa	kilodalton
l	liter
LB	lysogeny broth
LDS	lithium dodecylsulfate
M	molar
MDS	myelodysplastic syndromes
min	minute
ml	milliliter
mm	millimeter
mM	millimolar
mRNA	messenger RNA
MS2-MBP	MS2 loop binding - maltose binding protein fusion protein
NaOAc	sodium acetate
NE	nuclear extract
nM	nanomolar
nm	nanometer
nt	nucleotide
OD	optical density
PA	poly acrylamide
PAGE	poly acrylamide gel electrophoresis
PCI	Phenol-Chloroform-Isoamyl extraction
PCI	Phenol-Chloroform-Isoamyl alcohol
PEI	polyethyleneimine
PK	proteinase K
PlaB	Pladienolide B
pre-mRNA	premature messenger RNA
pY	poly pyrmidine
RC	Röder C buffer
rcf	relative centrifugal force
RD	Röder D buffer
RNA	ribonucleic acid
RNP	ribonucleoprotein particle
rpm	rounds per minute
RT	room temperature
SDS	sodium dodecyl sulfate
snRNA	small nuclear RNA
snRNP	small nuclear ribonucleoprotein
SOC	Super Optimal broth with Catabolite repression
T	thymine
TEAB	triethylammonium bicarbonate
TEMED	tetramethylethylenediamine
TMEM14C	Transmembrane Protein 14C

TMT	tandem mass tag
Tris	Tris-(hydroxymethyl)aminomethane
U	uracil
V	volt
v/v	volume per volume
w/v	weight per volume
Y	pyrimidine nucleotide
ZDHHC16	Zinc Finger DHHC-Type Containing 16

Acknowledgements

I would like to thank Prof. Reinhard Lührmann for the opportunity to work on this challenging and interesting project. I also would like to thank him for the time and energy he invested into the project, for the discussions and the support he provided.

I also would like to thank Dr. Cindy Will for the support she provided throughout the project in form of scientific advice and countless discussions.

A special thank you also to my thesis committee and the other members of examination board: Prof. Ralf Ficner, Prof. Patrick Cramer as well as Prof. Henning Urlaub, Prof. Ralph Kehlenbach and Prof. Detlef Doenecke.

I also want to thank Dr. Norbert Rigo for the help with the EM experiments, Dr. Olexandr Dybkov for his guidance with the crosslinking experiments and Dr. Kuan-Ting Pan for helping me with the proteomic analysis of the cell lines.

A special thanks also to Dr. James Manley for kindly providing the K562 cell lines used in this thesis.

In addition, I want to thank the group of Bioanalytical Mass Spectrometry for the support with all MS-related experiments and questions.

Finally, I would like to thank my family for supporting me throughout my life and for always being my safe haven.