Identification of a protein factor shared by the recycling pathways of U2- and U12-type spliceosomes

Identifizierung eines Proteinfaktors für das Recycling von U2- und U12-abhängigen Spleißosomen

Inaugural-Dissertation
zur
Erlangung des Doktorgrades
der Naturwissenschaftlichen Fachbereiche
der Justus-Liebig-Universität Gießen

vorgelegt von

Andrey Damianov

geb. am 22.09.1972 in Plovdiv, Bulgarien

Gießen, März 2004

Die vorliegende Arbeit wurde am Institut für Biochemie des Fachbereiches 08 der Justus-Liebig-Universität Gießen in der Zeit von Dezember 2000 bis März 2004 unter der Anleitung von Prof. Dr. Albrecht Bindereif angefertigt.

Dekan:	Prof. Dr. Jürgen Mayer Institut für Biologiedidaktik Justus-Liebig-Universität Gießen
1. Gutachter:	Prof. Dr. Albrecht Bindereif Institut für Biochemie Justus-Liebig-Universität Gießen
2. Gutachter:	Prof. Dr. Rainer Renkawitz Institut für Genetik Justus-Liebig-Universität Gießen

Table of contents

Zusammenfassung

mRNA-Spleißen findet in einem großen Ribonukleoproteinkomplex (RNP) statt, dem Spleißosom, das sich aus 5 kleinen nukleären RNPs (snRNPs) und über 100 Proteinfaktoren zusammensetzt. Es gibt zwei Klassen von Introns, die sogenannten *major*- und *minor*-Introns, die von zwei verschiedenen Spleißosomen prozessiert werden. Diese enthalten unterschiedliche Gruppen von snRNPs: U1, U2, U4 und U6 snRNPs werden für das Spleißen von *major*-Introns benötigt; U11, U12, U4atac und U6atac bilden die funktionsanalogen snRNPs, welche für das Spleißen von *minor*-Introns essentiell sind. Der U5 snRNP dagegen kommt in beiden Spleißosom-Formen vor. Die U4 und U6 snRNAs werden in das Spleißosom in Form des U4/U6.U5 tri-snRNPs integriert, in welchem sie eine ausgedehnte Basenpaarung aufweisen. Nach Ablauf der Spleißreaktion werden U4 und U6 separat voneinander freigesetzt. Um daraus den tri-snRNP für weitere Runden der Spleißreaktion zu bilden, muß zunächst der U4/U6-Hybrid durch eine Recycling-Reaktion wiederhergestellt werden. Im Hefesystem wurde das Prp24-Protein, im Humansystem dessen verwandtes Protein p110 als U4/U6-Recyclingfaktor identifiziert und als spezifische Proteinkomponente der U6 und U4/U6 snRNPs.

Diese Arbeit konzentriert sich auf das Recycling der U4atac und U6atac snRNPs. Mit Hilfe eines mit der Spleißreaktion gekoppelten *in vitro*-Assays konnte gezeigt werden, daß p110 auch als U4atac/U6atac-Recyclingfaktor wirkt. Das p110-Protein wurde im U6atac snRNP nachgewiesen; im Gegensatz zum normalen U4/U6 snRNP ist p110 nur minimal mit dem U4atac/U6atac snRNP assoziiert. Interessanterweise können die U4/U6- und U4atac/U6atac-Hybride auch unabhängig von der *in vitro*-Spleißreaktion gespalten werden. Dieser Prozeß stellt *in vivo* möglicherweise einen Kontrollmechanismus der Spleißreaktion dar.

Die für die p110-Interaktion notwendigen Sequenzelemente der U6 und U6atac snRNAs wurden kartiert: p110 benötigt Positionen 38-57 in U6 bzw. 10-30 in U6atac. Diese Bereiche sind zwischen den U6 und U6atac snRNAs am stärksten konserviert. Der singuläre U6 snRNP wurde außerdem biochemisch aus HeLa S100-Extrakt gereinigt. Als eine neue Komponente wurde ein Protein, p85, mittels Massenspektrometrie identifiziert (cDNA *GenBank accession* AK001239). Diese Protein befindet sich im Nukleoplasma und liegt nicht nur mit U6 snRNA gebunden vor, sondern auch mit U2 sowie 5S und 5.8S ribosomaler RNA. p85 zeigt eine geringe Ähnlichkeit mit dem an der ribosomalen RNA-Prozessierung beteiligten NOP4-Faktor aus der Hefe, was eine Funktion von p85 bei der Reifung der U2 und U6 snRNAs nahelegt.

Summary

Pre-mRNA splicing takes place in a large ribonucleoprotein complex, the spliceosome, which contains five small nuclear ribonucleoprotein complexes (snRNPs) and over 100 protein factors. There are two types of introns, the major- and the minor-type, that are excised by two different spliceosomes, assembled by two distinct sets of snRNPs: U1, U2, U4, and U6 snRNPs are required for splicing of the major-type introns; U11, U12, U4atac, and U6atac are their corresponding counterparts, involved in minor-type intron splicing. U5 snRNP is common for both spliceosomes. U4 and U6 snRNAs enter the spliceosome in the form of the U4/U6.U5 tri-snRNP, in which they are extensively base-paired. After splicing, U4 and U6 are released as singular snRNPs. In order to restore the tri-snRNP that can participate in another round of splicing, the U4/U6 hybrid must be recycled. The Prp24 protein in the yeast system, and its distant human homologue p110 have been identified as U4/U6 recycling factors and as specific protein components of U6 and U4/U6 snRNPs.

This work focuses on the recycling the minor spliceosome and of the minor spliceosomal U4atac/U6atac snRNP. Using a recycling assay coupled to splicing *in vitro*, p110 was demonstrated to function also as U4atac/U6atac recycling factor. p110 was detected in U6atac snRNP. In contrast to the major U4/U6, however, it is associated only at background levels with U4atac/U6atac snRNP. Interestingly, U4/U6 and U4atac/U6atac hybrids can be also disrupted independently of splicing *in vitro*. This process may provide a mechanism to control splicing rates *in vivo*.

The sequence elements of U6 and U6atac snRNA, required for interaction with p110 were identified: U6 nucleotides 38 to 57, and U6atac nucleotides 10 to 30 are sufficient for p110 binding. These regions are the most highly conserved between U6 and U6atac.

Finally, singular U6 snRNP was purified from HeLa S100 extract. It contained a novel protein, p85 (cDNA GenBank accession AK001239), identified by mass-spectrometric analysis. p85 localizes to the nucleoplasm, and associates not only with U6, but also with U2 snRNA, and 5S and 5.8S rRNAs. This protein is weakly similar to the yeast rRNA processing factor NOP4, which implicates a function of p85 in maturation of U2 and U6 snRNAs.

List of abbreviations:

ADP	adenosine diphosphate
ATP	adenosine triphosphate
BBP/SF1	branch point binding protein / splicing factor 1
BSA	bovine serum albumin
cDNA	copy-deoxyribonucleic acid
CMV	cytomegalovirus
CTD	C-terminal domain
DIG	digoxigenin
DMEM	Dulbecco's modified eagle medium
DMPC	dimethyl pyrocarbonate
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
EDTA	ethylenediaminetetraacetic acid
EGFP	enhanced green fluorescent protein
ESE	exonic splicing enhancer
GMP	guanosine monophosphate
GST	glutathione S-transferase
HEPES	N-2-hydroxyethylpiperazine
ISE	intronic splicing enhancer
ISS	intronic splicing silencer
NAD	nicotinamide adenine dinucleotide
NIS	nonimmune serum
NTC	nineteen complex; Prp19-associated complex
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate-buffered saline
PMSF	phenylmethylsulfonyl fluoride
PNK	polynuceotide kinase
PVA	polyvinyl alcohol
RNAi	RNA interference
rRNA	ribosomal ribonucleic acid
siRNA	small interfering ribonucleic acid
snRNA	small nuclear ribonucleic acid
snRNP	small nuclear ribonucleoprotein
RRM	RNA-recognition motif
SCAF	SR-related CTD-Associated Factor
SDS	sodium dodecyl sulfate
SMN	survival of motor neurons
TMG	trimethylguanosine
Tris	tris-hydroxymethylaminomethane
tRNA	transfer ribonucleic acid
U2AF	U2 snRNP auxiliary factor
UTP	uridine triphosphate

1. Introduction

1.1. Pre-mRNA processing

In eukaryotes primary RNA transcripts, synthesized by RNA polymerase II are subjected to a series of processing events in order to become functional messages (**Figure 1.1**). The mRNA precursors are 5' capped and their 3' ends are polyadenylated (Shatkin and Manley, 2000). The most dramatic changes however, occur inside the RNA chain. Substantial parts of the pre-mRNAs, called introns, are excised and the remaining RNA regions (exons) are joined together, coding for the protein sequence. This process of removal of the intervening introns is known as splicing.

After processing, the mRNA is exported to the cytoplasm. Intact 5' cap structure and poly(A) tail are required for targeting of the mRNA to the ribosome and the initiation of the translation. The cap and the poly(A) tail are also subject to regulatory mechanisms that control the mRNA stability and thereby the protein expression.

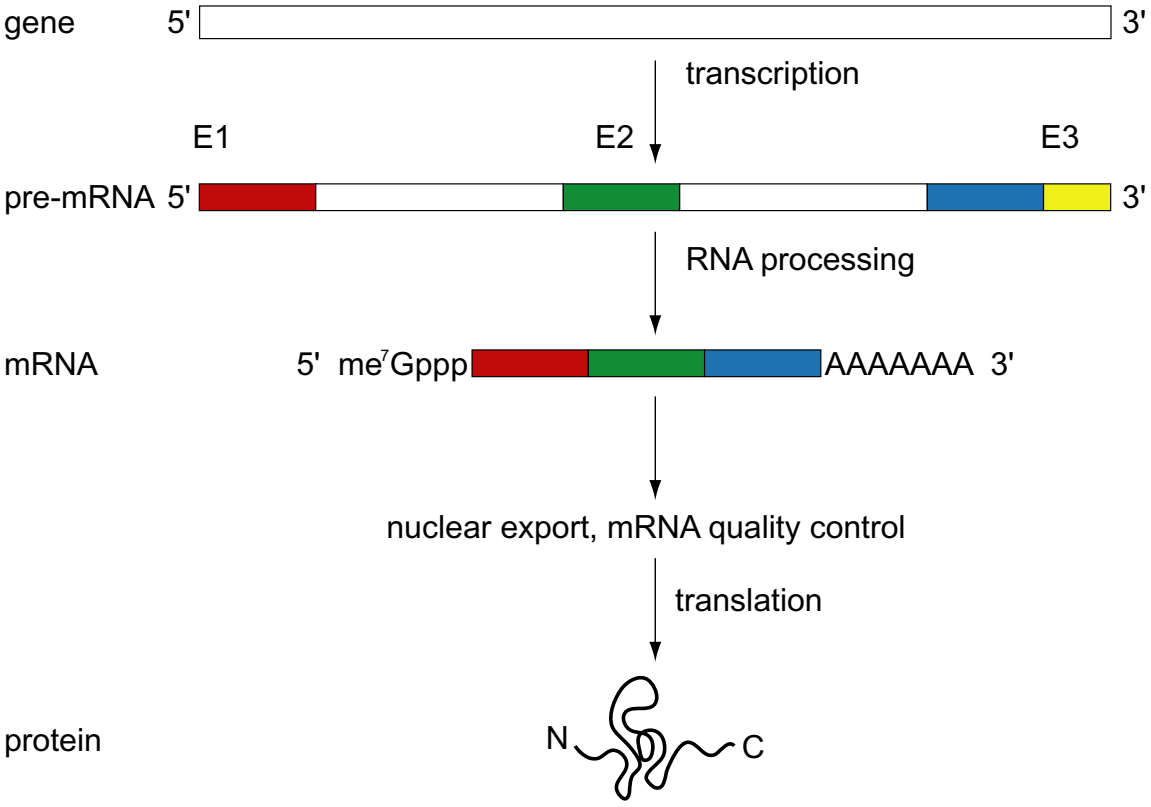


Figure 1.1. Eukaryotic gene expression.
Transcription, pre-mRNA processing, export and quality control are shown as separate consecutive events rather that coupled processes. The pre-mRNA exons are shown in colour, the introns are colourless. The sequence downstream of the polyadenylation signal is shown in yellow.

1.2. 5' capping and 3' polyadenylation

The 5' triphosphate of the pre-mRNA is first shortened to diphosphate and then GMP is transferred, generating an unusual triphosphate linkage: G(5')ppp(5')N. The guanine base is subsequently methylated at its N7 position (reviewed in Shuman and Schwer, 1995). The cap

can be further modified by methylating the 2' OH group of the ribose of the first and the second transcribed nucleotide.

3' end formation starts with cleavage of the RNA guided by an AAUAAA sequence 10 to 30 nucleotides upstream of the cleavage site and a less conserved GU- or U-rich stretch downstream of the cleavage site (Minvielle-Sebastia and Keller, 1999). Some pre-mRNAs require also U-rich motifs located upstream of the cleavage site. After cleavage a tail of 200-250 adenosines is added to the 3' end by poly(A) polymerase. Both the cap structure and the poly(A) tail are essential for the mRNA transport to the cytoplasm, the initiation of the translation, and the protection of the mRNA from exonucleases. They therefore serve to identify the RNA as a messenger.

1.3. Coupling transcription, pre-mRNA processing, and mRNA export

Although transcription and splicing can occur independently *in vitro*, they appear to be closely coupled *in vivo*. The key factor in this integration is the C-terminal domain (CTD) of the largest subunit of RNA polymerase II (reviewed in Manley, 2002; Rosonina and Blencowe, 2002). Factors involved in capping and 3' end polyadenylation are recruited by the CTD to pre-mRNA during transcription (Shatkin and Manley, 2000). A number of proteins, called SR-like CTD associated factors (SCAFs) contain a CTD-interacting domain as well as an RS domain (Yuryev *et al.*, 1996). It is believed that these proteins can recruit splicing factors to the transcribing polymerase. Basic spliceosome components such as U1 and U2 snRNPs are associated with TAT-SF1, a component of the transcription elongation factor P-TEFb (Fong and Zhou, 2001). The association of this processing factor with the transcription machinery ensures accurate and efficient processing of the nascent transcripts. CTD deletion prevents efficient capping, splicing, and 3'end cleavage *in vivo* (McCracken *et al.*, 1997a; McCracken *et al.*, 1997b). Moreover, the splicing pattern may depend on the transcription elongation rate (de la Mata *et al.*, 2003). Surprisingly it has been found that splicing (Fong and Zhou, 2001) and 5' capping (Myers *et al.*, 2002) in turn can influence transcription. There is also a mutual facilitation between the different processing events. Capping has been found to enhance the recognition of first intron's 5'splice site (Lewis *et al.*, 1996). Similarly, in the case of the 3' terminal exons the factors involved in cleavage and polyadenylation of pre-mRNA can play a role in directing the splicing machinery to the 5' splice site (Niwa and Berget, 1991). On the other hand, the presence of an upstream intron can stimulate polyadenylation (Niwa *et al.*, 1990).

The spliced mRNA is marked by a protein complex, deposited 20-24 nucleotides upstream of the exon-exon junction (Le Hir *et al.*, 2000). This so-called exon-junction complex (EJC) links pre-mRNA splicing with nuclear export and mRNA quality control. Two EJC proteins, UAP56 and Aly, are necessary for efficient mRNA export. Aly interacts directly with the mRNA export receptor TAP/p15 (Rodrigues *et al.*, 2001; Strasser and Hurt, 2001; Stutz *et al.*, 2000; Zhou *et al.*, 2000).

mRNA is subjected to quality control, which eliminates the messages with premature stop-codon through the nonsense-mediated decay (NMD) pathway. Two other EJC proteins, Y14 and RNPS1, are involved in that process (revieved in Singh and Lykke-Andersen, 2003). They recruit Upf

2 and Upf 3 proteins. Normally the first translating ribosome displaces the EJC-Upf complexes (Dostie and Dreyfuss, 2002). However, if a stop-codon is found more than 50-55 nucleotides upstream of the last exon-exon junction, at least one such complex will remain associated with the mRNA. The third protein of the NMD complex - the translation termination factors eRF1 and eRF3 - associated Upf1 (Czaplinski *et al.*, 1998) - is then recruited to the EJC-Upf complex. An activated NMD complex is formed, and the mRNA is targeted for rapid decay.

1.4. Splicing

Introns have been identified not only in pre-mRNA, but also in ribosomal and transfer RNAs (rRNA and tRNA). They are removed by three different mechanisms of splicing: pre-mRNA splicing, self-splicing, and pre-tRNA splicing.

1.4.1. Self-spliced introns and splicing of tRNA precursors

The self-spliced introns are two types of ribozymes, which catalyze their own excission. There are two groups of self-spliced introns, named group I and group II, with distinct secondary structures and mechanisms of catalysis. Both types of introns can self-splice *in vitro*, however, it has been demonstrated that *in vivo* many of them depend on *trans*-acting protein factors for efficient splicing (reviewed in Chech and Golden, 1999).

Group I introns are abundant in fungal and plant mitochondrial and chloroplast genes, but are also found in nuclear rRNA genes of lower eukaryotes, as well as in bacteriophage and several tRNA genes in eubacteria. Splicing is initiated by the 3' hydroxyl group of an external guanosine, which attacks the phosphodiester bond at the 5'end of the intron. Subsequently the released 3' hydroxyl group of the exon attacks the 3' end of the intron, resulting in ligation of the exons and release of the intron in linear form with the guanosine attached at its 5' end.

Group II introns are abundant in organellar genomes of plants and lower eukaryotes, in bacterial genomes, but have not yet been found in organelles of higher eukaryotes, or in nuclear genomes. Organellar group II introns are typically found in highly conserved genes such as cytochrome oxidase or rubisco subunit genes, and there are often many introns in a gene. In contrast to group I splicing, the excision of group II introns does not require an external guanosine, instead the reaction is initiated by a nucleophilic attack of the 2'hydroxyl group of the branch adenosine. As a result, a 2'-5' phosphodiester bound is formed between the branch adenosine and the 5' end of the intron, and a lariat intermediate is created. Then the released upstream exon attacks through its 3' hydroxyl group the phosphodiester bond at the boundary of the intron and the downstream exon. Exons are joined together and the intron-lariat is released.

Both groups of self-spliced introns possess conserved secondary and tertiary structures that are essential to bring the exons in close proximity and to form the catalytic center.

The introns found in the bacterial tRNA genes are self-spliced. However, there are introns found also in some of the tRNA genes of *Eukaryota* and *Archaea*, most frequently interrupting the anticodon loop. These introns are short – typically between 14 and 60 nucleotides and are spliced

through a three-step pathway (reviewed in Abelson *et al.*, 1998). Each step is catalyzed by a distinct enzyme, which can work interchangeably on all the substrates:

In the first step the pre-tRNA is cleaved at both splice sites by an endonuclease, which forms 5'-hydroxyl and 2'-3'-cyclic phosphate ends.

Then the two half-tRNA molecules are ligated by tRNA ligase. The reaction is carried out through the following mechanism: the 3' cyclic phosphate is hydrolyzed to 2'-phosphate and 3'-hydroxyl group; then the 5' end is phosphorylated by GTP; AMP is transferred to the 5'-phosphate; formation of the 5'-3'-phosphodiester bond proceeds, and AMP is released.

Finally the remaining 2'-phosphate is removed by an NAD-dependent phosphotransferase, which generates the unusual ADP-ribose 1"-2"-cyclic phosphate.

1.4.2. Pre-mRNA splicing

One of the most prominent features of eukaryotic genes is their discontinuity. With very few exceptions, notably the histone genes, the coding sequences of almost all of eukaryotic genes are disrupted by introns. The analysis of the working draft of the human genome showed that on average introns account for 95% of the pre-mRNA (Lander *et al.*, 2001). In higher eukaryotes exons are typically short (100-200 nucleotides), while the introns are much longer - around 3 500 nucleotides (Deutsch and Long, 1999). However there are introns that span hundreds of thousand nucleotides. In average there are 7-8 introns per human protein-encoding gene. The human dystrophin gene, the largest gene known so far, contains 79 exons, scattered over 2.3 megabases, to give an mRNA of approximately 14 kilobases (Nobile *et al.*, 1997).

Unlike the self-spliced introns and those of tRNA precursors, pre-mRNA introns vary greatly in size and do not possess conserved secondary structurures. The pre-mRNA introns are spliced by a large machinery, built up of over one hundred *trans*-acting factors, which recognize and assemble on conserved *cis* elements found both on the introns and the exons.

1.4.2.1. Conserved *cis* elements in pre-mRNA; GU-AG and AU-AC introns

There are three sequence elements - at the 5' and 3' splice sites, and the branch site (**Figure 1.2**) - that mark the introns and are essential for their removal (Reed and Palandjian, 1997). Two distinct types of introns are identified in higher eukaryotes, spliced by separate splicing machineries.

The first, major type of introns represents the vast majority of introns in the human genome. The 5' splice site conforms to the consensus YRG/GURRGU (the slash denoting the exon - intron border, R a purine (A or G) base, Y a pyrimidine (C or T) base). The two underlined nucleotides until recently were believed to be invariant. However, a database search showed that in 1% of all introns the first dinucleotide is GC instead of GU (Thanaraj and Clark, 2001).

The 3' splice site consensus is Y_{12}NYAG/, i.e. a highly conserved AG dinucleotide, preceded by a polypyrimidine stretch. The third element is the branch point (YNYURAY) which is located 18 to 20 nucleotides upstream of the 3' splice site. The highly conserved branch adenosine is

underlined.

The second, minor intron type has more conserved 5' splice site and branch site sequences, and 3'splice site that lacks the polypyrimidine tract. The consensus sequences of these elements are as follows: 5' splice site: /RUAUCCUUU; 3' splice site: YYA(G/C)/; branch point: UCCUURACY (Dietrich *et al.*, 1997; Hall and Padgett, 1994, reviewed in Patel and Steitz, 2003). Those introns could have GU-AG or AU-AC termini and they are much less abundant - only 404 such introns are found in the human genome (Levine and Durbin, 2001).

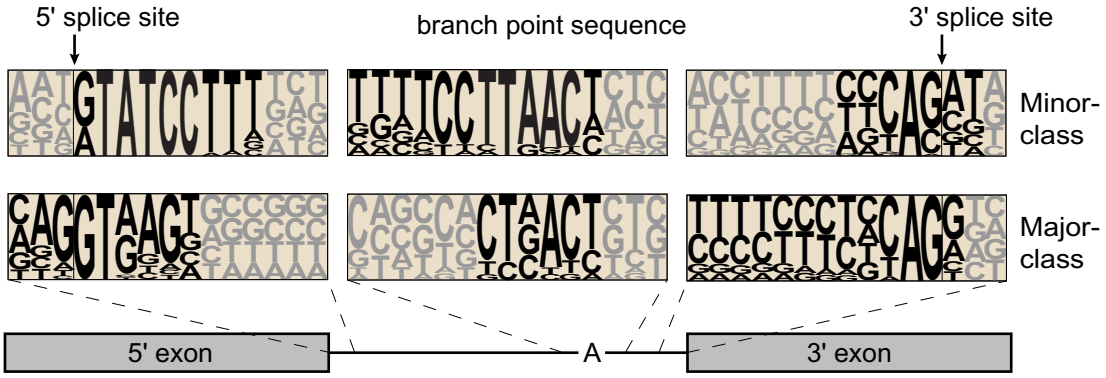


Figure 1.2. Conserved *cis*-elements of major-class and minor-class introns as shown by Patel and Steitz (2003).

The consensus sequences of the 5' splice site, branch site and 3' splice site are shown and the letter heights at each position represent the frequency of occurrence of the corresponding nucleotides at that position. The positions that are thought to be involved in intron recognition are in black, and the others in gray.

An interesting exception is the budding yeast *Saccharomyces cerevisiae*. Due to the massive wave of retrotransposition events, most of the introns in this organism have been eliminated. The remaining 217 introns strongly adhere to the consensus sequences, (A/C)AG/GURAGU and Y_{12}CAG/ for the 5' and 3' splice sites, respectively (Staley and Guthrie, 1998). In addition, the branch point is virtually invariant (UACUAAC) and is located 16 to 19 nucleotides upstream of the 3' splice site. There are no minor-type introns found in the genome of the the budding yeast. The classical splicing signals described above are often insufficient for recognition of all introns. It has been estimated that these sequence elements provide less than 50% of the information necessary for accurate removal of human introns (Lim and Burge, 2001). Additional elements, typically found nearby weak splice sites, provide the required information for intron recognition and splicing. These auxiliary signals are classified according to their location – exonic or intronic, and their functional effect on splicing – activation or repression.

Exonic splicing enhancers (ESEs) have been identified by different approaches in a number of studies (Coulter *et al.*, 1997; Fairbrother *et al.*, 2002; Liu *et al.*, 2000; Liu *et al.*, 1998; Schaal and Maniatis, 1999; Woerfel and Bindereif, 2001). Two classes of ESEs have been identified: purine-rich and A/C-rich. The ESEs are highly degenerate, which prevents them from interfering with the coding capacity of the exons. As in the case of the constitutive *cis*-elements discussed above, however, they carry very low information content. To overcome this, ESEs act cooperatively (Hertel and Maniatis, 1998; Smith and Valcarcel, 2000).

There is evidence that purine-rich ESEs are functional on both major- and minor- type introns (Wu and Krainer, 1998).

A number of intronic splicing enhancers (ISEs) and silencers (ISSs) have also been identified (for example, Fairbrother and Chasin, 2000; Hui *et al.*, 2003, listed by Ladd and Cooper, 2002). Their sequences are also degenerate. The ISSs appear to be widespread in introns (Fairbrother and Chasin, 2000).

1.4.2.2. The mechanism of pre-mRNA splicing

The intron excision (**Figure 1.3**) proceeds through two transesterification reactions (Padgett *et al.*, 1984; Ruskin *et al.*, 1984). Initially a nucleophilic attack of the 2' hydroxyl group of the branch point adenosine at the phosphodiester bond of the 5' splice site leads to its cleavage. As a result a 5'-2' phosphodiester bond forms between the branch point adenosine and the terminal guanosine. In the second step the newly formed 3' hydroxyl group of the 5' exon attacks the phosphodiester bond of the 3' splice site, resulting in a 5'-3' phosphodiester bond between two exons and the release of the lariat intron.

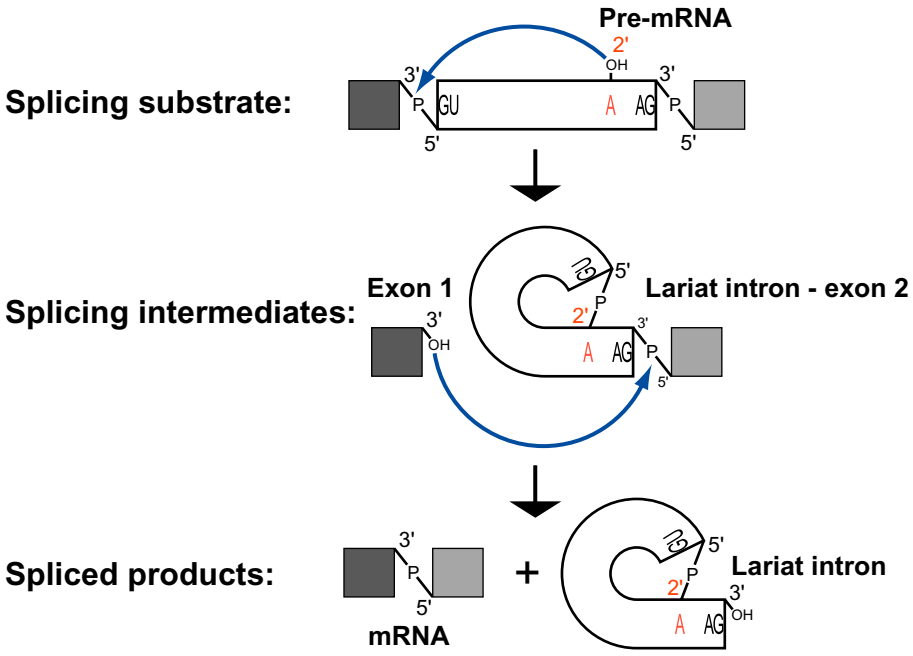


Figure 1.3. The mechanism of pre-mRNA splicing.
The upstream and downstream exons, the intron, the phosphodiester bonds that are cleaved or formed during the reaction, and the reactive hydroxyl groups are schematically represented. The intermediates and the splicing products are also shown. The highly conserved branch adenine is red; both transesterification reactions are indicated by blue arrows. The dinucleotides at the intron termini might be AU-AC as well, but the highly prevalent GU-AG are shown.

1.4.2.3. snRNPs and protein factors involved in pre-mRNA splicing

Pre-mRNA splicing occurs within a large ribonucleoprotein complex, the spliceosome. There are two distinct spliceosomes involved in the removal of introns. The spliceosome consists of five small nuclear ribonucleoprotein complexes (snRNPs), named according to their RNA component, and over a hundred non-snRNP protein factors (Hartmuth *et al.*, 2002; Zhou *et al.*, 2002).

U1, U2, U4, U5, and U6 snRNPs participate in major intron splicing. U11, U12, U4atac, U5, and U6atac snRNPs (Tarn and Steitz, 1996a) are involved in the excision of the minor-type introns

(**Table 1.1**). Because of the catalytic components involved in their excision, the major introns are also termed U2-dependent, the minor introns U12-dependent (Sharp and Burge, 1997).

major-class introns :	U1	U2	U4	U6	U5
minor-class introns :	U11	U12	U4atac	U6atac	U5

Table 1.1. snRNPs involved in removal of the major-class and minor-class pre-mRNA introns. The snRNPs grouped in columns act analogously in major and minor spliceosomes. U5 snRNP is common for both of them.

snRNP	protein components	references
U1	Sm, U1-70K, U1A, U1C	Patton *et al.*, 1989
12S U2	Sm, U2A', U2B"	Kleinschmidt *et al.*, 1989
17S U2	Sm, U2A', U2B", SF3B (160, 150, 120, 49, 14b, 10K), SF3A (120, 66, 60K), U2AF65, U2AF35K, ASF/SF2, hPrp5, hPrp43,SPF45, SPF31, SPF30, SR140,p14, BRAF35, PUF60, CHERP, Hsp75, Hsp60	Will *et al.*, 2002
U11/U12	Sm, U11-35K, SF3B (160, 150, 120, 49K), U11/U12 (65, 64, 48, 35, 28.5, 20K)	Will *et al.*, 1999
20S U5	Sm, 220, 200, 116, 102, 100, 40, 15K	Achsel *et al.*, 1998
35S U5	Sm, 220, 200, 116, 40K, hPrpr19, CDC5, hSyf1, hSyf3, hIsy1, PRL1, SPF27, SKIP, hECM2, G10, GCIP p29, AD-002, KIAA0560, MGC23918, Cyp-E, PPIase-like1, hPrp17, DDX3, Hsp73, FLJ10206, KIAA0773, C25A1.10p, S100 A8, S100 A9	Makarov *et al.*, 2002
U4/U6	Sm, 90, 61, 60, 20, 15.5K	Nottrott *et al.*, 2002
U4/U6.U5	Sm, 90, 61, 60, 20, 15.5K, 220, 200, 116, 102, 100, 40, 15K, LSm, 110, 65, 27K,	Behrens and Lührmann 1991
U4atac/U6atac. U5 *	Sm, 61, 60K, 220, 116, 100, 40K, LSm, 110, 65, 27K,	Schneider *et al.*, 2002
U6	LSm, p110	Achsel *et al.*, 1999 Bell *et al.* 2002

Table 1.2. The protein composition of the human snRNPs.
The characterized human snRNPs and their protein components are listed; the references are indicated on the right side. The Sm complex is built up of seven Sm proteins – Sm B/B', D1, D2, D3, E, F, and G. The LSm complex contains proteins LSm 2 to LSm8. Some proteins that participate in different snRNP complexes are highlighted in colours.
*The U4atac/U6atac.U5 complex has not been purified biochemically; the presence of the U4/U6.U5-specific proteins is only confirmed by immunoprecipitation.

Each snRNP contains a single small nuclear uridine-rich RNA (snRNA) and several proteins - specific or common to all spliceosomal snRNPs except for U6. The basic spliceosomal snRNPs also form larger complexes, stabilized by RNA-RNA (U4/U6 and U4atac/U6atac snRNPs) or protein-protein interactions (U4/U6.U5, U4atac/U6atac.U5, and U11/U12 snRNPs). A detailed protein composition of the spliceosomal snRNPs is given on **Table 1.2.**

There is a direct functional correspondence between the snRNPs that participate in splicing of the two types of introns: U11 is the functional analogue of U1, U12 that of U2, U4atac and U6atac correspond to U4 and U6, respectively. U5 is present in both spliceosomes. The snRNPs that form the major spliceosome are approximately 2 orders of magnitude more abundant than those involved in the formation of the minor spliceosome (Yu *et al.*, 1999). The corresponding major and minor snRNAs have conserved secondary structures and participate in analogous interactions during splicing, however, they are not similar in sequence (Tarn and Steitz, 1996).

Only limited information is available regarding the protein constituents of the minor spliceosome. Most of the snRNP proteins are shared between the major and minor snRNPs (Schneider *et al.*, 2002), however, there are exceptions such as the U11-specific 35K protein, which is the minor spliceosome analogue of U1-70K (Will *et al.*, 1999), see **Table 1.2.**).

A number of other proteins, such as SR proteins, hnRNPs, and RNA helicases are not stably associated with any of the snRNPs, but also participate in pre-mRNA splicing.

The SR proteins (Zahler *et al.*, 1992) constitute a family of splicing factors with a characteristic structural organization: RNA-binding motifs (RRMs or RBDs) in their N-terminal part, and a variable-length C-terminal serine/arginine rich (RS) domain (reviewed in Kramer, 1996; Valcarcel and Green, 1996). The N-terminal part of the SR proteins is involved in RNA recognition and binding, while the RS motif is involved in protein-protein interactions. Ten human SR proteins have been found so far, a number of other proteins, termed SR-like proteins are structurally and functionally related.

All of the SR protein family members have been shown to activate splicing *in vitro*. Several distinct mechanisms of their action have been proposed, but perhaps the best-characterized function of the SR proteins is to recognize ESEs (Liu *et al.*, 2000; Liu *et al.*, 1998) and to bridge exons and spliceosome components (Blencowe *et al.*, 2000; Eperon *et al.*, 1993; Jamison *et al.*, 1995; Kohtz *et al.*, 1994; Liu *et al.*, 2000; Liu *et al.*, 1998; Tarn and Steitz, 1995; Zahler and Roth, 1995). ESE sequences recognized by the SR and SR-like proteins are short (5 to 8 nucleotides) and highly degenerate (Cooper, 1999; Coulter *et al.*, 1997). As SR and SR-like proteins attach to multiple nearby sites, their binding is stabilized by protein/protein interactions, forming a complex that marks the exon.

The hnRNP proteins are another group of proteins that bind to pre-mRNA and play an important role in splicing (reviewed in Krecic and Swanson, 1999). They were first described as packaging factors. The binding of the hnRNP proteins to RNA shows some sequence preference which could serve to ‘phase’ the packaging of the pre-mRNA. hnRNP proteins in general antagonize the splicing-activating function of the SR and SR-like proteins. However, hnRNPs can also facilitate splicing by binding to ISEs (Hui *et al.*, 2003). Exon recognition and splice site selection

is a combinatorial process involving both positive and negative signals provided by SR proteins and hnRNP proteins respectively.

The RNA helicases (or RNA-dependent ATPases) are members of a large protein family with a conserved domain that includes the motif DExD/H (reviewed in Linder *et al.*, 2001; Silverman *et al.*, 2003; Tanner and Linder, 2001). These proteins can unwind double-stranded RNAs in an ATP-dependent manner, and it has been recently shown that they work in a similar way to disrupt RNA-protein complexes (Jankowsky *et al.*, 2001).These enzymes are involved in virtually every aspect of the expression of genetic information, including transcription, RNA modification and export, NMD, ribosome biogenesis and translation. DExD/H helicases are essential factors at every step of splicing – from the recognition of the splice sites, the assembly of the spliceosome, through the catalytic steps and to the dissociation of the spliceosome complex. At least eight RNA helicases are involved splicing. Some of them like U5-220K (yeast Brr2) are stable protein components of the spliceosomal snRNPs, others (hPrp28, hPrp43) associate only transiently with the spliceosome to facilitate a specific conformational transition or catalytic step of splicing.

1.4.2.4. The spliceosome cycle; spliceosome assembly and catalysis

There are three stages of the catalytic excision of pre-mRNA introns (**Figure 1.4.**): First the spliceosome **assembles** on pre-mRNA. During that stage the boundaries of the intron to be removed are recognized. Then the spliceosome is triggered towards its **catalytically active** state, the two transesterification reactions quickly follow, i.e. the intron is excised, and the exons are joined together. Finally the spliceosome disassembles and the catalytic components of splicing are **recycled** to join another round of splicing.

The spliceosome is a dynamic ribonucleoprotein complex; extensive structural rearrangements of the RNA network (**Figure 1.5.**) and the protein components occur throughout the spliceosome cycle.

The spliceosome assembly and catalysis proceed through the formation of four distinct complexes. Each complex transition is ATP-dependent and accompanied by significant structural rearrangements of the individual components.

Spliceosome assembly is initiated by the ATP-independent formation of the early (E) complex: First the U1 snRNP associates with the 5’ splice site by hybridization of the 5’end of U1 snRNA with the intron nucleotides +1 to +6. Interestingly, the U1C protein recognizes a consensus sequence, very similar to the one of the 5’ splice site (Du and Rosbash, 2002; Rossi *et al.*, 1996), which led to the hypothesis that U1C/5’ splice site interaction precedes pre-mRNA/U1 snRNA base pairing. Another factor that facilitates the recruitment of U1 snRNP to the 5’ splice site is the SR protein ASF/SF2: It binds to the pre-mRNA and to U1-70K protein (Kohtz *et al.*, 1994).

The polypyrimidine tract and the 3’ terminal dinucleotide are recognized by the auxiliary proteins U2AF65 and U2AF35, respectively. Binding of the two U2AF subunits to the 3’ splice site is facilitated by the U1 snRNP located at the 5’ splice site of the exon (Cunningham *et al.*, 1995; Hoffman and Grabowski, 1992) and could also be assisted by SR proteins (Wang *et al.*, 1995; Wu and Maniatis, 1993).

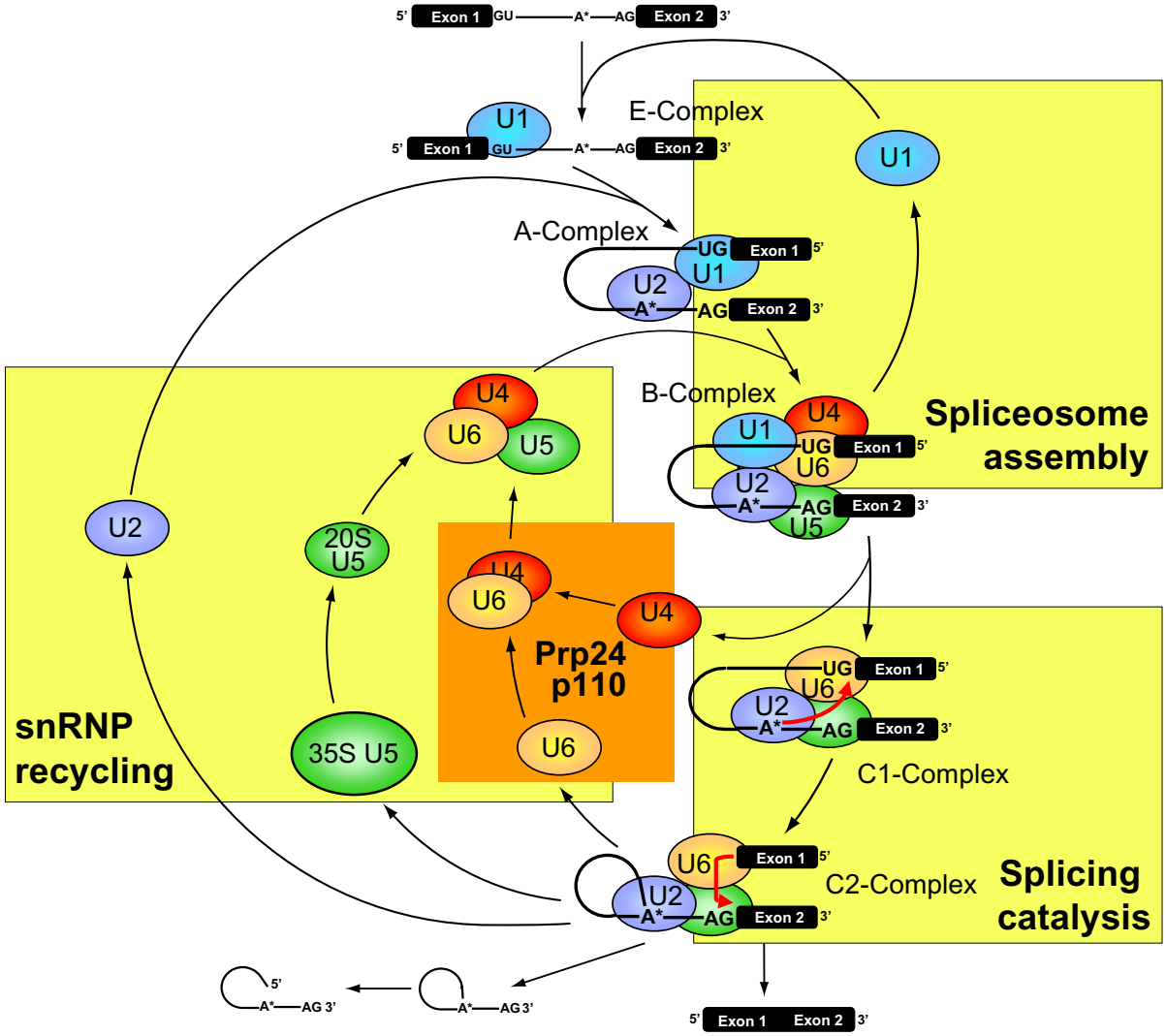


Figure 1.4. The spliceosome cycle.
The spliceosomal complexes formed during the spliceosome assembly and splicing catalysis, as well as the known steps of snRNP recycling after dissociation of the spliceosome are schematically represented. The position of U4/U6 snRNP recycling within the spliceosome cycle is indicated by an orange box. The U4/U6 recycling factors in yeast (Prp24) and human (p110) system are also shown.

The branch point is recognized by the splicing factor BBP/SF1 (Berglund *et al.*, 1997). SF1 interacts with U2AF and this interaction is essential for its binding to the branch point (Berglund *et al.*, 1998). Later studies have shown that the U2 snRNP is also included in the E-complex through interactions with U1 snRNP, U2AF65 and possibly SF1 (Das *et al.*, 2000). In addition, the U4/U6.U5 tri-snRNP can associate with the first exon near the 5' splice site in the E-complex in an ATP-dependent manner (Maroney *et al.*, 2000).

The transition to the next complex (A complex) involves RNA duplex formation between the U2 snRNP and the branch site with the branchpoint adenosine bulging out of the duplex. This conformational transition is ATP-dependent and assisted by the RNA helicases UAP56 (Fleckner *et al.*, 1997) and Rrp5 (Xu *et al.*, 2004). The U2 interaction is further stabilized by binding of the U2-specific multiprotein factors SF3a and SF3b upstream of the branch site (Gozani *et al.*, 1996). A characteristic feature of the early spliceosome is the interaction between the U1 snRNP bound to the 5' splice site and the U2 snRNP/U2AF complex at the 3' splice site. SC35 (Fu and Maniatis, 1992; Wu and Maniatis, 1993), a member of the SR protein family, and Prp5 (Xu *et*

al., 2004) have been shown to mediate that interaction. In case of the minor spliceosome, U11 and U12, the functional analogues of U1 and U2, form a stable U11/U12 complex (Prasad *et al.*, 1999; Wassarman and Steitz, 1992) due to interaction between the U11-specific 59K and the U12-specific 65K proteins (Will *et al.*, 2003). The U11/U12 particle associates with the U12-dependent introns, simultaneously recognizing the 5' splice site and the branch site.

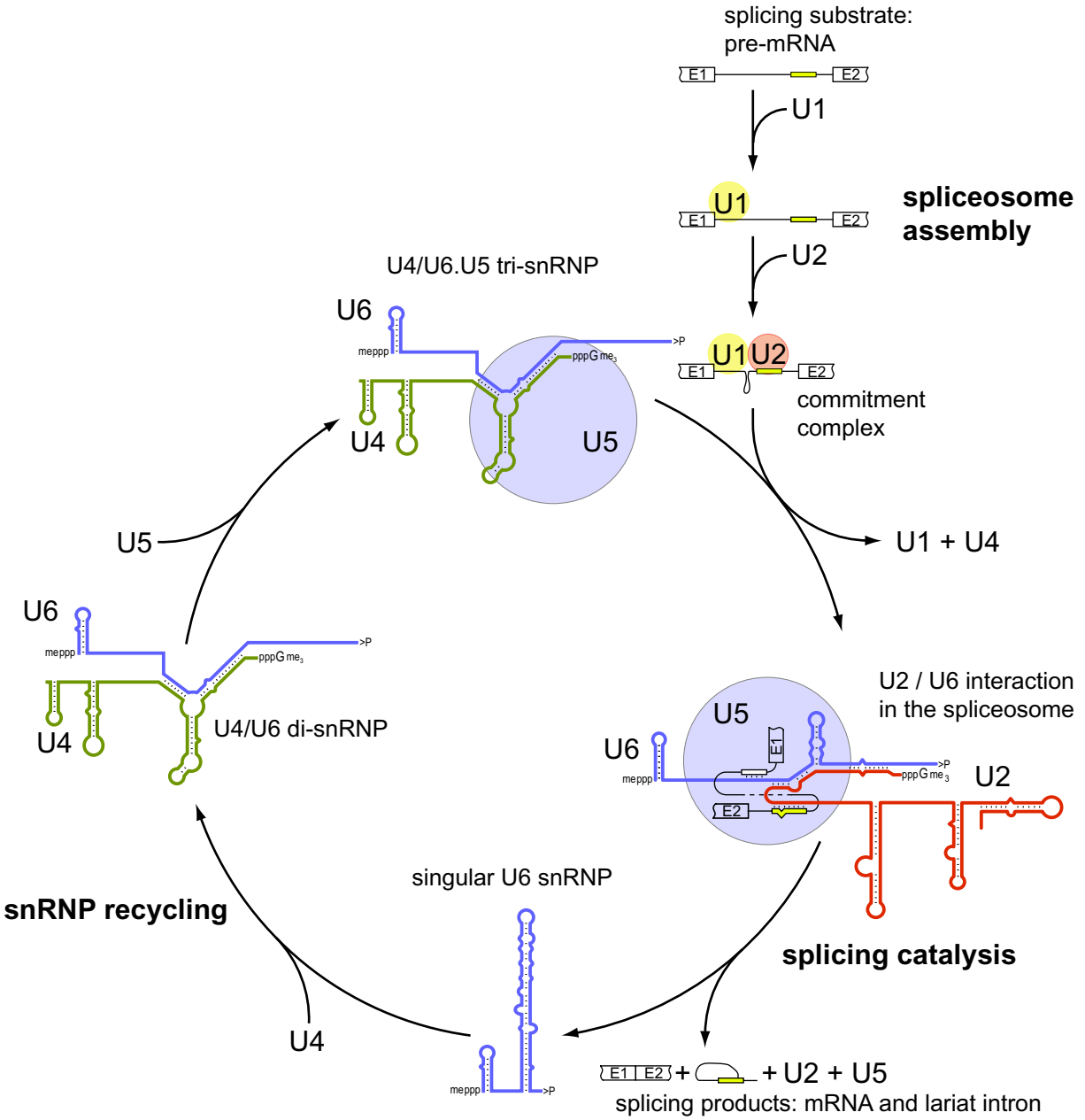


Figure 1.5. The spliceosome is a dynamic RNA network.
The dynamic nature of the spliceosome becomes particularly apparent when following the U4 and U6 snRNPs throughout the spliceosome cycle: Most of U4 and U6 are stably base-paired with each other; before entering the assemblying spliceosome, the U5 snRNP joins the U4/U6 snRNP, resulting in the U4/U6.U5 tri-snRNP. In the active spliceosome the U4-U6 base-pairing is disrupted and isomerizes to a mutually exclusive U6-U2 interaction. Finally, after splicing catalysis, the spliceosome disassembles, the U4 and U6 snRNPs are released in their singular forms and U4/U6 snRNP recycled. Significantly, all these dynamic RNA-RNA interactions are conserved between the major and minor snRNAs.

The A complex progresses to B complex when the large pre-assembled U4/U6.U5 snRNP (also known as the tri-snRNP) enters the spliceosome. In the tri-snRNP U4 and U6 snRNA are extensively base-paired, while U5 is associated by protein-protein interactions involving the

U4/U6 specific 61K and the U5-specific 102K (Makarova *et al.*, 2002; Schaffert *et al.*, 2003). Docking of the tri-snRNP onto the A complex is mediated by SPF30, an essential splicing factor that interacts also with U2 snRNP (Rappsilber *et al.*, 2001).

In contrast to the stepwise spliceosome assembly described above, an alternative model has been proposed. A large 45S snRNP complex, containing U4/U6.U5, U2, and U1 snRNPs has been found under low-salt conditions in yeast (Stevens *et al.*, 2002) and HeLa nuclear extracts (Malca *et al.*, 2003). This complex, called the penta-snRNP, contains the majority of the protein factors found in purified spliceosomes, with the notable exception of some SR proteins and most of the RNA-dependent ATPases, which drive the conformational changes required for the activation of the spliceosome, splicing catalysis, and the subsequent dissociation of the complex (see below). It has been shown that the purified penta-snRNP complex is able to recognize the 5' splice site, forming an RNA-RNA duplex with it via the 5' end of U1 snRNA, and promoting splicing upon addition of soluble protein factors. At present it is not known which of these two modes of spliceosome assembly is preferentially used. As U1 snRNP exists in excess over U2 and even larger excess over the tri-snRNP, it is likely that both assembly pathways could be utilized at the same time, possibly with some substrate-specific preferences.

Following the recruitment of all snRNPs to the pre-mRNA, either by the stepwise U1-U2-tri-snRNP pathway or by direct association of the penta-snRNP, two important RNA rearrangements take phase: First, U1 dissociates and a new duplex between the 5' splice site and the U6 snRNA is formed. Genetic studies in yeast identified Prp28, a DExH/D box helicase to mediate the exchange of U1 by U6 at the 5'splice site (Staley and Guthrie, 1999). Second, the U4/U6 duplex is disrupted by another helicase - the U5-specific 200K protein (Brr2) and U4 snRNP leaves the spliceosome (Kim and Rossi, 1999; Laggerbauer *et al.*, 1998; Raghunathan and Guthrie, 1998a). The 5' splice site is brought close to the branch point and the 3' splice site through U2/U6 snRNA base-pairing and through the interaction of U5 snRNP with both exons near the splice sites (Chiara *et al.*, 1997). After the release of U4 snRNP, the spliceosome is stabilized by the multiprotein complex NTC, which includes Prp19 and CDC5 (Chan *et al.*, 2003). That induces a rearrangement of the 5' splice site/U6 duplex, and destabilization of the U6-specific LSm heptamer, which also leaves the spliceosome.

At this point, the catalytically active C complex is formed, in which U6 snRNA is annealed with the 5' splice site, and U2 snRNA - with the branch site; at the same time, U2 and U6 are also base-paired with each other. The first catalytic step of the splicing is triggered by the RNA helicase Prp2 (Kim and Lin, 1996; Roy *et al.*, 1995), creating the intron lariat. The two exons are kept in close proximity by the U5 snRNP. Another RNA helicase, Prp16, is required for the second final step of the splicing (Ortlepp *et al.*, 1998; Schwer and Guthrie, 1991; Zhou and Reed, 1998). The exact mechanism of the catalysis is still poorly understood. However there is evidence suggesting the reaction is catalyzed by RNA and is assisted by the U5-specific 220K protein (Prp8) (Collins and Guthrie, 1999), (rewiewed in (Collins and Guthrie, 2000). The conservation of the RNA secondary structure of the snRNAs and more remarkably, the absolute conservation of the nucleotides in the proposed catalytic center strongly supports this hypothesis.

In addition a Mg^{2+} ion is believed to be involved in catalysis, coordinating highly conserved nucleotides in the U6 snRNA.

Interestingly, an abundant U2/U5/U6 snRNP complex is found in *Schizosaccharomyces pombe* extracts (Huang *et al.*, 2002). This is an indication that the slowest step of splicing in this organism might be the completion of the second step and/or release of the spliced exons.

After this step the spliceosome dissociates. The RNA helicases Prp22 (Company *et al.*, 1991; Schwer and Meszaros, 2000) and Prp43 (Arenas and Abelson, 1997) are involved in the release of mRNA and the lariat-intron, respectively. The intron-lariat is debranched (Chapman and Boeke, 1991; Ruskin and Green, 1985) and rapidly degraded (Chen and Okayama, 1987). The U2 and U6 snRNPs are released in their singular state. It is currently unknown how the U2/U6 duplex is disrupted. The U5 snRNP leaves the spliceosome in the form of a large 35S particle (Makarov *et al.*, 2002), which lacks some of the U5-specific proteins found the 20S U5 snRNP as well as in the tri-snRNP (see **Table 1.2** for details), but contains non-snRNP factors found in the activated spliceosome and also components that until now have not been identified in other complexes related to splicing.

Finally, the released snRNP are recycled to complexes that are able reenter the spliceosome cycle.

1.4.2.5. The recycling step of the spliceosome cycle

As it has been discussed above, after intron excision, the spliceosomal snRNPs are released as complexes that are not able to participate directly in assembly of another spliceosome. The singular U4, U5, and U6 snRNPs are recycled to reform the splicing-competent U4/U6.U5 snRNP. Recycling is the least well-studied phase of the spliceosomal cycle. The picture of the recycling pathway is just beginning to emerge. Until now it is not known in what form the U1 snRNP is released. Could it participate directly in another round of splicing or a remodeling is required? No data are available also about the U2 snRNP, although there is an indication that it is released as a larger 25S particle (Makarov *et al.*, 2002) after dissociation of the C complex. The components of the tri-snRNP are dissociated at different stages of the spliceosome cycle as singular U4, U5, and U6 snRNPs. The postspliceosomal U4 complex has not been studied so far. The U5 snRNP exits the spliceosome in the form of large 35S complex. First it has to be converted to the 20S U5 snRNP, which is capable to participate in the tri-snRNP formation. During spliceosome activation the LSm proteins are displaced from U6 (Chan *et al.*, 2003). It is not known how and at what stage they reassociate with the 3'end of U6. Finally, the tri-snRNP is restored from 20S U5 snRNP and 11S U4/U6 snRNP (di-snRNP). Two protein factors – the U5-specific 102K and the U4/U6-specific 61K interact with each other and are required for tri-snRNP formation (Golas *et al.*, 2003; Makarova *et al.*, 2002). Prior to that, the U4/U6 heteroduplex needs to be restored.

The reassociation of U4 and U6 is perhaps the most extensively studied recycling step (see below), however, almost exclusively in the yeast system. Only recently the recycling of human U4/U6 snRNP has been brought into focus. The latter process involves a dramatic rearrangement of the secondary structures of U4 and U6 snRNA (**Figure 1.6.**).

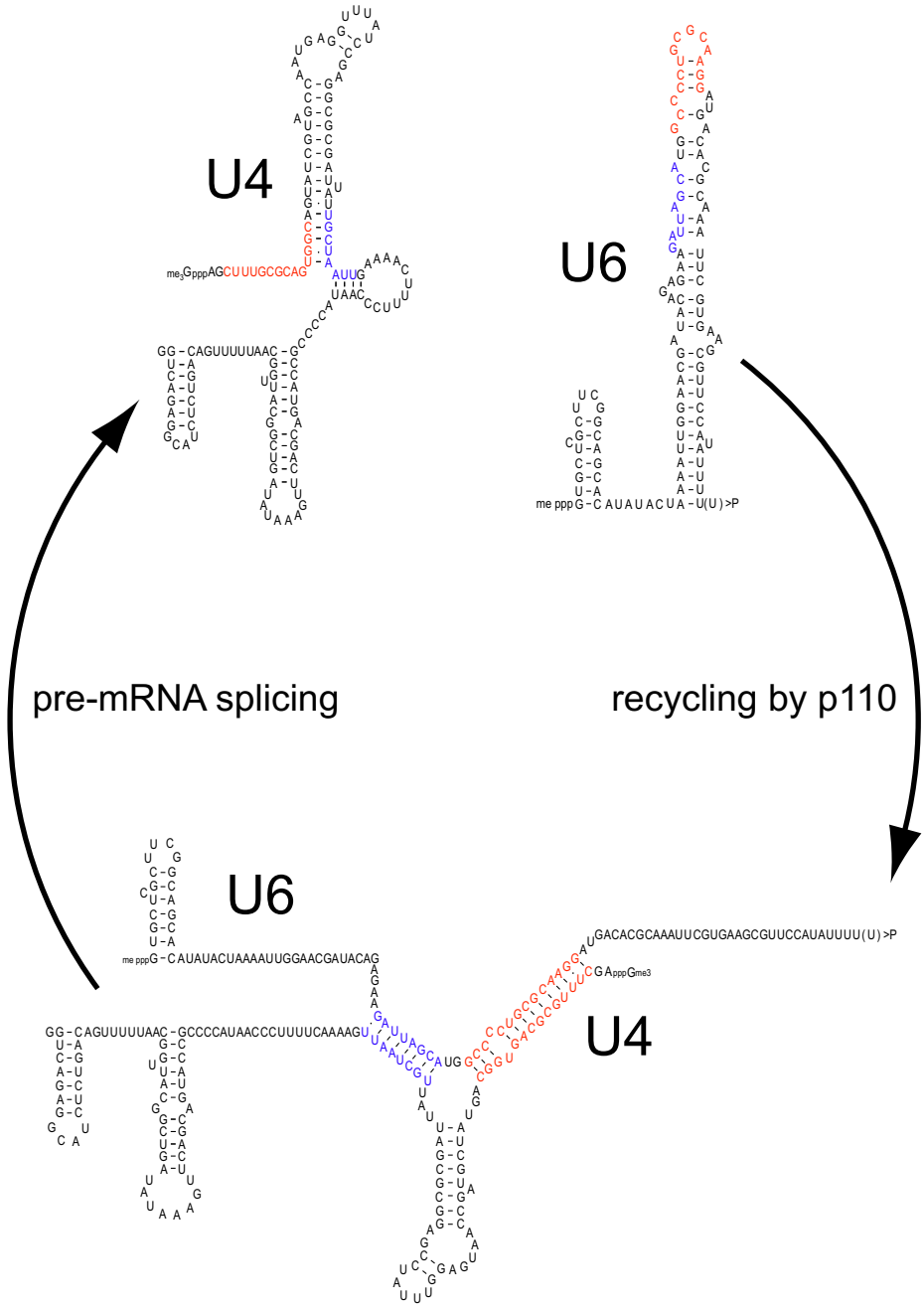


Figure 1.6. Transitions in human U4 and U6 snRNA secondary structures during the recycling of the di-snRNP.
The entire 3' stem-loop of U6 snRNA is disrupted and part of the 5' terminal stem-loop of U4 snRNA is opened; new intermolecular RNA-RNA interactions are formed. The nucleotides involved in the U4/U6 stem I are shown in blue, and those participating in stem II in red.

1.4.2.5.1. U4/U6 recycling in the yeast system

It has been proposed more than ten years ago that a yeast protein, Prp24, is involved the recycling of the U4/U6 snRNP (Shannon and Guthrie, 1991). Prp24 was originally identified by genetic suppression studies. In yeast it has been shown that it is primarily associated with U6 snRNP and to a lesser extent with U4/U6 snRNP, but is not present in the tri-snRNP (Rader and Guthrie, 2002 ; Shannon and Guthrie, 1991). Hydroxyl radical footprinting and *in vitro* complementation with mutant U6 snRNAs revealed that Prp24 binds to U6 in the region between nucleotides 30

and 65 (Ghetti *et al.*, 1995; Jandrositz and Guthrie, 1995; Ryan *et al.*, 2002). Prp24 has four RRM motifs and a highly conserved 12-amino-acid sequence at its C-terminal end (Rader and Guthrie, 2002). RRMs 2 and 3 are shown to be involved in direct interaction with U6 snRNA (Rader and Guthrie, 2002; Vidaver *et al.*, 1999), and the C-terminal motif (also called SNFL box) binds to the LSm core at the 3' end of U6 (Fromont-Racine *et al.*, 2000; Ryan *et al.*, 2002). Besides the genetic studies, Raghunathan and Guthrie provided direct biochemical evidence that Prp24 function as a U4/U6 recycling factor in yeast (Raghunathan and Guthrie, 1998). The LSm proteins enhance both the association of Prp24 with U6 (Ryan *et al.*, 2002) and U4/U6 snRNP recycling (Rader and Guthrie, 2002).

1.4.2.5.2. U4/U6 recycling in the human system

Although Prp24 had been known as a U4/U6 recycling factor in *S. cerevisiae* for over a decade, its human counterpart had not been found. It is known that the reassociation of human U4 and U6 snRNAs *in vitro* requires protein factors separate from the U4 snRNP (Wolff and Bindereif, 1993), and this process appeared to be ATP-independent. A study by Achsel *et al.*, 1999 revealed that the human LSm core itself can facilitate U4/U6 duplex formation *in vitro*, raising the question whether other factors in addition to the LSm proteins are required to recycle the di-snRNP in the human system. Recently a human protein, distantly related to Prp24, called p110^{nrb}/SART-3 (or p110 in the following), has been discovered. It carries two RRMs in its C-terminal domain and the SNFL box at the very C-terminal end. The N-terminal part of that protein contains seven half-a-TPR motifs (HATs) not found in Prp24. These motifs also exist in other RNA-processing factors and are thought to be involved in protein-protein interactions (Blatch and Lässle, 1999; Preker and Keller, 1998). A database search revealed that there are proteins from other species that share the domain organization of p110 rather than that of Prp24 (Bell *et al.*, 2002). The result of the search is summarized on **Figure 1.7.** The lack of N-terminal HAT motifs emphasizes the yeast Prp24 as an exception rather than as a conserved member of that group of structurally related proteins.

Previous reports on p110 demonstrated that it is a nuclear protein and is expressed in high levels in various tumor tissues and cell lines (Gu *et al.*, 1998; Yang *et al.*, 1999). Work by Bell *et al.* (2002), has identified p110 as a specific protein component of U6 and U4/U6 snRNP. The protein has not been found in the tri-snRNP nor in spliceosomes. p110 binds directly to U6, but not to U4 snRNA. Finally, using an *in vitro* recycling system coupled to splicing, the authors have shown that p110 is human U4/U6 recycling factor.

A study by Stanek *et al.* (2003) revealed that p110 is enriched in the Cajal bodies (CBs) and the N-terminal domain containing the HAT motifs is required for that localization. The same group proposed that U4/U6 snRNP recycling *in vivo* takes place in the CBs. The Cajal bodies are subnuclear compartments enriched in nuclear factors including snRNPs (reviewed in Ogg and Lamond, 2002). The Cajal bodies are located close to the foci of transcription and splicing. Intriguingly, the Cajal bodies have recently been implicated also as centers for recycling of the tri-snRNP (Schaffert *et al.*, 2003).

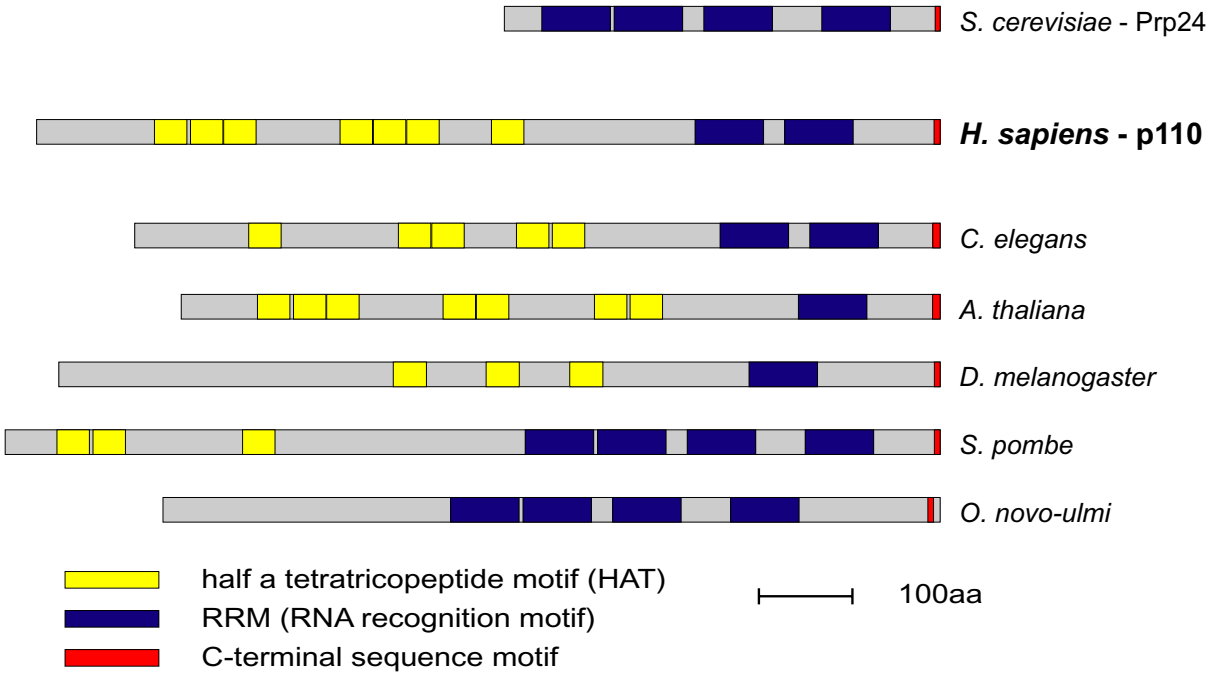


Figure 1.7. Conserved domain structures of the *S.cerevisiae* Prp24, human p110, and related proteins from other species.
The domain organisation of the *S.cerevisiae* Prp24 protein and human p110 is schematically compared with that of Prp24-related proteins from the following other species: *C.elegans*, *A.thaliana*, *D.melanogaster, S.pombe*, and *O.novo-ulmi.* The proteins are aligned relative to their C-terminal ends. The RRM (in blue) and HAT motifs (in yellow) as well as the short C-terminal region (in red) are indicated. Adapted from Bell *et al*. (2002).

1.4.2.6. Phylogeny and evolution of pre-mRNA introns

There is a close mechanistic similarity between pre-mRNA splicing and group II intron self-splicing. It is widely believed that the pre-mRNA introns have evolved from group II introns and the catalytic components of splicing function analogously to the conserved secondary structure of group II introns.

U2- and U12-type introns are likely derived from a common ancestor and have either co-diverged within the same host or have evolved independently and later were brought together as a result of merging of genetic material. Evidence in favour of that hypothesis came from the comparison of the catalytic components participating in splicing of the two types of introns. There is a direct correspondence of the RNA networks, the snRNA functional elements, and the secondary structures within the major and the minor spliceosome (reviewed in Tarn and Steitz, 1997) and similarities in the promoters of the snRNA genes (Tarn and Steitz, 1996a). Most of the protein components involved in splicing of the major and minor introns appear to be identical (Schneider *et al.*, 2002; Will *et al.*, 1999).

By comparing introns in homologous positions Burge *et al.* concluded that the U12-type introns occurred much more frequently in early evolution history and were either lost or converted to U2-type introns (Burge *et al.*, 1998). Yet the few remaining U12-dependent introns have resisted conversion and have been remarkably conserved (Spafford *et al.*, 1998), AT-AC introns have been found at non-homologous positions in several paralogous genes (Burge *et al.*, 1998). A large proportion of the genes containing minor-type introns encode proteins that are involved

in information processing, including DNA replication and repair, transcription, splicing and translation (Burge *et al.*, 1998).

These observations suggest that the U12-type introns were evolutionary retained because their presence is important for the genes harbouring them. Patel *et al.* (2002) had demonstrated that the AT-AC introns are excised more slowly than the major-type GU-AG introns (Patel *et al.*, 2002): Therefore the U12-dependent splicing may be a rate-limiting step in gene expression. Even if the U12-dependent introns are so rare, there are examples of genes that contain two and even three U12-type introns (Burge *et al.*, 1998; Levine and Durbin, 2001).

1.4.2.7. Alternative pre-mRNA splicing and splicing regulation

The recently released draft of the human genome sequence revealed a surprisingly low number of genes, about 35,000, with more recent unpublished estimates of 23,299 protein-coding genes (human ENSEMBL release 15.33.1). To generate the estimated proteome of at least 250,000 proteins, diverse posttranscriptional mechanisms are used. One major mechanism is alternative pre-mRNA splicing (reviewed in Black, 2003). It has been estimated that 59% of all human genes on chromosome 22 are alternatively spliced (Hide *et al.*, 2001). Alternative splicing can generate an astonishing diversity. For example, the *Drosophila* Dscam gene contains 95 alternatively spliced exons, organized in four clusters. The exons within each cluster are alternatively spliced in a mutually exclusive manner such that the mRNAs contain only one variable exon from each cluster. As a result, up to 38,016 different mRNAs could be generated by alternative splicing of a single pre-mRNA (Celotto and Graveley, 2001). Virtually all possible splice sites combinations are used in the alternative splicing of pre-mRNA. **Figure 1.8.** summarizes the major modes of alternative splice site usage.

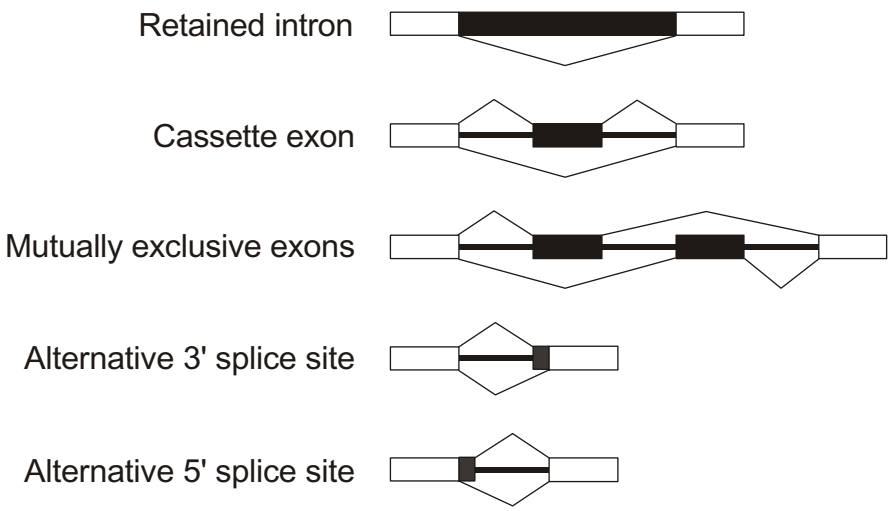


Figure 1.8. Types of alternative splicing. The constitutive exons are represented by open boxes, the alternative exons by black boxes. The introns are shown as black lines. In each case the alternative splicing events are indicated by thin lines above and below the exon-intron diagram. Adapted from Cartegni *et al* (2002).

Alternative U12-dependent splicing occurs with a lower frequency: 3.2 % of the human U12-dependent introns are alternatively spliced, compared with 14% of the U2-dependent introns (Levine and Durbin, 2001). An interesting splicing event occurs in the *D. melanogaster* prospero gene. Here a U2-dependent intron is flanked by U12-type splice sites and this “intron within an intron” pre-mRNA can be spliced by either the U2-dependent or U12-dependent pathway (Otake *et al.*, 2002).

A substantial part of the alternatively spliced exons show tissue- or cell type-specific patterns of expression (Stamm *et al.*, 1994) and/or are regulated during development or in response to external stimuli. For example, insulin administration influences the incorporation of the alternative exon 11 of the insulin receptor (Sell *et al.*, 1994) and activates exon ßII inclusion in the PKC gene (Patel *et al.*, 2001). Serum deprivation alters usage of the SRp20 exon 4 (Jumaa and Nielsen, 1997). In the brain, stress changes splicing patterns of potassium channel (Xie and Black, 2001; Xie and McCobb, 1998) and acetylcholin esterase genes (Meshorer *et al.*, 2002). The splicing of tumor necrosis factor ß is regulated by src kinases (Gondran and Dautry, 1999).
Tissue- and cell type-specific alternative splicing is still poorly understood, but it is regulated in two major ways: through tissue specific splicing factors or through regulation of the relative levels of the general splicing factors.
A few tissue-specific splicing factors have been described like the neuron-specific Nova and PTB (Jensen *et al.*, 2000; Polydorides *et al.*, 2000), and the mouse testis-specific hnRNP G protein RMB (Elliott *et al.*, 2000). Not many data are available on their mechanism of action.
The ubiquitously expressed general splicing factors also participate in regulation of the tissue-specific splicing. Although they are expressed in virtually every tissue, it has been shown that the expression levels of several SR proteins and hnRNP A are not the same in different tissues (Hanamura *et al.*, 1998). As discussed above, exon recognition by the splicing machinery is based on cooperative assembly of SR and SR-like protein complexes on pre-mRNA that is antagonized by the hnRNP proteins. Deviations in the concentration of these proteins will affect the efficiency of exon recognition. Apart from the expression levels of splicing factors, their posttranslational modifications also play a significant role in alternative splicing regulation.
Protein phosphorylation has been shown to modulate alternative splicing of a number of exons (Du *et al.*, 1998; Gondran and Dautry, 1999; Prasad *et al.*, 1999; Sanford and Bruzik, 1999). It affects the ability of several splicing factors to interact with RNA (Derry *et al.*, 2000) and with other proteins (Xiao and Manley, 1998; Yeakley *et al.*, 1999). Protein phosphorylation controls the release of the SR-proteins from the storage compartments in the nucleus (Hartmann *et al.*, 1999) and causes relocalization of hnRNP proteins to the cytoplasm (van der Houven van Oordt *et al.*, 2000). In all cases the active concentration of splicing factors will be altered and as result, alternative splicing patterns will change. Several protein kinases have been shown to specifically phosphorylate the SR repeats of splicing factors (reviewed in Stojdl and Bell, 1999). Some of them are ubiquitously expressed like the general splicing factors, the expression levels of others, however, vary significantly between different tissues.
Alternative splicing can also be regulated by through signal transduction cascades (Lynch and Weiss, 2000; van der Houven van Oordt *et al.*, 2000; Weg-Remers *et al.*, 2001), although the mechanisms, by which this is achieved, are still poorly understood.

1.4.3. snRNP biogenesis

The biogenesis of the spliceosomal snRNPs is a complex process, many aspects of which remain poorly understood (reviewed in Will and Lührmann, 2001). It involves processing of the nascent

snRNA transcripts, side chain modifications and protein assembly.

Apparently U1, U2, U4, U5 and their minor spliceosome analogues follow one general pathway of snRNP biogenesis. Their RNA components are transcribed by RNA polymerase II as pre-snRNAs. These transcripts receive an m^7G cap and are exported by a U snRNA-specific export complex (Ohno *et al.*, 2000). In the cytoplasm, a complex of seven Sm proteins is being deposited to the pre-snRNAs in an ordered, stepwise manner. The Sm proteins form three distinct heteromeric complexes: SmB/B'-SmD3, SmD1-SmD2, and SmE-SmF-SmG (Raker *et al.*, 1996). The Sm proteins are symmetrically dimethylated at their C-terminal arginine residues (Friesen *et al.*, 2001a; Friesen *et al.*, 2001b) and then associate with an oligomeric complex, containing SMN, Gemin 2, 3, 4, and other proteins. The SMN complex assists in the deposition of the Sm proteins to the conserved Sm binding site ($YAU_{4-6}GR$) of the U snRNAs (Fischer *et al.*, 1997; Liu *et al.*, 1997). After the Sm core assembly, the m^7G cap is converted to a 2,2,7-trimethylated guanosine cap (Mattaj, 1986) and the pre-snRNA undergoes 3' end maturation (Seipelt *et al.*, 1999). The assembled snRNP complexes are then imported back in the nucleus, which is dependent on the presence of both the Sm complex and the hypermethylated 5' cap (Fischer *et al.*, 1993; Huber *et al.*, 1998).

Before the association of the particle-specific proteins, the U snRNAs are internally modified at several positions, mainly by pseudouridinylation or 2'-*O*-methylation.

Some of these modifications are essential for splicing (Yu *et al.*, 1998). The minor class snRNAs are modified at fewer positions than their major class counterparts (Massenet and Branlant, 1999).

The biogenesis of the U6 and U6atac snRNPs differ in many ways from that of the other spliceosomal snRNP and is believed to take place entirely in the nucleus. U6 and U6atac are transcribed by RNA polymerase III. The La protein is a known chaperone of the polymerase III products and binds to the 3' end and stabilizes nascent U6 and probably U6atac snRNAs (Pannone and Wolin, 2000; Pannone *et al.*, 1998). U6 and U6atac snRNAs acquire a γ-monomethyl cap structure (Singh and Reddy, 1989; Tarn and Steitz, 1996). U6 snRNA is transiently located to the nucleolus, where it is internally pseudouridinylated and 2'-*O*-methylated (Ganot *et al.*, 1999; Tycowski *et al.*, 1998). Finally, a complex of seven proteins (LSm 2 to 8), similar to the Sm proteins common for all other U snRNPs, is deposited at the 3' end of U6 and U6atac snRNAs (Achsel *et al.*, 1999; Vidal *et al.*, 1999). At present there is no indication that the LSm core assembly is assisted in a way like the SMN complex does for the Sm core.

The 3' end U6 snRNA is processed and modified. Four terminal uridines are encoded by the human U6 snRNA gene. U6-specific terminal uridylyl transferase (Trippe *et al.*, 1998) and nuclease (Booth and Pugh, 1997) activities can add and remove uridines from the 3' end of U6 post-transcriptionally, generating a heterogeneous population of U6 snRNAs with a 3' poly U stretch varying in length up to 12 nucleotides. These processing events might be coupled to splicing (Tazi *et al.*, 1993). It has been found that 90% of the human U6 snRNA has a terminal cyclic 2'-3' phosphate and only the remaining 10% bear a free 3' hydroxyl group (Lund and

Dahlberg, 1992). The major form of U6 snRNA has five terminal uridines ending in a cyclic phosphate. The La protein is not associated with the cyclic phosphate-containing fraction of U6 (Terns *et al.*, 1992). Therefore blocking of the 3'terminal hydroxyl groups might be a signal to switch from the maturing La-U6 complex to the LSm-U6 complex, which is functional in splicing.

The functional significance of U6 processing is unclear and the enzymes involved in it are not identified.

2. Project aims

After the identification of human p110 as a novel U6 and U4/U6-specific protein, as well as a U4/U6 recycling factor, an important question was raised: **Is p110 also involved in recycling of the minor spliceosomal U4atac/U6atac di-snRNP?**
In the present work, this question is being addressed. More specifically, the main goals of the study were as follows:

- To test whether p110 associates with U6atac and U4atac/U6atac snRNP.
- To determine the RNA-binding specificity of p110.
- To establish an *in vitro* recycling assay and to check whether p110 is involved in the restoration of U4atac/U6atac snRNP.
- To study the similarities and the differences of major and minor snRNP recycling pathways and possibly to evaluate their functional consequences for splicing of U2- and U12-dependent introns.

An additional aim of the study was to purify and analyze the protein composition of the singular U4 and U6 snRNPs, which are the substrates for U4/U6 snRNP recycling.

3. Materials and methods

3.1. Materials

3.1.1. Chemicals and reagents

Acetic acid	Roth, Germany
Acetone	Roth, Germany
Acetonitrile	Merck, Germany
Acrylamide	Bio-Rad, USA
Acrylamide : bisacrylamide 30, 37.5:1	Roth, Germany
Acrylamide : bisacrylamide 40, 19:1	Roth, Germany
Agarose ultra pure	Roth, Germany
Ammonium persulfate	Bio-Rad, USA
Ampicillin	Roche, Germany
Aprotinin	Roche, Germany
Bacto-agar	Difco Lab., USA
Bacto-tryptone	Difco Lab., USA
Bacto-yeast extract	Difco Lab., USA
Bisacrylamide	Sigma, USA
Blocking reagent for nucleic acid hybridization and detection	Roche, Germany
Boric acid	Merck, Germany
Bovine serum albumin (BSA), RNase free	Roche, Grmany
Bromophenol Blue	Merck, Germany
Calcium chloride dihydrate	Merck, Germany
Chloroform	Merck, Germany
Coomassie brilliant blue R250	Merck, Germany
Creatine phosphate	Roche, Germany
Crystal violet	Sigma, USA
Dimethylsulphoxide (DMSO)	Merck, Germany
Dimethyl pyrocarbonate (DMPC)	Sigma, USA
Di-sodium hydrogenphosphate dihydrate ($Na_2HPO_4.2H_2O$)	Merck, Germany
Dithiotreitol (DTT)	Roche, Germany
Ethanol 100%	Roth, Germany
Ethidium bromide	Roche, Germany
Ethylenediaminetetraacetic acid (EDTA)	Sigma, USA
Formaldehyde, 37 %	Sigma, USA
Formamide	Roth, Germany
Glucose, anhydrous, cell culture tested	Sigma, USA
Glycerol, >99.5%	Roth, Germany
Glycine	Roth, Germany
Glycogen	Roche, Germany
Guanidinium isothiocyanate	Sigma, USA
7-methylguanosine	Sigma, USA
Hydrochloric acid	Roth, Germany
Isoamyl alcohol	Roth, Germany
Isopropanol	Roth, Germany
Kanamycin	Roche, Germany
Leupeptin	Roche, Germany
Magnesium chloride	Merck, Germany

β-mercaptoethanol	Bio-Rad, USA
Methanol	Roth, Germany
N-2-hydroxyethylpiperazine (HEPES)	Sigma, USA
N,N, N', N' tetramethylethylenediamine (TEMED)	Bio-Rad, USA
Nonidet P-40 (NP-40)	Sigma, USA
Phenylmethylsulfonylfluoride (PMSF)	Sigma, USA
Polyoxyethylenesorbiten monolaurate (Tween 20)	Sigma, USA
Polyvinylalcohol	Merck, Germany
Potassium chloride	Merck, Germany
Roti-Block	Roth, Germany
Roti-Blue	Roth, Germany
Roti-phenol	Roth, Germany
Roti-phenol/chloroform	Roth, Germany
Silver nitrate ($AgNO_3$)	Sigma, USA
Sodium acetate	Merck, Germany
Sodium azide	Sigma, USA
Sodium carbonate 10-hydrate ($Na_2CO_3.10\ H_20$)	Roth, Germany
Sodium chloride	Roth, Germany
Sodium citrate	Merck, Germany
Sodium dihydrogen phosphate dihydrate ($NaH_2PO_4.2H_2O$)	Merck, Germany
Sodium dodecyl sulfate (SDS)	Merck, Germany
Sodium hydroxide	Merck, Germany
Tris-hydroxymethylaminomethane (Tris)	Roth, Germany
Triton X-100	Boehringer Mannheim, Germany
Urea	Roth, Germany
Yeast tRNA	Roche, Germany

3.1.2. Nucleotides

[γ-^{32}P]ATP 3 000 Ci/mmol, 5 μCi/μl	Hartmann Analytic, Germany
Deoxyribonucleotidetriphosphate set (dNTP), 100 mM	Roche, Germany
m^7GpppG cap analogue	New England Biolabs, USA
5' [^{32}P] pCp, 800 Ci/mmol, 10 μCi/μl	Hartmann Analytic, Germany
PCR DIG-labeling mix, 2 mM dNTP/0.2 mM DIG-dTTP	Roche, Germany
Ribonucleotidetriphosphate set (NTP), 100 mM	Roche, Germany
[α-^{32}P]UTP, 800 Ci/mmol, 10 μCi/μl	Hartmann Analytic, Germany

3.1.3. Enzymes and enzyme inhibitors (supplied with reaction buffer)

Calf intestinal alkaline phosphatase (CIP), 1U/μl	Fermentas, Lithuania
Expand long template PCR system, 3.5 U/μl	Roche, Germany
Expand reverse transcriptase, 50 U/μl	Roche, Germany
GST-tagged p110, 100 ng/μl	Provided by J. Medenbach & S. Schreiner, JLU-Gießen
His-tagged p110, 100 ng/μl	Provided by S. Schreiner, JLU-Gießen
Proteinase K, 10 μg/μl	Roche, Germany
Restriction endonucleases, 20 U/μl	New England Biolabs, USA
RNase H, 1 U/μl	Roche, Germany
RNase T1, 1 U/μl	Fermentas, Lithuania
RQ1 RNase-free DNase, 1U/μl	Promega, USA
Shrimp alkaline phosphatase (SAP), 1U/μl	New England Biolabs, USA

SP6 RNA polymerase, 20 U/µl	New England Biolabs, USA
SuperScript II reverse transcriptase, 200 U/µl	Invitrogen, USA
T4 DNA ligase, 400 U/µl	New England Biolabs, USA
T4 polynucleotide kinase (PNK), 20 U/µl	New England Biolabs, USA
T4 RNA ligase, 20 U/µl	New England Biolabs, USA
T7 RNA polymerase, 20 U/µl	New England Biolabs, USA
Taq DNA polymerase, 5 U/µl	Promega, USA
RNasin RNase inhibitor, 40 U/µl	Promega, USA

3.1.4. Buffers not described in the text below

TBE 100 mM Boric acid, 100 mM Tris, 2 mM EDTA pH 8.8

D_{50} 20 mM HEPES-KOH pH 7.5, 50 mM KCl, 1.5 mM $MgCl_2$, 0.5 mM DTT

3.1.5. Molecular weight markers

DNA DIG-labeled molecular weight marker V	Roche, Germany
GeneRuler DNA ladder mix	Fermentas, Lithuania
^{32}P-labeled pBR322 *Hpa* II fragments	see 3.2.6.
peqGold protein marker IV	PeqLab, Germany
Rainbow protein molecular weight marker (RPN756)	Amersham, UK

3.1.6. Kits

QIAGEN plasmid maxi kit	QIAGEN, Germany
QIAprep spin miniprep kit	QIAGEN, Germany
QIAquick gel extraction kit	QIAGEN, Germany
QIAquick nucleotide removal kit	QIAGEN, Germany
TOPO TA cloning kit	Invitrogen, USA
Silver stain kit	Bio-Rad, USA

3.1.7. Materials for bacterial culture

Cell culture paltes — TPP, Germany

LB plate

1.5% (w/v) in LB medium, autoclaved
supplemented with 100 ng/ml ampicillin

Luria-Bertani (LB) medium

1% (w/v) Bacto-tryptone
0.5% (w/v) Bacto-yeast extract
1% (w/v) sodium chloride
adjustet to pH 7.5 by NaOH, autoclaved, and
supplemented with 25 ng/ml kanamycin for
pEGFP-F, and with 100 ng/ml ampicillin for all
other plasmids used.

SOC medium — Invitrogen, USA

3.1.8. Materials for mammalian cell culture

Dulbecco's modified Eagle's medium (DMEM)	Invitrogen, USA
Fetal calf serum	Invitrogen, USA
Opti-MEM	Invitrogen, USA
10x Phosphate-buffered saline (PBS)	Invitrogen, USA
Tissue culture dishes	TPP, Germany
1x Trypsin-EDTA, 0.5 g/l trypsin; 0.2 g/l EDTA	Invitrogen, USA

3.1.9. Plasmid vectrors

pBR322	New England Biolabs, USA
pEGFP-F	Clontech, USA
pCR2.1 TOPO	Invitrogen, USA
pHIL (modified pCMV-GST vector)	provided by Aria Baniahmad, JLU-Giessen
pQE30 UA	QIAGEN, Germany
pSP64-SCN4AENH12	Described in Wu and Krainer 1998
pSP65-MINX	Described in Zillman *et al.* 1988
pUC13-SP6-U4-*Dra* I	Described in Wersig and Bindereif 1990
pUC13-T7-U6-*Dra* I	Described in Wolff and Bindereif 1993
pUC13-T7-U6 38-83	Described in Bindereif *et al.* 1990

3.1.10. *E. coli* strains and mammalian cell lines

JM109 high-competent cells	Promega, USA
TOP 10 high-competent cells	Invitrogen, USA
HeLa (human cervix carcinoma cells)	ATCC No. CCL-2
Human embryonic kidney 293 cells (HEK293)	ATCC No. CRL-1573

3.1.11. Antisera and antibodies

Anti-Digoxigenin Fab fragment-Alkaline phosphatase	Roche, Germany
Anti-GST goat immunoglobulin	Sigma, USA
Anti-goat Immunoglobulin-Fluorescein	Sigma, USA
Anti-goat Immunoglobulin-Peroxidase	Roche, Germany
Anti-La autoimmune antiserum	Provided by CDC, USA
Anti-human Immunoglobulin-Peroxidase	Boehringer Mannheim, Germany
Anti-mouse Immunoglobulin-Peroxidase	Roche, Germany
Anti-p110 rabbit antiserum	Eurogentec, Belgium
Anti-rabbit Immunoglobulin-Fluorescein	Sigma, USA
Anti-rabbit Immunoglobulin-Peroxidase	Boehringer Mannheim, Germany
Anti-γ-tubulin monoclonal antibody	Sigma, USA
Anti-SC35 affinity purified rabbit antibody	Provided by S. Stamm, Friedrich-Alexander-Universität, Erlangen
H20-Sepharose (immobilized anti-TMG antibody)	Provided by R. Lührmann, Max-Planck-Institut, Göttingen
Y12 monoclonal antibody (anti-Sm)	Provided by I. Mattaj, EMBL, Heidelberg

3.1.12. DNA oligonucleotides

U1-F	5′-GGG GAG ATA CCA TGA TCA CG-3′
U1-R	5′-GTC GAG TTT CCC ACA TTT GG-3′
U2-F	5′-CTC GGC CTT TTG GCT AAG AT-3′
U2-R	5′-TGC AAT ACC AGG TCG ATG C-3′
U4-F	5′-AGC TTT GCG CAG TGG CAG T-3′
U4-R	5′-CCG TAG AGA CTG TCA AAA ATT GC-3′
U5-F	5′-TGG TTT CTC TTC AGA TCG CAT A-3′
U5-R	5′-CCA AGG CAA GGC TCA AAA-3′

U6-F	5′-CGC TTC GGC AGC ACA TAT AC-3′
U6-R	5′-AAA ATA TGG AAC GCT TCA CGA-3′
U4atac-F	5′-AAC CAT CCT TTT CTT GGG GTT G-3′
U4atac-R	5′-TAT TTT TCC AAA AAT TGC ACC AA-3′
U6atac-F	5′-TGT ATG AAA GGA GAG AAG GTT A-3′
U6atac-R	5′-AAA AAC GAT GGT TAG ATG C-3′
U4 83-66	5′-GGG GTA TTG GGA AAA GTT-3′
U4atac 82-67	5′-GGG TGT GTT GTT CAG GCG-3′
U6 105-86	5′-AAA TAT GGA ACG CTT CAC GA-3′
U6atac 122-103	5′-AAC GAT GGT TAG ATG CCA CG-3′
U2 49-27	5′-ATA AGA ACA GAT ACT ACA CTT GA-3′
SP6-U4atac 1-16	5′-ATT TAG GTG ACA CTA TAG AAC CAT CCT TTT CTT G-3′
NsiI-U4atac 131-113	5′-ATG CAT ATT TTT CCA AAA ATT GCA-3′
T7-U6atac 1-18	5′-TAA TAC GAC TCA CTA TAG GTG TTG TAT GAA AGG AGA-3′
T7-U6atac 38-59	5′-TAA TAC GAC TCA CTA TAG ACA AGG ATG GAA-3′
DraI-U6atac 125-105	5′-TTT AAA AAC GAT GGT TAG ATG CCA-3′
SalI-AK001239-fwd	5′-CTA CAT GTC GAC TGC CGG CCT GAC CTT ATT TGT G-3′
AK001239-NotI-rvs	5′-CTA CAT GCG GCC GCT CAT CAA CTA TCA AAC CAT TTG CT-3′
Oligo dT_{18}	5′-TTT TTT TTT TTT TTT TTT-3′
S1 anti-U4	5′-GGG GTA TTG GGA AAA GTT TTC AAT TAG CAA TA-3′
S2 anti-U4	5′-ACT GCC ACT GCG CAA AGC T-3′

These DNA oligonucleotides were ordered from Sigma-ARK, Germany or Sigma Genosys, UK

3.1.13. RNA oligonucleotides

U1 14-1-bio	5‘-UGC CAG GUA AGU AU-3‘-Biotin
U4 bio-84-69	Biotin-5′-UGG GGU AUU GGG AAA AG-3′

These RNA oligonucleotides are 2’-*O*-methylated and biotinylated at their 5’ or 3’ end. They were ordered from Eurogentec, Belgium

U6 bio-101-82 5′-UXX XXA UGG AAC GCU UCA CGA AUU U-3′

This RNA oligonucleotide is 2’-*O*-methylated. “X” denotes a biotinylated cytosine (Wolff and Bindereif 1992).

AKK001239 siRNA dimer

```
5′-GGC UGG CAG AAU AUC AAG CdTdT-3′
   ||| ||| ||| ||| ||| ||| |
3′-dTdTCCG ACC GUC UUA UAG UUC G-5′
```

These RNA oligonucleotides have deoxythymidine dinucleotide overhangs at their 3’ termini. They were synthesized, annealed, and purified by Dharmacon, USA.

3.1.14. Other materials

Coverslips	Roth, Germany
CSPD chemiluminescence substrate	Roche, Germany
ECL Western blotting detection reagent, RPN2106	Amersham, UK
Eppendorf tube, 1.5 ml, 2.0 ml	Eppendorf, Germany
Falcon tubes, 15 ml, 50 ml	Roth, Germany
Filter, 0.22 μm	Ambion, USA
Glutathione-Sepharose	Amersham, UK
HeLa nuclear extract	4C Biotech, Belgium
HeLa S100 extract	4C Biotech, Belgium
Hybond ECL nitrocellulose membrane	Amersham, UK

Neutravidin-agarose	PIERCE, USA
Nickel-nitrilotriacetic acid (Ni-NTA) agarose	QIAGEN, Germany
Oligofectamine reagent	Invitrogen, USA
Protein A-Sepharose	Amersham, UK
Roti-Nylon Plus membrane	Roth, Germany
Sephadex G50	Pharmacia LKB, Sweden

3.2. Methods

3.2.1. Preparation of total RNA from HeLa nuclear or S100 extract

HeLa nuclear or S100 extract was diluted ten times in 1x proteinase K buffer (50 mM Tris-HCl pH 8.0, 5 mM EDTA, 0.5 % SDS) and incubated for 1 hour at 50 °C in the presence of 0.2 μg/μl proteinase K. After extraction with equal volume of phenol/chloroform the RNA was precipitated by 85 mM sodium acetate pH 5.2, 70% ethanol, and 10-20 μg glycogen for 1 hour at -20 °C. The RNA was pelleted by centrifugation at 12,000 g for 20 min at 4 °C, washed by 75% ethanol, dried, and dissolved in desired volume of DMPC-treated distilled water.

3.2.2. Preparation of total RNA from HeLa cells

The RNA preparation protocol described here is a modification of the method developed by Chomczynski and Sacchi, 1987.

HeLa cells grown in 10-cm dishes were washed three times by 5 ml ice-cold 1x PBS. The cells were then lysed by 5 ml of freshly prepared solution A (50% Roti-phenol, 2 M guanidinium isothiocyanate, 12.5 mM sodium citrate pH 7.0, 100 mM sodium acetate pH 4.0, 100 mM β-mercaptoethanol). The lysate was transferred into a 15-ml falcon tube, 500 μl chloroform: isoamyl alcohol (24:1) were added and mixed by vortexing, The lysate was incubated for 20 min on ice and then centrifuged at 12,000 g for 20 min at 4 °C. The upper aqueous phase was then transferred into fresh Eppendorf tubes and precipitated by adding an equal volume of isopropanol and incubation for at least 1 hour at -20 °C. The RNA was pelleted, washed by 75% ethanol, dried and dissolved in DMPC-treated distilled water. The RNA concentration was determined by UV light absorption at 260 nm. The RNA quality was tested both by spectrophotometry (A_{260}/A_{280} absorption ratio) and by electrophoresis in 1% agarose/0.5x TBE/0.5 μg/ml ethidium bromide gel.

3.2.3. DNA Cloning

For the purposes of this study, three plasmid constructs, named pCR-U4atac-NsiI, pQE30-U6atac-DraI, and pHIL-AK001239, were made: U4atac and U6atac snRNA templates for *in vitro* transcription were cloned into pCR 2.1 TOPO and pQE30 UA, respectively. The cDNA of AK001239 gene was cloned into pHIL vector for expression *in vivo*. All three inserts were generated by RT-PCR.

3.2.3.1. RT-PCR

U4atac and U6atac were amplified from total HeLa nuclear extract RNA and AK001239 cDNA – from total HeLa RNA using the following primers: SP6-U4atac 1-16 and NsiI-U4atac 131-113 for U4atac; T7-U6atac 1-18 and DraI-U6atac 125-105 for U6atac. The second listed primer from each pair was used to generate the first strand cDNA in both cases.

The AK001239 cDNA was amplified using the primers SalI-AK001239-fwd and AK001239-NotI-rvs. The first strand cDNA was reverse-transcribed by oligo dT_{18} primer.

The reverse transcriptase reactions were carried out by SuperScript II according to manufacturer's instructions. The PCR amplifications of U4atac and U6atac templates were done by *Taq* DNA polymerase, and for AK001239 cDNA by the Expand long template PCR system.

3.2.3.2. Generation of sticky DNA ends, ligation, and transformation into *E. coli* cells

The U4atac and U6atac RT-PCR products were not cut prior to ligation. These PCR products had 3'end dA-overhangs left by the polymerase. The pCR 2.1 TOPO and pQE30 UA vectors are supplied in linear form, bearing T or U overhangs, respectively. The PCR products were ligated with the corresponding vectors according to manufacturer's instructions.

The AK001239 RT-PCR product and pHIL plasmid were 5-fold overdigested by *Sal* I and *Not* I restriction endonuleases. The cut vector was subsequently dephosphorylated by incubation for 1 hour at 37 °C in the presence of 2 units CIP per 5 μg plasmid. The PCR product and the open vector were then run on 0.8% agarose/0.5x TBE/40 μg/ml crystal violet gel and gel-purified. The DNA was recovered using QIAqick gel-extraction kit.

A ten-fold molar excess of the insert was ligated overnight at 16 °C with 100 ng of the open vector in a 20 μl reaction in the presence of 400 units T4 DNA ligase and the supplied reaction buffer.

Aliquots (1.5 μl) of the ligation reactions were transformed into TOP10 or JM109 high-competent cells according to the manufacturer's protocol. The transformed cells were inoculated on LB plates containing 100 ng/ml ampicillin. Single-cell colonies were picked up and grown in LB-Amp liquid medium. Plasmids from these clones were prepared by QIAGEN plasmid maxi kit and confirmed by sequence analysis (SeqLab, Germany). The precise 3' ends of U4atac and U6atac templates for *in vitro* transcription were generated by cutting with *Nsi* I and *Dra* I restriction endonucleases, respectively.

3.2.4. Immunoprecipitation

For immunoprecipitation experiments 50 μl packed volume of protein A Sepharose CL-4B in 50 μl N_{100} (50 mM Tris-HCl, pH 8.0, 100 mM NaCl, 0.05% NP-40, 1 mM DTT, 1 mM PMSF) was incubated with 175 μl immune- or pre-immune rabbit antiserum, rotating overnight at 4 °C, and washed five times with 1 ml N_{100}. Then the beads were incubated for 4 h together with 500 μl N_{100} and 500 μl HeLa nuclear or S100 extract, followed by washing five times with 1 ml of N_{200} buffer (containing 200 mM NaCl). Coprecipitated RNAs were eluted in 300 μl proteinase K buffer for

10 min at 80°C, phenolized, precipitated, and analyzed by Northern blot or primer extension.

3.2.5. Northern blot

RNA was separated by running at 21 W in 17 cm-long denaturing polyacrylamide-urea gels (8% acrylamide, 0.42% bisacrylamide, 50% urea, 1X TBE buffer) and transferred for 1 hour at 40 mA on Roti-Nylon Plus membrane using Trans-Blot SD Semi-dry transfer cell (Bio-Rad, USA) and 0.5x TBE as transfer buffer. The RNA was crosslinked to the membrane by irradiation with 254 nm UV light on transilluminator for 10 min. The membrane was pre-hybridized for at least one hour in Church buffer (50 mM sodium phosphate pH 7.0, 7 % SDS, 2% blocking reagent, 5x SSC buffer, 0.1% N-lauroyl sarcosin, 50% formamide), rotating at 50 °C.

PCR products of U1, U2, U4, U5, U6, U4atac, and U6atac snRNAs were amplified by RT-PCR from HeLa nuclear extract total RNA or by PCR from plasmid templates using the corresponding F / R primer pairs. Digoxigenin-labeled probes directed against the snRNAs mentioned above were obtained by multiple-cycle extension of the R primers including the corresponding PCR products as templates and PCR DIG-labeling mix. Finally the dTNPs including the DIG-dTTP were separated from the DIG-labeled probes by the QIAquick nucleotide removal kit.

The probes were hybridized to the crosslinked RNA overnight by slow rotation in Church buffer at 50 °C. The membrane was washed twice with 2x SSC buffer (30 mM sodium citrate pH 7.0, 300 mM NaCl) and twice with 0.5x SSC buffer each time for 15 min. The membrane is briefly rinsed with maleic acid buffer (100 mM maleic acid-NaOH pH 7.5, 150 mM NaCl) and blocked for 1 hour by 2% blocking reagent in maleic acid buffer.

Anti-Digoxigenin Fab fragment conjugated with alkaline phosphatase was added to the same solution in a 1:10 000 (v/v) ratio, and the membrane was incubated for another 1 hour. After washing three times for 10 min in maleic acid buffer containing 0.3 % Tween-20, the membrane was briefly rinsed in detection buffer (100 mM Tris-HCl pH 9.5, 100 mM NaCl) and incubated for 5 min in detection buffer containing 1% CSPD chemiluminescence substrate. The membrane was then quickly dried, and the hybridization signals were detected by autoradiography.

3.2.6. Preparation of ^{32}P-labeled pBR322 / *Hpa* II marker

2 µg pBR322 plasmid was cleaved by *Hpa* II restriction endonuclease. The purified plasmid fragments were dephosphorylated for 1 hour at 37 °C by 1 unit of SAP in the supplied buffer. The alkaline phosphatase was subsequently heat-inactivated for 20 min at 65 °C and the DNA fragments were 5' end-labeled for 2 h at 37 °C in the presence of 10 µCi [γ-^{32}P]ATP, 20 units of PNK and the provided buffer in a total volume of 20 µl. The labeled fragments were then diluted up to 100 µl in 1x loading buffer (0.5x TBE, 45 % formamide, 0.01% bromophenol blue, 0.01% xylene cyanol). 1-5 µl aliquots were loaded as molecular weight markers for genaturing gel electrophoresis of ^{32}P-labeled samples.

3.2.7. Primer extension analysis

For primer extension analysis the following primers were used: U6atac 122-103; U4atac 82-67; U6 105-86; U4 83-66. The primers (2 pmole) were labeled in the presence of 10 μCi [γ-^{32}P] ATP and 10 units of PNK in the supplied reaction buffer according to the manufacturer's protocol. The labeled primers were then purified on Sephadex-G50 column containing 1x TE buffer (10 mM Tris/1 mM EDTA - HCl pH 8.0).

The primer extension was carried in the presence of 10 – 25 nmole (2500 cpm/μl) 5'[^{32}P]-labeled primer, anti-p110 precipitated RNA corresponding to 5-50 μl of S100 extract, and 50 units Expand RT reverse transcriptase according to the manufacturer's instructions. The reverse transcription products were precipitated, separated by electrophoresis in 8% denaturing gel, and detected by autoradiography.

3.2.8. Native gel assays of p110-RNA binding

U4atac, U6atac, U4, and U6 snRNA were transcribed *in vitro* from the corresponding linearized templates and by the corresponding RNA polymerases in the presence or absence of [α-^{32}P] UTP. The full-length [^{32}P]-labeled snRNAs were purified from a 6% denaturing gel.

To study the affinity of RNA-protein interaction, 49 fmole of *in vitro* transcribed RNA or pre-annealed RNA duplex were incubated with the indicated amounts of N-terminally tagged GST-p110 in a 12-μl reaction containing 20 mM HEPES-KOH, pH 7.5; 140 mM KCl; 2.0 mM $MgCl_2$; 1 mM DTT; 83.3 ng/μl yeast tRNA, and 3.3 U/μl RNasin for 30 min at 30 °C. The p110-RNA complexes were analyzed by native gel electrophoresis (6% acrylamide; 0.075% bisacrylamide; 0.5X TBE buffer; run at 1.4 V/cm, 4°C). The quantitation of the protein-free and p110-bound RNA was performed by Gel-Pro Analyzer 3.0 (Media Cybernetics).

To prepare U4atac/U6atac duplex RNA, 4.7 pmole of ^{32}P-labeled U6atac snRNA was mixed with a 10-fold molar excess of unlabeled U4atac snRNA in a 24-μl reaction containing 50 mM Tris-HCl pH 7.5, 150 mM NaCl, 5 mM $MgCl_2$, and 1 mM DTT. After a 1-min incubation at 80 °C, the reaction was allowed to cool down slowly to 30°C. U4/U6 duplex RNA was prepared under the same conditions as described for the atac snRNAs, except that a 1000-fold molar excess of oligonucleotide U6-R was added.

3.2.9. Glycerol gradient fractionation of HeLa nuclear extract

10-30% glycerol gradients containing buffer G (20 mM HEPES-KOH pH 8.0, 150 mM KCl, 1.5 mM $MgCl_2$, 0.5 mM DTT, 0.5 mM PMSF, 4 μg/ml leupeptin) were prepared in 10 ml polyallomer ultracentrifuge tubes by gradient mixer. The gradients were stabilized for 1 hour at 4 °C, overlayed with 100-250 μl HeLa nuclear extract diluted with equal volume of buffer G, and centrifuged at 32 0000 RPM for 17 hours at 4 °C in Beckman L-60 ultracentrifuge with SW-40 rotor. Fractions of 500 μl are collected from top to bottom.

3.2.10. snRNA mutational analysis and p110 binding *in vitro*

Full-length versions, 5'- and 3'- truncated mutants and short derivatives of U6 and U6atac

snRNAs were *in vitro* transcribed and tested for p110 binding.

The full-length snRNAs were transcribed from their corresponding templates, U6 38-83 mutant was transcribed from pUC13 T7-U6 38-83 *Bam* HI-cut plasmid template. The template for U6atac 38-125 was obtained by PCR using a plasmid bearing the full-length U6atac sequence and the primers T7-U6atac 38-59 / U6atac-R. The U6 and U6atac snRNA derivatives containing short internal regions were transcribed *in vitro* from double-stranded oligonucleotide templates, not listed in the DNA oligonucleotide section. After *in vitro* transcription by SP6 or T7 RNA according the manufacturer's protocols, the DNA templates were degraded by incubation for 1 hour at 37 °C with 1 unit RQ1 of DNase. The transcribed RNAs were phenolized and precipitated. An aliquot from each reaction was denatured for 5 min at 85 °C in 45% formamide, 0.5x TBE buffer, and quantitated by running on 1% agarose gel along with yeast tRNA standards. 15 pmole of each snRNA were then 3' end-labeled by incubation with 10 µCi 5' [^{32}P] pCp, 10 units of T4 RNA ligase, the supplied reaction buffer and 10% DMSO for 5 hours at 16 °C. The labeled snRNAs were purified from a 12% denaturing gel.

3.2.10.1. Partial alkaline hydrolysis of RNA

5 pmole of pCp-labeled U6 38-54+78-83 and U6atac 1-30 snRNAs were subjected to alkaline hydrolysis at 100 °C in a 20 µl-reaction containing 100 mM NaOH and 1.5 mM EDTA. After 30 sec or 45 sec the hydrolysis was terminated by addition of 330 µl 25 mM Tris-HCl pH 8.0. The digested RNA was precipitated and dissolved in appropriate volume of DMPC-treated distilled water.

As control, 0.4 pmole of these snRNAs were partially digested for 10 min at 55 °C with 0.02, 0.01, and 0.003 units of T1 RNase (50-, 100-, and 300-fold dilutions of the stock 1 U/µl T1 RNase). The enzymatic digestion was carried out in the supplied reaction buffer and a final reaction volume of 6 µl. The reaction was stopped by ethanol precipitation.

3.2.10.2. p110-RNA *in vitro* binding assay

2 pmole of each RNA sample was incubated for 30 minutes at 30 °C with 4.5 pmole of His-tagged p110 in a 30-µl reaction containing buffer D_{50}, 0.5 mM ATP, and 40 U RNasin. After the addition of 240 µl buffer D_{50}, p110 complexes were precipitated by 30 µl packed Ni-NTA agarose beads, and washed four times by N_{250} buffer. The co-precipitated RNAs were analyzed by electrophoresis in 8%, 15%, or 20% denaturing gels.

3.2.11. p110 immunodepletion from HeLa nuclear extract

For p110 immunodepletion, 100 µl packed protein A Sepharose beads were incubated overnight at 4°C with 1 ml polyclonal anti-p110 antiserum. Beads were then washed 5 times with 1 ml each of N_{100} buffer, followed by the addition of 400 µl HeLa nuclear extract and incubation at 4 °C for 2 h. As control, mock-depleted extract was prepared in parallel with non-immune serum. The efficiency of p110 depletion was analyzed by Western blot.

3.2.12. Western blot

Protein samples were separated by SDS-PAGE and transferred for 1 hour at 40 mA on Hybond ECL nitrocellulose membrane using Trans-Blot SD Semi-dry transfer cell and protein transfer buffer (50 mM Tris / 380 mM glycine pH 8.8, 20 % methanol, 0.02 % SDS). The membrane was washed in PBS-T (1x PBS, 0.2% Tween 20) and blocked overnight at 4 °C in blocking solution (PBS-T, 1x Roti-Block). The first antibody, diluted in blocking solutions as follows, was incubated for 2 hours at room temperature with the membrane: anti-La autoimmune antiserum, 1:2,000; anti-p110 rabbit antiserum, 1:2,000; anti-γ-tubulin monoclonal antibody, 1:20,000, Y12 monoclonal antibody, 1:100.

The membrane was washed three times for 10 min with PBS-T and then incubated for 1 hour at room temperature with anti human, mouse, or rabbit Immunoglobulin-Peroxidase conjugate, diluted 1:20,000 in blocking solution. After four washes with PBS-T the membrane was overlayed with peroxidase chemiluminescent substrate (1x ECL: 30% H_2O_2 1,000:1) and the signals visualized by autoradiography.

3.2.13. Pre-mRNA splicing *in vitro*

SCN4AENH1 and Minx pre-mRNAs were *in vitro* transcribed by SP6 RNA polymerase from the plasmid templates pSP64-SCN4AENH12, cut by *Eco* RI, and pSP65-MINX, digested with *Bam* HI, respectively.

Pre-mRNA splicing *in vitro* was done in reaction containing 60% mock-depleted or p110-depleted nuclear extract, 0.5 mM ATP, 20 mM creatin phosphate, 3.2 mM $MgCl_2$, 1.6 U/µl RNasin, and 2.66 % PVA. For complementation assays recombinant p110 protein (50 ng per 25 µl-reaction) was added to depleted extract. Before the reaction U2-dependent splicing was blocked by incubation with 1.2 mM U2 49-27 oligonucleotide for 20 min at 30 °C. Then ^{32}P-labeled SCN4AENH1 pre-mRNA was added (0.2 ng or 10 ng per 25 µl-reaction).

Alternatively, the U2-dependent splicing was assayed without preincubation with DNA oligonucleotide by addition of ^{32}P-labeled Minx pre-mRNA (10 ng per 25 µl-reaction). *In vitro* splicing activity was assayed after incubation at 30 °C for 1.5, 2, 3 and 4 hours.

3.2.14. U4atac/U6atac snRNP recycling *in vitro*

In vitro splicing reactions with unlabeled SCN4AENH1 pre-mRNA (10 ng per 25 µl-reaction) or Minx pre-mRNA (100 ng per 25 µl-reaction) in mock-depleted or p110-depleted nuclear extract were done as described in section 3.2.13. For complementation assays recombinant p110 protein was added to depleted extract (25, 50, or 100 ng per 25 µl-reaction). After splicing for 1.5 h at 30 °C, the entire reaction (50 µl) was fractionated through by CsCl gradient centrifugation:

2.36 µl 1M $MgCl_2$ was added the splicing reaction, the volume was adjusted to 200 µl by buffer D (20 mM HEPES-KOH pH 7.5, 20% glycerol, 100 mM KCl, 1.5 mM $MgCl_2$, 0.5 mM DTT), and mixed with 300 µl of 1.55 g/ml CsCl-buffer D solution, containing 0.5 mM DTT and 0.5 mM PMSF. The reaction was then carefully overlayed over 500 µl of the same CsCl-buffer D solution in 1 ml thick-walled polycarbonate tube and centrifuged at 90 000 RPM for 20 hours

at 4 °C in Beckman TLX tabletop ultracentrifuge with TLA 100.2 rotor. Five fractions of 0.2 ml each were obtained (#5 including the pellet), RNA was purified and analyzed on a 8% denaturing polyacrylamide gel, followed by Northern hybridization with digoxygenin-labeled U4atac and U6atac probes.

3.2.15. p110 RNAi in HeLa cells

HeLa cells were grown in DMEM supplemented with 10% FBS. The evening before the transfections they were plated in 10 cm dishes. On the next day, immediately before the transfection, the cells were approximately 40% confluent and growing exponentially. The old medium was removed, the cells were washed once and covered by 4.8 ml Opti-MEM without serum. 1050 µl of Opti-MEM was mixed with 60 µl of 20 µM AKK001239 siRNA dimer. As a mock-transfection control, 60 µl of water was added instead of the RNA hybrid. In another tube, 24 µl oligofectamine was diluted with Opti-MEM to a final volume of 90 µl. After a 7-min incubation at 20 °C the content of both tubes was gently mixed and incubated for another 20-25 min. The RNA/oligofectamine complexes were then applied dropwise to the cell medium. The cells were incubated for 4 hours at 37 °C / 5 % CO_2, then supplemented with 3 ml Opti-MEM, containing 30% FBS. The cells were harvested 24, 48, and 72 hours after the transfection. The mock control was harvested after 48 hours. The cells were resuspended in 360 µl lysis buffer (20 mM HEPES-KOH pH 7.5, 20 % glycerol, 100 mM KCl, 1.5 mM $MgCl_2$, 0.5 mM DTT, 1.5% Triton X-100) or in 500 µl lysis buffer for the mock-transfected control and lysed by rotation for 15 min at 4 °C. The cell debris was removed by centrifugation at 12,000 g for 5 min at 4 °C and the lysates were quick-frozen in liquid nitrogen.

The efficiency of the p110 knockdown in was tested by Western blotting, the levels of γ-tubuline were checked as internal control. The base-pairing status of U4, U6, U4atac, and U6atac snRNAs was determined by CsCl gradient fractionation and Northern blot as described in section 3.2.14.

3.2.16. Silver staining

The RNA samples were separated by denaturing electrophoresis in 8 % acrylamide gel. The gel was fixed for at least 1 hour in 40% methanol, 10% acetic acid and incubated twice for 15 min in 10% ethanol, 5% acetic acid solution. After rinsing for 30 sec with distilled water the gel was impregnated for 30 min with 8 mM AgNO3, 0.01% formaldehyde. The gel was washed again with distilled water for 30 sec, and developed in 280 mM sodium carbonate, 0.02% formaldehyde until the RNA bands reached an appropriate intensity and then stopped for at least 15 min in 5% acetic acid. To decrease the background staining, the developer was changed twice or three times after one minute of incubation.

3.2.17. snRNP affinity selection

3.2.17.1. Selection of U1 snRNP and postspliceosomal U4 snRNP

U1 and U4 snRNPs were affinity-selected from p110-depleted HeLa nuclear extract, incubated under recycling conditions, described in section 3.2.14. The extract was subsequently fractionated through 10-30 % glycerol gradients, the fraction containing the peak of U4 snRNP was collected and the glycerol concentration was decreased to less than 5 % by dilution with washing buffer 100 (20 mM HEPES-KOH pH 7.5, 100 mM KCl, 15 mM $MgCl_2$, 0.5 mM DTT, 0.01 % NP-40).

The diluted postspliceosomal extracts were then rotated slowly on a wheel for 1 hour at room temperature in the presence of biotinylated RNA oligonucleotides that specifically anneal to U4 or U1 snRNAs. For selection of postspliceosomal U4 snRNP, 2.0 μg U4 bio-84-69 RNA oligonucleotide was used per 675 μl of nuclear extract. For selection of U1 snRNP, 5.0 μg of U1 14-1-bio RNA oligonucleotide were incubated with diluted gradient fraction, corresponding to 225 μl of nuclear extract.

The selected snRNPs were then immobilized by slow rotation for 2 hours at 4 °C on neutravidin-agarose beads, preblocked overnight in washing buffer 100, containing 1 μg/μl BSA, 0.2 μg/μl yeast tRNA, and 0.2 μg/μl glycogen. 15 μl of packed neutravidin-agarose beads were used per 1 μg of biotinylated RNA oligonucleotide. The beads were washed five times for 2 min with washing buffer 100. The associated proteins were eluted by two subsequent incubations for 15 min at room temperature with 125 μl 9M and 6M urea and precipitated overnight at -20 °C in the presence of five volumes of acetone. The RNAs were eluted by incubation for 10 min at 80 °C with 1x proteinase K buffer and precipitated by ethanol.

3.2.17.2. Selection of U6 snRNP from HeLa S100 extract

U6 snRNP was selected similarly from HeLa S100 extract, diluted with 1.5 volumes of washing buffer 100. To destroy the U4/U6 snRNP, the extracts were preincubated for 30 min at 30 °C in the presence of 1.5 μM S1 antiU4, S2 antiU4 DNA oligonucleotides and 0.01 U/μl RNase H.

For preparative-scale U6 selection, 20 ml S100 was diluted to 50 ml with washing buffer 100 and incubated for 1 hour at room temperature with 10 μg U6 bio-101-82 RNA oligonucleotide. The selected snRNP was immobilized on 80 μl preblocked neutravidine-agarose beads by slow rotation for 2 hours at 4 °C. The beads were washed four times for 2 min with washing buffer 200 (containing 200 mM KCl) and one last time with washing buffer 100. The proteins and RNAs were eluted by incubation for 10 min at 85 °C with 1x proteinase K buffer. 5% of the eluate was phenolized, the RNA precipitated, and analyzed by electrophoresis in denaturing gel and silver staining. The rest was run on SDS-PAGE, the gel was stained by Coomassie blue R-250 for 5 min and destained until the background was clear. The bands of interest were cut out from the gel, incubated twice for 2 min in 50 % acetonitrile and sent on dry ice for mass-spectrometric identification (W. S. Lane, Harvard Microchemistry Facility).

3.2.18. anti-TMG immunoaffinity chromatography

The peak fraction of postspliceosomal U4 snRNP was prepared and diluted as described in section 3.2.17.1. The fraction was run through a 150 µl H20-Sepharose column, equilibrated in washing buffer 100. The flow-through was loaded for a second time. The immobilized snRNPs were washed by 20 volumes of washing buffer 100 and eluted for 15 min in 200 µl washing buffer 100, containing 15 mM 7methylguanosine. The column was regenerated by washing with 5 volumes of 6 M urea and stored at 4 °C in 1x PBS containing 0.2 % NaN_3.

3.2.19. Transient transfection in HEK293 cells

The transfection with calcium phosphate precipitate is based on the protocol of Chen and Okayama (Chen and Okayama, 1987). HEK293 cells were replated the day before the transfection in 15 cm plates and grown in 30 ml DMEM containing 10% FBS at 37 °C / 5 % CO_2. Immediately before the transfection the cells were exponentially growing and have reached 50-60 % confluency. 75 µg of expression construct (pHIL or pHIL-AK001239) was mixed with 375 µl of 1 M $CaCl_2$ in 1.5 ml final volume. An equal volume of 2 x HBS buffer (50 mM HEPES-NaOH pH 6.95, 280 mM NaCl, 10 mM KCl, 1.5 mM Na_2HPO_4, 12 mM glucose) was added dropwise, with constant vortexing. The solution was allowed to stay at room temperature for 20 min to form a precipitate and then added dropwise to the growth medium. To express the transfected plasmid, cells were grown at 37 °C, 3 % CO_2. Typically the transfection efficiency ranged between 50 and 80 %, as determined by control transfections with pEGFP-F plasmid.

3.2.20. *In situ* immunofluorescence

Cells grown on glass coverslips and transfected with pHIL or pHIL-p85 plasmids were fixed 24 hours after transfection with freshly prepared 4 % paraformaldehyde in 1x PBS, for 30 minutes at 4 °C. The cells were washed twice with 1x PBS and were permeabilized in 0.5 % Triton X-100 / PBS for 5 min. The coverslips were washed with PBS-T and blocked for 1 hour in blocking solution (5% BSA in PBS-T). The coverslips were then incubated for 1 hour in anti-GST goat antibody diluted 1:100 in BPS-T. A control incubation with anti-SC35 rabbit antibody was carried out in paralel. After three washes with PBS-T the coverslips were incubated for 1 hour in fluorescein-labeled anti-goat or anti-rabbit secondary antibodies, diluted 1:1000 in blocking solution. The coverslips were washed five times with PBS-T and were mounted on microscope slides with antifading Gel/Mount medium. Slides were stored at 4 °C. The localization of GST, GST-p85, as well as of the endogenous SC35 was determined by fluorescence microscopy with a 488 nm filter.

3.2.21. GST-p85 pull down

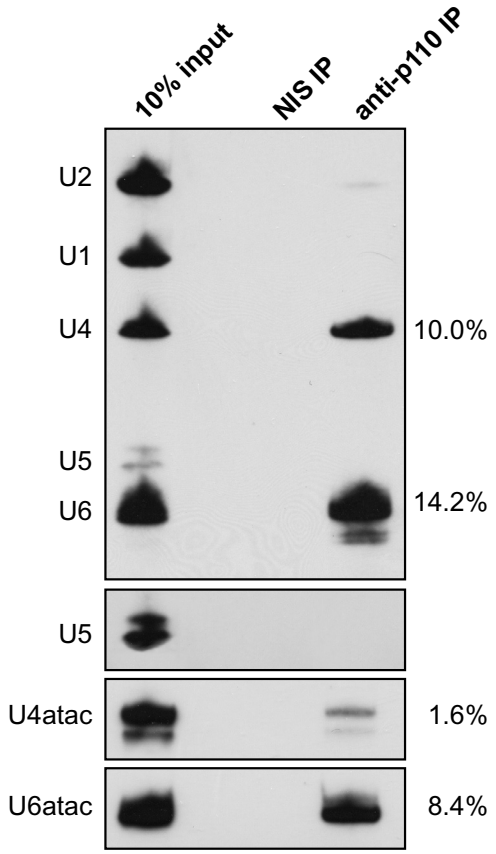
Preparative-scale transfections with pHIL or pHIL-p85 plasmids were carried out in four 15-cm dishes. The cells were trypsinized 36 hours after the transfection, harvested, and washed by PBS. The cells were allowed to swell in one packed cell volume of buffer A (10 mM HEPES-KOH pH 8.0, 10 mM KCl, 1.5 mM $MgCl_2$, 1 mM DTT) for 15 min on ice, then broken by passing 4-5

times through a thin hypodermic needle. Two cell volumes of buffer C (20 mM HEPES-KOH pH 8.0, 25 % glycerol, 420 mM NaCl, 1.5 mM $MgCl_2$, 0.2 mM EDTA, 1 mM DTT, 5.0 µg/ml aprotinin, 0.5 mM PMSF) were added and the cell homogenate was incubated with stirring for 30 minutes at 4 °C. The nuclear debris was removed by centrifugation at 12 000 g for 5 min at 4 °C and the cell extract (the supernatant) was dialyzed for 2 hours at 4 °C against buffer D. The extracts were diluted with four volumes of washing buffer 100 and slowly rotated for 2 hours at 4 °C with 60 µl glutathione-Sepharose beads. The beads were blocked prior to that as described in section 3.2.17.1. Then the beads were split into three aliquots of 20 µl and washed four times with washing buffers containing 100 mM, 250 mM, and 500 mM KCl, respectively. After one final wash with washing buffer 100, the immobilized GST and GST-p85 complexes were eluted for 10 min at 85 °C by 1x proteinase K buffer. A one-quarter-aliquot of these eluates was analyzed by SDS-PAGE followed by Roti-Blue staining. The other three quarters were treated with proteinase K, phenolized, the RNA was precipitated, separated on 8% denaturing gels, and detected by silver staining and Northern hybridization.

4. Results

4.1. p110 is a protein component of U6, U6atac, and U4/U6 snRNPs, but not of U4atac/U6atac snRNP

Using immunoprecipitation with anti-p110 antibodies from both HeLa nuclear and S100 extracts, followed by pCp-labeling or silver staining, it has been shown previously that p110 is specifically associated with U6 and U4/U6 snRNPs (Bell *et al.*, 2002). U4atac and U6atac snRNAs occur at a much lower abundance and were not detected. To test whether p110 is also a protein component of U6atac and U4atac/U6atac snRNPs, anti-p110 immunoprecipitates from HeLa nuclear extract were analyzed by Northern hybridization with U4atac and U6atac-specific probes (**Figure 4.1.**, lanes anti-p110 IP). Non-immune antiserum was used in control immunoprecipitations (lanes NIS IP). Consistent with the previous studies, anti-p110 antibodies, but not the non-immune antiserum, efficiently precipitated both U4 and U6 snRNAs. In contrast, no signals for U1, U2, nor U5 snRNA were detectable. U6atac snRNA was coimmunoprecipitated by anti-p110 antibodies, with an efficiency comparable to that of U6 snRNA; however, only background levels of U4atac could be detected.



Similar coimmunoprecipitation assays were carried out from HeLa S100 extract with the only difference that the RNAs were detected by primer extension instead of Northern. It gave essentially the same results (**Figure 4.2.**): U4, U6 and U6atac snRNAs were associated with p110 (compare lanes 1 and 3; 7 and 9; 10 and 12), but only trace amounts of U4atac snRNA coimmunoprecipitated with p110 (compare lanes 4 and 6).

Therefore, in contrast to U4 and U6 snRNAs of the major spliceosome - p110 is primarily associated with the U6atac snRNP, but is almost undetectable in the U4atac/U6atac snRNP.

Figure 4.1. Human p110 protein is associated primarily with the singular U6atac snRNP, but not the U4atac/U6atac di-snRNP in HeLa nuclear extract.

Immunoprecipitations were done from HeLa nuclear extract, using anti-p110 antibodies (lanes anti-p110 IP) or non-immune antiserum (lanes NIS IP). RNA was prepared from the immunoprecipitates and 10% of the input (lanes 10% input), followed by Northern hybridization analysis, using a mixed probe detecting U1, U2, U4, U5, and U6 snRNAs, or probes specific for U4atac and U6atac. Since U5 snRNA was more difficult to detect than the other snRNAs, an additional, longer exposure is shown for this snRNA in the middle panel. In addition, immunoprecipitation efficiencies were quantitated and listed on the right (in % of the input).

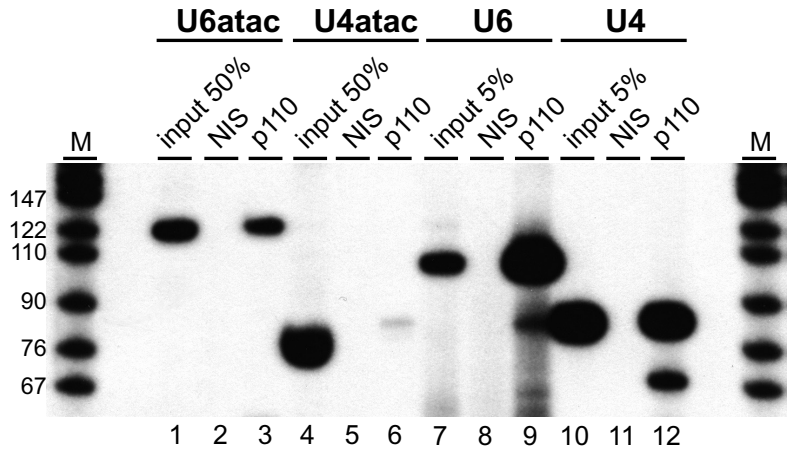


Figure 4.2. Human p110 protein is a component of the U6atac snRNP, but not of the U4atac/U6atac di-snRNP in HeLa S100 extract.
Immunoprecipitations were done from HeLa S100 extract, using anti-p110 antibodies (lanes p110) or non-immune antiserum (lanes NIS). RNA was prepared from the immunoprecipitates and the input (lanes input for U4atac and U6atac: 50%; for U4 and U6: 5%) and analyzed by primer extension, using snRNA-specific primers as indicated. M, pBR322/*Hpa* II marker fragments (67, 76, 90, 110, 122, 147 nucleotides).

4.2. snRNA sequence requirements for p110 binding

After identifying the U6atac snRNA as a second p110 target, the RNA sequence requirements for p110 binding of U6 and U6atac snRNAs were determined. Importantly, U6 and U6atac share only 45% overall sequence identity, and therefore comparing the sequence requirements for p110 RNA binding on these two functionally related snRNAs should reveal more about the RNA binding specificity of p110. It has been demonstrated previously (Bell *et al.*, 2002) that the 3' and 5'-terminal parts of U6 snRNA are not required for p110 binding. In this study, p110 binding to the central part of U6 snRNA has been examined in more detail.

Wild-type U6 and U6atac snRNAs, derivatives containing internal regions of U6, as well as the 5' terminal stem-loop of U6atac (U6atac 1-30) and the 3' terminal nucleotides 38-125 were transcribed by T7 RNA polymerase, labeled with [^{32}P]pCp, and incubated with recombinant His-tagged p110 protein (**Figure 4.3., panel A**; see **panel B** for a schematic representation of the RNAs). Bound RNAs were selected on Ni-NTA agarose.

Wild-type U6 and U6 38-83 mutant efficiently associated with p110. The U6 38-83 mutant was then split into two shorter derivatives: U6 54-79, which comprises most of the stem-loop, and U6 38-57/78-83, which contains the rest of U6 38-83. Only U6 38-57/78-83 was recognized by p110; U6 54-79 was bound at background levels.

Full-length U6atac as well as the mutant containing only the 5' stem-loop (U6atac 1-30) were also bound by p110, whereas the long 3'-terminal fragment (U6atac 38-125) was not associated with p110.

Three smaller U6 and U6atac derivatives were tested for p110 binding: U6 38-57 (lacking the 3' part of the stem-loop), U6atac 6-30, and U6atac 10-30 (**Figure 4.4., panel A**; see **panel B** for a schematic representation of the RNAs). U6 38-57+78-83 and U6 54-79 were included in this assay as positive and negative control, respectively. Each of these shorter derivatives was efficiently selected on Ni-NTA agarose, dependent on the presence of His-tagged p110.

To map more precisely the 5' ends of the sequence elements recognized by p110, U6 38-57/78-83 and U6atac 1-30 snRNAs were 3'-end labeled by [^{32}P]pCp, and partially digested by alkaline hydrolysis, generating a pool of oligonucleotides shortened at their 5' ends.

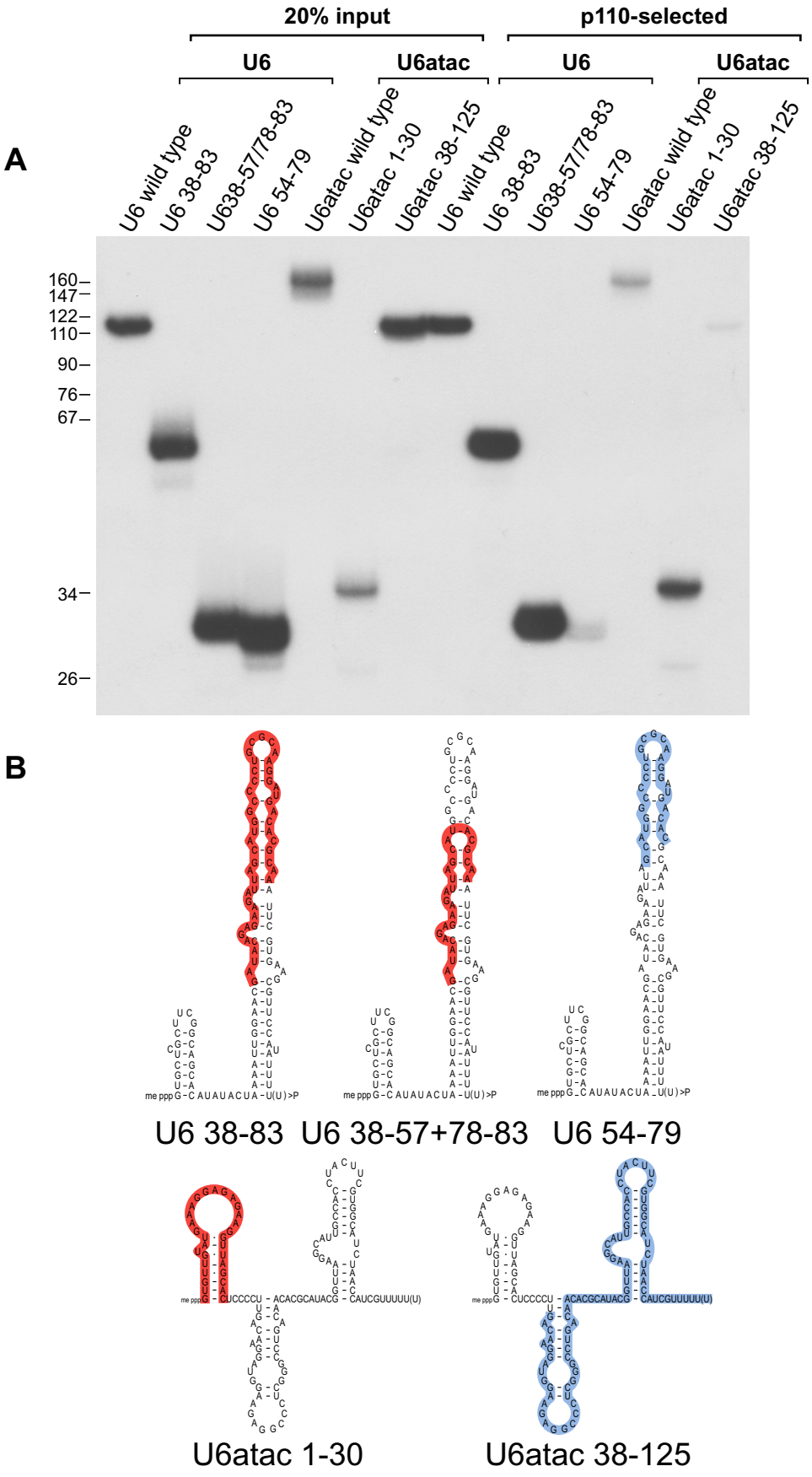


Figure 4.3. Sequence requirements for p110 binding. p110 interacts with the central part of U6 snRNA and the 5' stem-loop of U6 snRNA.

(A) p110 binding to internal U6 fragments, as well as to 5'- or 3'-truncated mutants of U6atac snRNA. ^{32}P-labeled wild-type U6 and U6atac as well as three derivatives containing internal fragments of U6

and two 5' and 3' truncations of U6atac (as indicated above the lanes; see panel B) were incubated with recombinant His-tagged p110 protein, followed by Ni-NTA precipitation and recovery of bound material. That material and 20% of the RNA input were analyzed on a 8% denaturing gel, followed by autoradiography. The positions of pBR322/*Hpa* II marker fragments are schematically indicated on the left.

(B) U6 and U6atac secondary structure models. The internal fragments of U6 as well as the 5' and 3' truncations of U6atac snRNAs are schematically represented, and their p110 binding properties shown in colour code - those highlighted in red were associated with p110, the others in blue were not

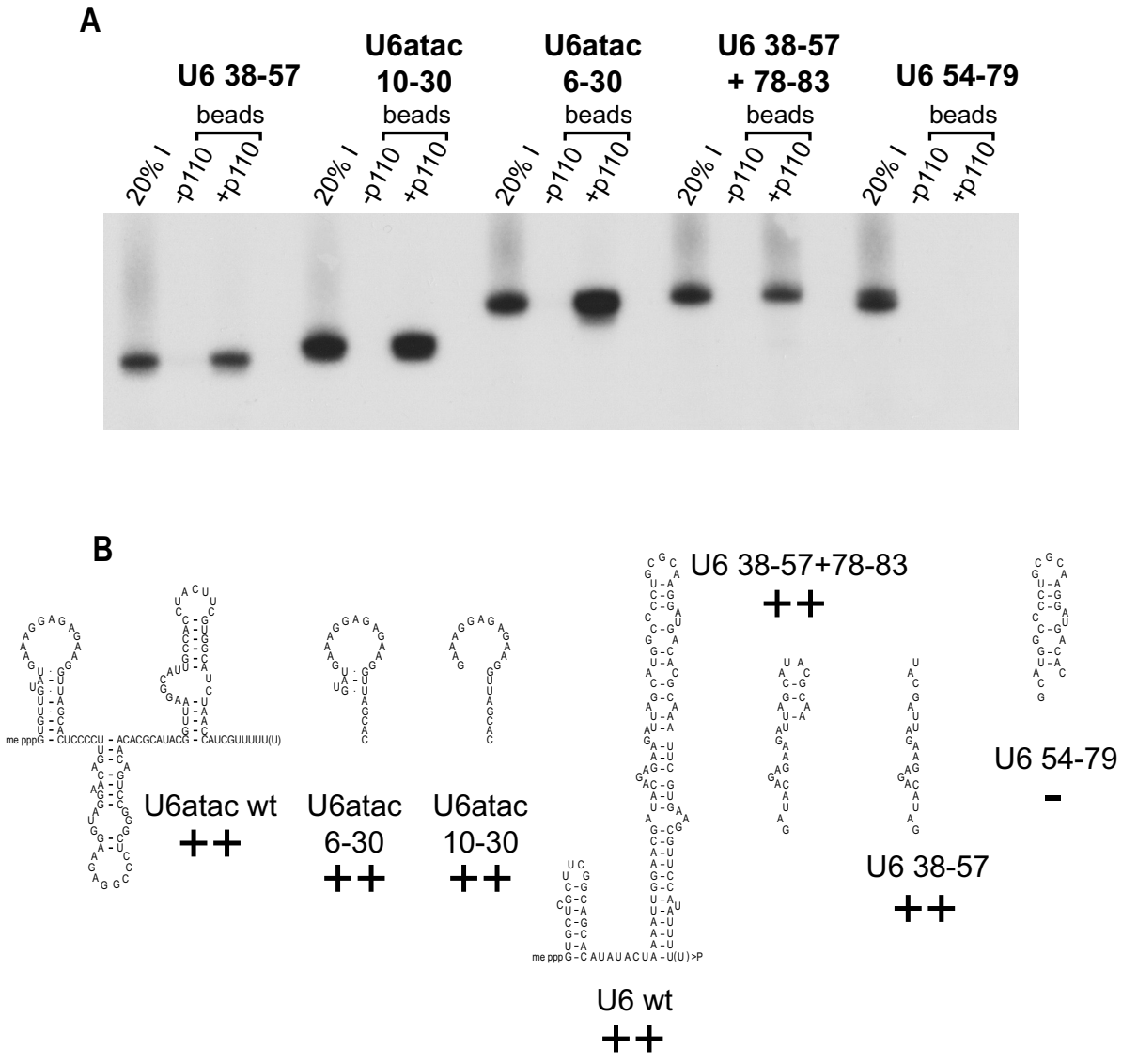


Figure 4.4. The sequence elements U6 38-57 and U6atac 10-30 are sufficient for p110 binding.

(A) p110 binding to short U6 and U6atac snRNA derivatives. ^{32}P-labeled short derivatives of U6 and U6atac snRNA (as indicated above the lanes; see panel B) were incubated in the presence (lanes +p110) or absence (lanes -p110) of p110, and selected on Ni-NTA agarose. 20% of the inputs (20% I) as well as the RNAs recovered from the agarose beads were analyzed by electrophoresis in a 15% denaturing gel and autoradiography.

(B) The short derivatives tested for p110 binding are presented as fragments of the secondary structure models of U6 and U6atac snRNAs. Their p110-binding properties are also indicated (++, > 50 % binding; -, undetectable).

These RNAs were tested for p110-binding (**Figure 4.5.**). Partial RNase T1-digests of these snRNAs were run in parallel to mark the positions of the guanines. Shortening the 5' end U6 38-57/78-83 caused rapid decrease of p110-binding (compare lanes B and I). Those RNAs with more than 2 nucleotides removed were selected only at background levels. In contrast, the 5' terminal nucleotides of U6atac 1-30 were not essential for the association with p110. The removal of the first 8 nucleotides had no significant effect; the truncation of U6atac snRNA downstream of nucleotide U_9 caused a partial decrease of p110-binding efficiency. However, when U6atac 1-30 was shortened further, its efficiency of p110 association was dramatically decreased. It could therefore be concluded that the nucleotides G_{38} in U6 and G_{10} in U6atac snRNA represent the 5' termini of the motifs required for interaction with p110.

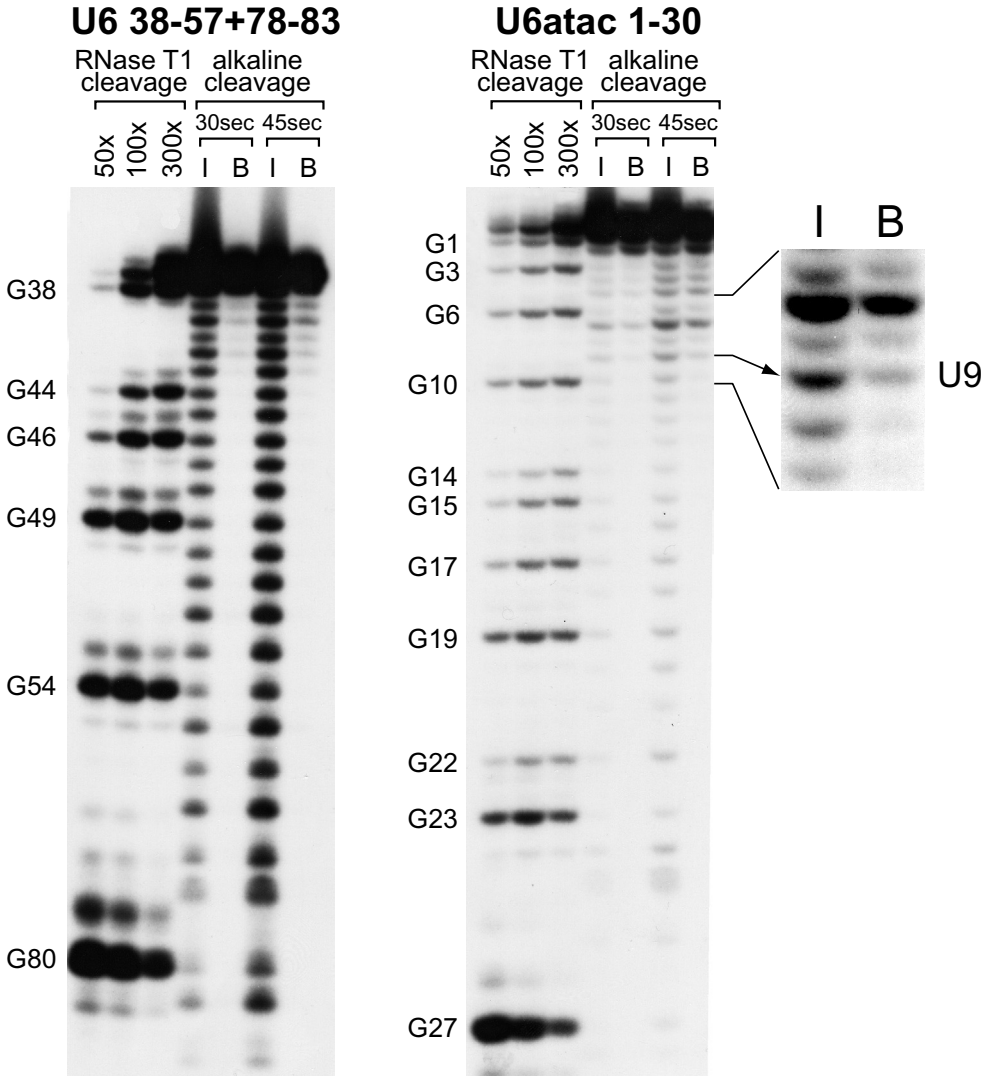


Figure 4.5. The 5' ends of U6 38-57 and U6atac 10-30 derivatives are required for p110 binding.

[^{32}P] 3' end-labeled U6 38-57+78-83 and U6atac 1-30 derivatives were partially digested by alkaline hydrolysis for 30 and 45 seconds. These RNAs were incubated with p110, selected on Ni-NTA beads, recovered, and separated (lanes B) on a 20 % denaturing gel along with 20 % inputs (lanes I). The same RNAs were also partially cleaved by increasingly diluted RNase T1 (50, 100, and 300 times) and run in parallel on the same gel. The positions of the guanosine nucleotides determined by the RNase T1 digestion are indicated on the left side. Part of the gel, including the border of U6atac fragments that are bound or not bound by p110 is shown enlarged on the right side. The last nucleotide (U_9) that can be removed without a complete loss of p110-binding is also indicated.

Comparing U6 and U6atac by sequence alignment (**Figure 4.6.**) revealed that the shortest U6 and U6atac derivatives able to bind p110, namely U6 38-57 and U6atac 10-30, represent the longest stretch of sequence homology between U6 and U6atac (76% identity). These sequence elements also have corresponding roles in U2-dependent and U12-dependent splicing.

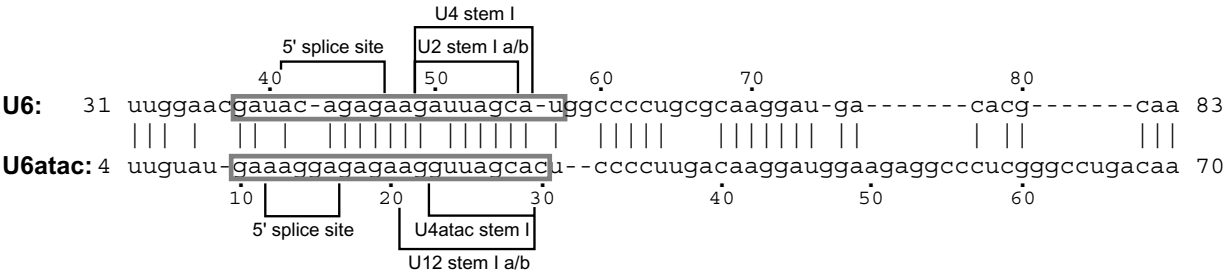


Figure 4.6. Sequence alignment of U6 and U6atac snRNAs.
The alignment of U6 nucleotides 31 to 83 and U6atac nucleotides 4 to 70 is shown. p110-binding regions, identified in both RNAs are boxed. In addition, sequences important for 5' splice site interaction (5' splice site), U4 base pairing (U4 or U4atac stem I), and U2 base pairing (U2 or U12 stem Ia/b) are indicated.

4.3. p110 binds U4/U6 and U4atac/U6atac duplex snRNAs

As shown in section 4.1., p110 is a protein component of U6 and U6atac snRNP. p110 is also present in U4/U6, but not in U4atac/U6atac snRNP. To investigate this interesting difference further, p110-RNA binding was assayed by native gel electrophoresis, titrating for each RNA binding substrate the p110 concentration (molar excess indicated above the lanes) and determining the apparent dissociation constant K_D (**Figure 4.7.**).

Initially binding of p110 to U6 and U6atac snRNA was compared (**Figure 4.7., panel A**). ^{32}P-labeled, *in vitro* transcribed U6 and U6atac RNAs were incubated in the absence of or with increasing amounts of recombinant p110 protein, followed by native gel electrophoresis. Quantitative evaluation gave for the p110 complexes of U6 and U6atac snRNAs K_D values of 138 ± 23 and 5.20 ± 0.69 nM, respectively. Therefore p110 has an approximately 25-fold higher affinity for U6atac than for U6 snRNA.

Next, the singular U6 and U6atac snRNAs were directly compared with their corresponding snRNA duplexes (U4/U6 and U4atac/U6atac) in terms of their affinity to p110 protein; in addition, p110 binding to U4 and U4atac snRNAs was assayed (**Figure 4.7, panels B and C**). ^{32}P-labeled, *in vitro* transcribed RNAs were incubated in the absence of or with increasing amounts of recombinant p110 protein, followed by native gel electrophoresis. The U4/U6 and U4atac/U6atac snRNA duplices were ^{32}P-labeled in their U6 (U6atac) component and were formed *in vitro* prior to the binding reaction. Note that in **Figure 4.7., panel B**, both singular U6 and U4/U6 duplex had a DNA oligonucleotide bound at the 3' terminal region of U6, which was necessary for efficient formation of the U4/U6 snRNA duplex (see schematic **Figure 4.8.**). Binding of oligonucleotide U6-R apparently prevents destabilization of the U4/U6 duplex by intramolecular base-pairing (see also (Mougin *et al.*, 2002). Only this way it was possible to directly compare the affinity of p110 to the singular versus the duplex snRNAs. The p110

complexes of the singular forms of U6 and U6atac snRNAs and the corresponding duplex forms could be resolved in this gel system (**Figure 4.7.**, compare **panels B and C**).

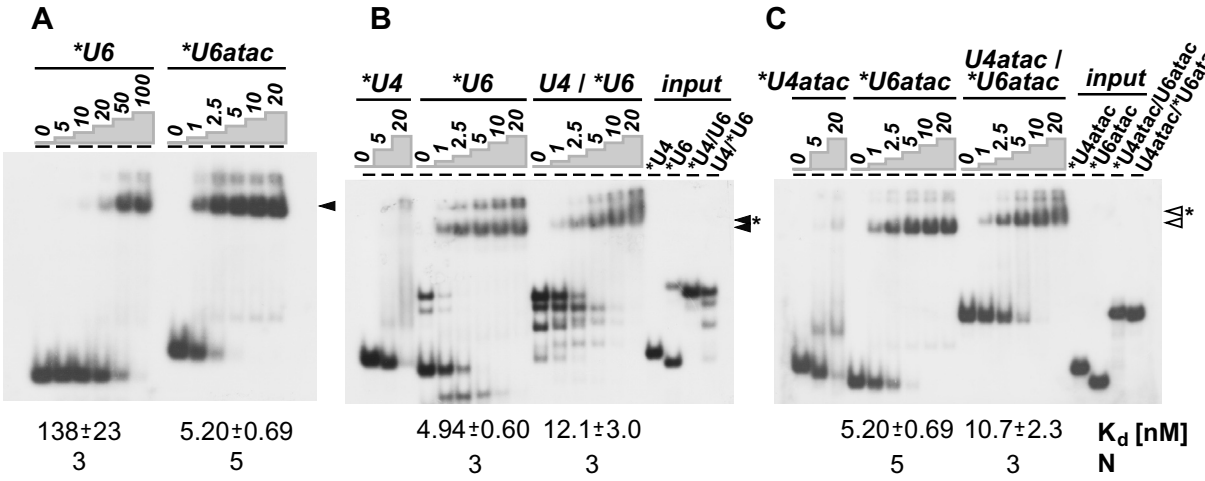


Figure 4.7. *In vitro* binding of p110 protein to singular and duplex forms of U6 and U6atac snRNAs.

The apparent dissociation constants K_D are listed below the corresponding panels. Standard deviations were derived from 3-5 independent experiments (N, number of experiments). The signals above the marked complexes are mostly due to material that had not entered the gel.

(A) Binding to singular U6 and U6atac snRNA. ^{32}P-labeled U6 and U6atac snRNAs were incubated without or with recombinant p110 protein (molar excess indicated above the lanes). RNA-protein complex formation was analyzed by native gel electrophoresis and visualized by autoradiography. The mobilities of the p110 complexes with U6 and U6atac snRNAs are marked by an arrowhead.

(B) Binding to singular U6 and U4/U6 duplex snRNA. ^{32}P-labeled U4 snRNA, ^{32}P-labeled U6 snRNA, and U4/U6 snRNA duplex (^{32}P-labeled in the U6 snRNA) were incubated without or with recombinant p110 protein (molar excess indicated above the lanes). RNA-protein complex formation was analyzed as in panel A. The input RNAs are also shown (lanes *input*: ^{32}P-labeled U4 and U6 snRNAs; U4/^{32}P-U6 and for comparison ^{32}P-U4/U6 snRNA duplex). The mobilities of the p110-U6 complex (filled arrowhead) and the p110 complex with the U4/U6 duplex snRNA (filled arrowhead with asterisk) are marked. The multiple lower bands were observed only after oligonucleotide annealing and most likely reflect different conformations of U6 and U4/U6 snRNAs.

(C) Binding to singular U6atac and U4atac/U6atac duplex snRNA. ^{32}P-labeled U4atac snRNA, ^{32}P-labeled U6atac snRNA, and U4/U6 snRNA duplex (^{32}P-labeled in the U6atac snRNA) were incubated without or with recombinant p110 protein (molar excess indicated above the lanes). RNA-protein complex formation was analyzed as in panel A. The input RNAs are also shown (lanes *input*: ^{32}P-labeled U4atac and U6atac snRNAs; U4atac/^{32}P-U6atac and for comparison ^{32}P-U4atac/U6atac snRNA duplex). The mobilities of the p110-U6atac complex (open arrowhead) and the p110 complex with the U4atac/U6atac duplex snRNA (open arrowhead with asterisk) are marked.

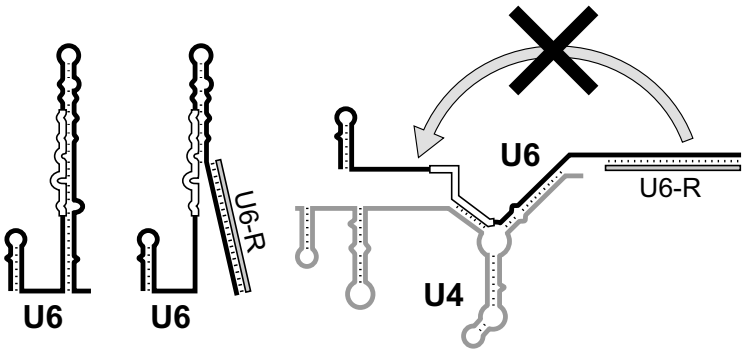


Figure 4.8. Structural alterations of singular U6 and U4/U6 duplex snRNAs by binding of oligonucleotide U6-R.

The singular U6 secondary structure model is given with the p110 binding region indicated by open lines. Binding of oligonucleotide U6-R is expected to open up the extended 3' stem of the singular form of U6. For the U4/U6 duplex, annealing of oligonucleotide U6-R should prevent the intramolecular interaction of U6 and thereby stabilize the U4/U6 duplex in the absence of protein.

In sum, the following apparent K_D values were derived:
U6 (with U6-3' oligonucleotide bound): 4.94 ± 0.60 nM;
U4/U6 (with U6-3' oligonucleotide bound): 12.1 ± 3.0 nM;
U6atac: 5.20 ± 0.69 nM;
U4atac/U6atac: 10.7 ± 2.3 nM;
In contrast to these specific U6- or U6atac-containing complexes, U4 and U4atac snRNAs formed only at the highest protein concentration complexes of diffuse mobility, which probably reflect non-specific p110 binding.
The affinity of p110 to the singular form of either U6 and U6atac snRNA is significantly higher than that for the corresponding duplex forms: For both the normal and the atac version of U6 the K_D values differ approximately two-fold. More importantly, there was a strong effect of binding the 3' terminal oligonucleotide to U6 on p110 affinity: The K_D was lowered more than 25-fold (from 138 to 4.94 nM), most likely reflecting the extensive structural alteration introduced by oligonucleotide binding (see schematic **Figure 4.8.**). However, there was no significant difference observed between the p110-binding of U4/U6 and U4atac/U6atac heteroduplex RNAs, therefore the absence of p110 in the atac di-snRNP is not due to low binding affinity for this RNA structure.

4.4. U6atac is found predominantly in the U4atac/U6atac di-snRNP form in HeLa nuclear extract

It was shown in sections 4.1. and 4.3. that in contrast to the major U4/U6 snRNP, atac di-snRNP-p110 complex was detected at low levels, however, that was not due to low-affinity binding of p110 to U4atac/U6atac heteroduplex RNA.
Another possible explanation is that U4atac/U6atac snRNP comprises a minor portion of the complex containing U6atac snRNA, most of which should then be present in singular U6atac and U4atac/U6atac.U5 snRNPs. To examine that possibility, the spliceosomal snRNPs were fractionated through 10-30 % glycerol gradient, RNA was prepared from the odd fractions and analyzed by denaturing gel electrophoresis and Northern blotting with a mixed probe detecting U1, U2, U4, U5, and U6 snRNAs, as well as with probes specific for U4atac and U6atac snRNAs (**Figure 4.9.**). The U4/U6 snRNP was less abundant than the U4/U6.U5 snRNP. These two snRNPs had peaks in gradient fractions 11 and 17, respectively. The minor spliceosomal snRNPs, however, were distributed in a much different proportion: the U4atac/U6atac di-snRNP was much prevalent than the atac tri-snRNP. Interestingly, these two snRNPs peaked in fractions 9 and 15, respectively, and did not coincide with the peaks of the major di- and tri- snRNPs.
To investigate these differences further, anti-p110 immunoprecipitations were carried out from gradient fractions 5, 7, 9, 11, 15, and 17, containing U6, U4/U6, and U4/U6.U5 snRNPs as well as their minor spliceosome analogues. The co-precipitated RNAs were purified and analyzed by denaturing gel electrophoresis and Northern blot as above (**Figure 4.10.**). Ten percent input is shown next to each immunoprecipitate. A considerable fraction of the U4/U6 snRNP (approximately 30%) was associated with p110 (see fraction 11). The protein was not found in

the singular U6 nor the tri-snRNP, which is consistent with the immunoprecipitation shown on **Figure 4.1.** This likely reflects the low abundance of the singular U6 snRNP in HeLa nuclear extract. p110-U6 snRNP complex is detected in S100 extract (**Figure 4.2.**) Importantly, most of the co-precipitated U6atac was found in fraction 9; the singular U6atac snRNP in fractions 5 and 7 did not contain p110. This is an indication that there are two minor spliceosomal snRNPs present fraction 9: U6atac-p110 snRNP and U4atac/U6atac snRNP, which was not associated with p110.

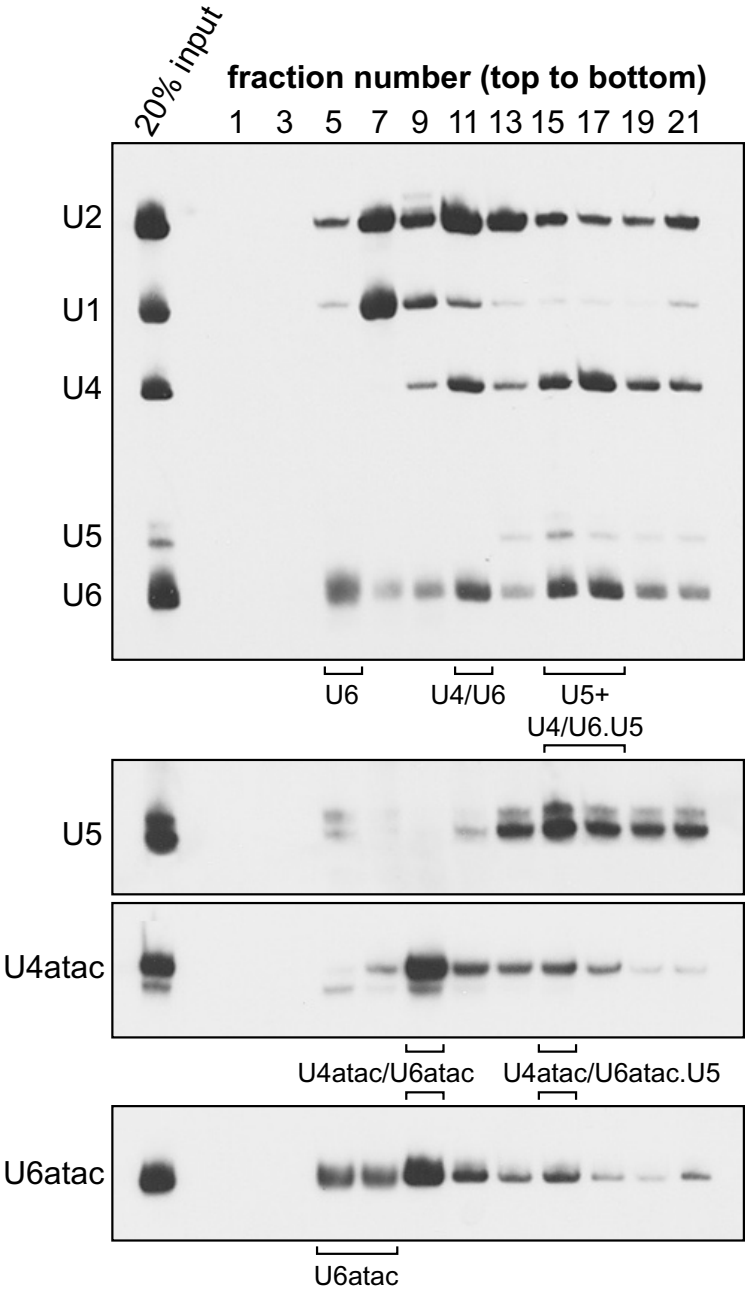


Figure 4.9. U4atac/U6atac snRNP is found in large excess over U4atac/U6atac.U5 snRNP in HeLa nuclear extract.

HeLa nuclear extract was fractionated through a 10-30 % glycerol gradient. RNA was prepared from 20 % of the input as well as from aliquots of the odd gradient fractions, and analyzed by Northern hybridization. A mixed probe detecting U1, U2, U4, U5, and U6 snRNAs was used (top panel), or probes specific for U4atac and U6atac (the two bottom panels). Due to lower sensitivity of the U5 probe, an additional, longer exposure is shown for this snRNA in the middle panel. The peaks of the singular U6 and U6atac snRNPs, as well as of the the major and minor di- and tri-snRNPs are indicated.

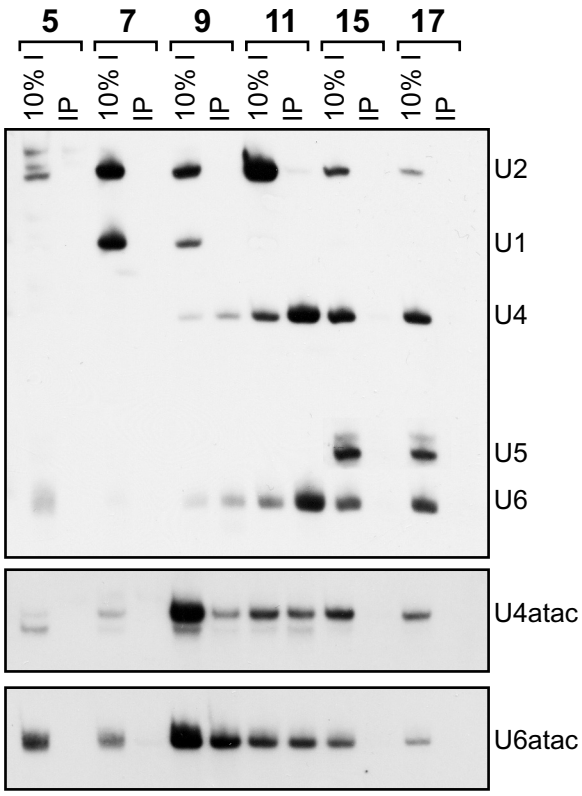


Figure 4.10. p110 is associated primarily with U4/U6 snRNP (gradient fraction 11) and with U6atac snRNP (gradient fraction 9).

Immunoprecipitations were carried out from glycerol gradient fractions 5, 7, 9, 11, 15, and 17, containing the peaks of U6, U6atac, U4/U6, U4atac/U6atac, U4/U6.U5, and U4atac/U6atac.U5 snRNPs (see **Figure 4.9.**). RNAs, prepared from 10 % of the inputs (lanes 10% I) as well as the co-precipitated RNAs (lanes IP) were analyzed by Northern blotting using a mixed probe detecting U1, U2, U4, U5 and U6 snRNAs (top panel) or probes specific for U4atac (middle panel) and U6atac snRNA (bottom panel).

Taken together, the experiments shown on **Figures 4.9. and 4.10.** reveal that there is a profound difference in the relative abundances of major and minor spliceosomal snRNPs. Both U4/U6 and U4atac/U6atac heteroduplexes are efficiently formed since there is no significant fraction of singular U4 (U4atac) snRNP detected. However, the predominant part of U4/U6 exists in a complex either with U5 snRNP (as tri-snRNP) or with p110, while most of the U4atac/U6atac complex is associated neither with U5 snRNP nor with p110.

4.5. p110 is not essential for U12-dependent splicing *in vitro* but increases splicing efficiency with excess of pre-mRNA

Despite the differences shown above, p110 associates in a similar way with major and minor spliceosomal snRNPs: p110 is a transient component of both U6 and U6atac singular snRNPs and the U4/U6 snRNPs, but is absent in the tri-snRNP complexes. Moreover, as revealed in section 4.2., it recognizes a sequence element highly conserved between U6 and U6atac, which is also involved in analogous functional interactions during U2- and U12-dependent splicing. p110 is related to the yeast recycling factor Prp24 and is required for U4/U6 snRNP recycling *in vitro* (Bell *et al.*, 2002). Therefore the question raises whether it also plays a corresponding role in the restoration of U4atac/U6atac snRNP.

To study the putative function of p110 as U4atac/U6atac recycling factor, at first the effect of p110 depletion on *in vitro* splicing of a U12-type intron has been examined. U12-type introns are much less abundant than the major U2-type introns (see 1.4.2.1.) *In vitro* splicing of U12-class introns tested so far was inefficient compared to that of the major-class introns.

In vitro splicing reactions were prepared with SCN4AENH1 pre-mRNA, which contains a single AT-AC intron, derived from the human voltage-gated skeletal muscle sodium channel α subunit. A purine-rich enhancer element, inserted at the end of the second exon, enhances splicing *in vitro* (Wu and Krainer, 1998). When incubated under splicing conditions, the SCN4AENH1 pre-

mRNA was not spliced exclusively through the U12-dependent pathway; cryptic, U2-dependent splice sites were also selected (**Figure 4.11.,** right side) To inhibit the U2-dependent splicing, the reaction was preincubated with an antisense DNA oligonucleotide that inhibits branch site recognition by U2 snRNA (Wu and Krainer, 1998). Minx pre-mRNA containing U2-type intron, derived from the adenoviral major-late transcription unit, was spliced *in vitro* efficiently; after the addition of anti-U2 oligonucleotide Minx pre-mRNA splicing was inhibited completely (**Figure 4.11.**, left side).The SCN4AENH1 pre-mRNA added after the preincubation was spliced exclusively by the U12-dependent spliceosome (**Figure 4.11.,** right side).

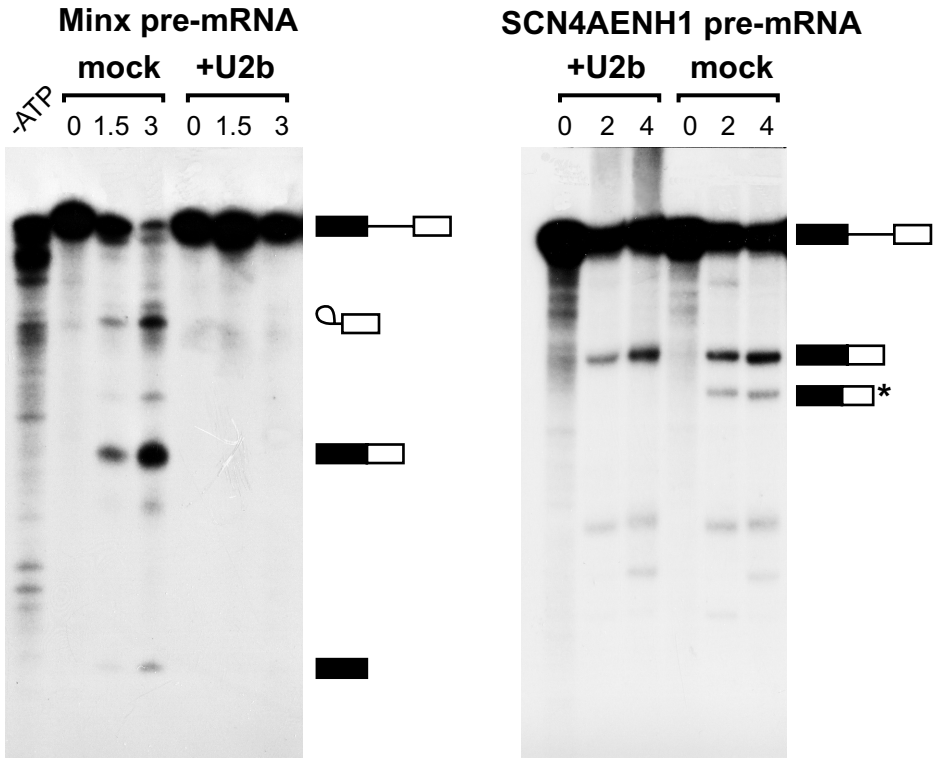


Figure 4.11. *In vitro* splicing of U2-type and U12-type introns.
HeLa nuclear extracts were incubated for 20 min in the absence (lanes mock) or in the presence of 1.2 μM U2 49-27 DNA oligonucleotide (lanes +U2b). Subsequently,
Minx pre-mRNA, containing a U2-type intron, was incubated in these extracts under splicing conditions for 0, 1.5, and 3 hours (left panel). A control incubation for three hours in the absence of ATP was also carried out (lane -ATP).
SCN4AENH1 pre-mRNA, containing a U12-dependent intron, was incubated in these extracts under splicing conditions for 0, 2, and 4 hours (right panel).
The positions of the pre-mRNA, the spliced mRNA, and the splicing intermediates (exon1 and lariat-intron - exon2) are indicated on the right. The SCN4AENH1 spliced product, indicated by an asterisk, represents an aberrant product due to U2-dependent splicing (Wu and Krainer, 1998).

Splicing of SCN4AENH1 pre-mRNA was assayed under the established U12-dependent splicing conditions at two different concentrations, 2 ng and 10 ng per 25-μl reaction. At each pre-mRNA concentration the activity of splicing *in vitro* was compared 2 and 4 hours hour after starting the reaction in mock-depleted (NEΔmock) and p110-depleted extract (NEΔp110), as well as in p110-depleted extract complemented with 50 ng of p110 protein (NEΔp110+50ng). Immunodepletion with anti-p110 antibodies routinely resulted in the selective removal of at least 95% of p110, as determined by Western blot analysis (**Figure 4.12.**). The overall *in vitro* splicing activity of

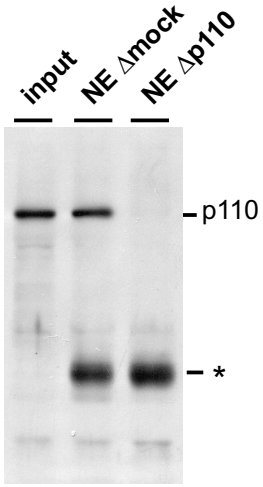


Figure 4.12. p110 is efficiently immunodepleted from HeLa nuclear extract.
Aliquots from untreated extract (input), nuclear extract after mock depletion (NE Δmock), and after p110-depletion (NE Δp110) were analyzed by Western blotting with anti-p110 antiserum. The asterisk indicates the heavy chain of the immunoglobulin, which leaked in the extract during depletion.

SCN4AENH1 pre-mRNA was relatively low, already at 2 ng and even more so at 10 ng pre-mRNA per 25-µl reaction (**Figure 4.13.**). Splicing activity after 4 hours at the low pre-mRNA concentration (2 ng per 25-µl reaction) was very similar, whether p110-depleted extract was used or after complementation with 50 ng of p110 protein (**Figure 4.13.**, compare lanes 2 and 4). Unfortunately, at 2 ng per 25-µl reacton, splicing of SCN4AENH1 pre-mRNA was less efficient in mock-depleted than in p110-depleted extract concentration (compare lanes 2 and 6), which is probably an experimental artifact. At the high pre-mRNA concentration, however, a significant, two-fold stimulation was observed by addition of p110 (**Figure 4.13.**, compare lanes 8 and 10). Consistently, the splicing activity of the mock-depleted extract at the high pre-mRNA concentration was comparable to that of the depleted extract, complemented with p110 (compare lanes 9-10 and 11-12). The splicing-stimulatory effect observed at the high pre-mRNA concentration therefore suggests that a p110-dependent recycling step becomes limiting for splicing of the ATAC-type intron.

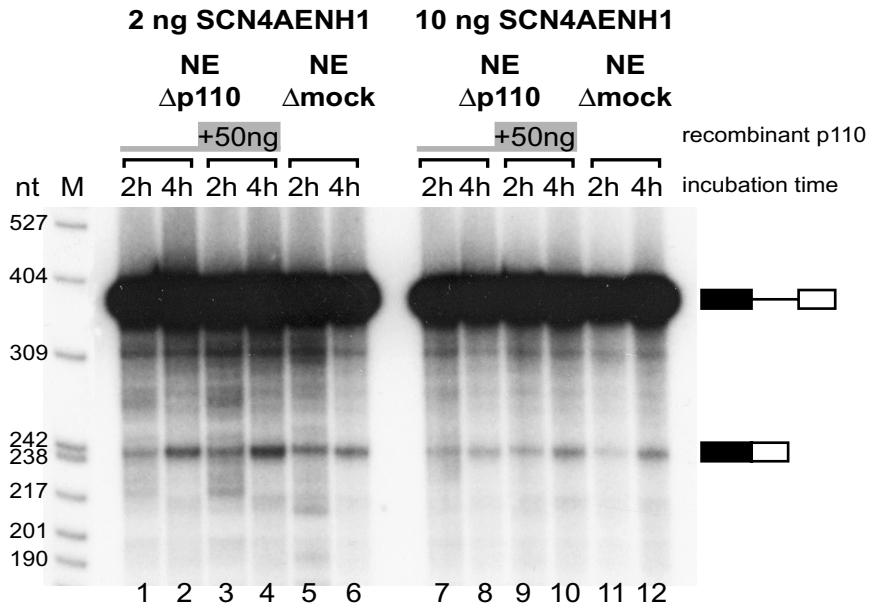


Figure 4.13. Addition of recombinant p110 to depleted nuclear extract stimulates U12-dependent *in vitro* splicing with excess of pre-mRNA.
In vitro splicing of ^{32}P-labeled SCN4AENH1 pre-mRNA (left part: 2 ng per 25-µl reaction; right part: 10 ng per 25-µl reaction). SCN4AENH1 pre-mRNA was incubated in p110-depleted nuclear extract (NEΔp110; without and with 50 ng complementing recombinant p110 protein per 25-µl reaction) as well in mock-depleted (NEΔmock) nuclear extract. All extracts were incubated for 20 min with U2 49-27 oligonucleotide before the addition of the pre-mRNA. Splicing was analyzed after incubation for 2 and 4 hours. Pre-mRNA and spliced mRNA are indicated on the right. M, pBR322/*Hpa* II marker fragments.

4.6. p110 is required for U4atac/U6atac snRNP recycling

To test that possibility directly, a U4atac/U6atac-recycling assay coupled to splicing *in vitro* has been employed, similar to the assay developed to investigate the U4/U6-recycling role of p110 (Bell *et al.*, 2002). *In vitro* splicing reactions were carried out using a high concentration (10 ng per 25-µl reaction) of unlabeled SCN4AENH1 pre-mRNA. After a 1.5-hour incubation under splicing conditions, the relative levels of U4atac/U6atac as well as of free U4atac and U6atac snRNPs were quantitatively measured, using CsCl density gradient centrifugation (see **Figure 4.14.** for a schematic outline).

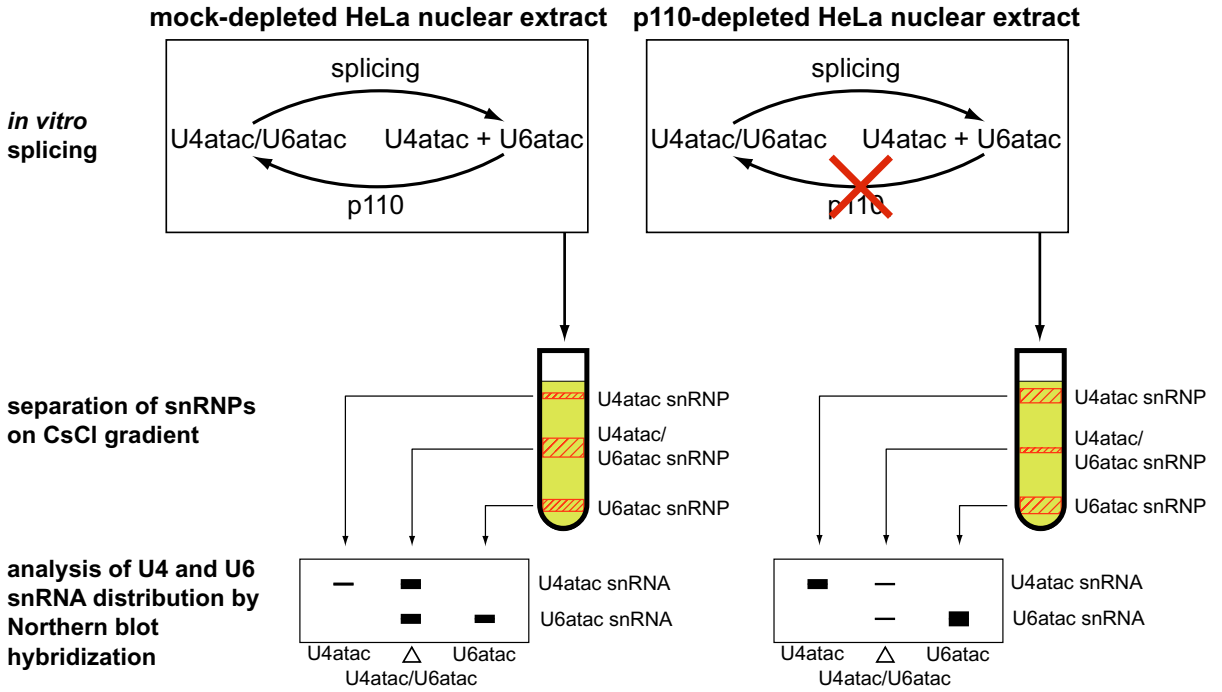


Figure 4.14. Schematic representation of splicing-coupled *in vitro* assay for U4atac/U6atac recycling.
In vitro splicing with excess of unlabeled SCN4AENH1 pre-mRNA was carried out in mock-depleted and p110-depleted HeLa nuclear extract. The content of the splicing reactions was fractionated through CsCl density gradient centrifugation, and the distribution of U4atac and U6atac was analyzed by denaturing gel electrophoresis and Northern blotting. If p110 is required for recycling of U4atac/U6atac snRNP, there will be more singular U4atac and U6atac snRNPs and less U4/U6 dimer in the p110-depleted extract, compared to the mock-depleted extract.

Under these highly stringent conditions, core complexes of snRNPs are characteristically stable and fractionate according to their density, which is determined by their RNA/protein ratio (Lelay-Taha *et al.*, 1986). The larger U4atac/U6atac.U5 complexes and spliceosomes dissociate. The entire splicing reaction was fractionated by CsCl density gradient centrifugation, RNA was prepared from the resulting five fractions (#1-5, from top to bottom, the last including the pellet), separated by denaturing gel-electrophoresis, and analyzed by Northern blot hybridization with probes specific for U4atac and U6atac snRNAs. Clearly, the U4atac and U6atac snRNAs cofractionated with a peak in #3, representing the U4atac/U6atac core snRNP. This gradient fraction contained approximately 70 % of the total U4atac and half of the total U6atac material, as determined by quantitation of the Northern blot signals (see the first gradient profiles shown on **Figures 4.15. and 4.17.**; the peak fractions are indicated by arrowheads on the top). U6atac

shows an additional peak at the bottom of the gradient, representing approximately one third of the total U6atac and most likely consisting of material derived from free U6atac snRNP. Only a small part (approximately 5 %) of U4atac fractionated in #1, where no U6atac snRNA could be detected. Therefore this material at the top of the gradient represents the singular U4atac snRNP. In sum, the relative levels of base-paired U4atac/U6atac, free U4atac, and free U6atac snRNPs can be quantitatively monitored by this assay. Gradient fractions #1, #3, and #5 are indicative for the abundance of free U4atac, base-paired U4atac/U6atac, and free U6atac snRNPs, respectively.

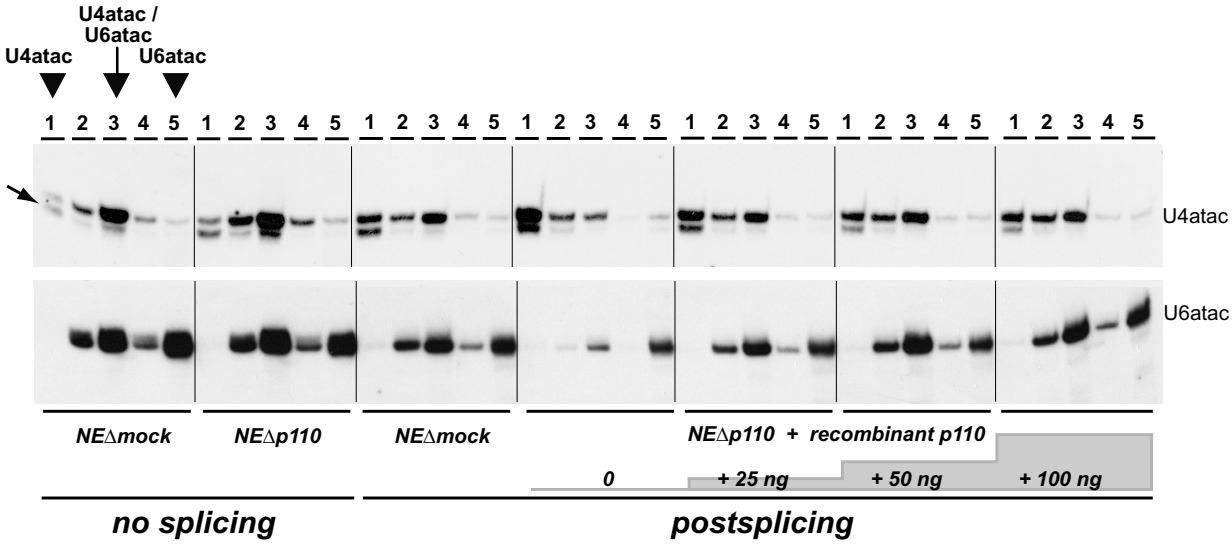


Figure 4.15. ***In vitro*** **recycling of the U4atac/U6atac snRNP.**
Mock-depleted (NEΔmock) and p110-depleted nuclear extract (NEΔp110) without prior incubation were fractionated by CsCl gradient centrifugation (panel no splicing). Similarly, mock-depleted (NEΔmock) and p110-depleted nuclear extract (NEΔp110) were analyzed after a 1.5-hour splicing reaction with SCN4AENH1 pre-mRNA under recycling conditions. In addition, incubations were done in p110-depleted extract with complementation by 25, 50, or 100 ng recombinant p110 protein per 25-µl reaction as indicated (panel postsplicing). For each reaction the distribution of U4atac and U6atac throughout the entire CsCl gradient has been determined: RNA was prepared from fractions #1 to 5 (top to bottom; as indicated at the top) and analyzed by denaturing gel electrophoresis and Northern blotting, using separate probes for U4atac (top part) and U6atac (bottom part). The positions of the U4atac/U6atac snRNP (fraction #3) and the singular forms of U4atac (#1) and U6atac (#5) within the CsCl gradient are given for the first panel on the left. The electrophoretic mobilities of U4atac and U6atac snRNAs are marked on the right. The minor form of U4atac snRNA (31) is indicated by an arrow on the left side.

Interestingly, there is a minor U4atac form, which had been noted previously by (Tarn and Steitz, 1996) and which is slightly shortened at the 5' end (**Figure 4.15.**, see minor band marked by arrow). Most of this 5'-shortened U4atac occurs in fraction #1, where no U6atac snRNA could be detected, and very little is present in the di-snRNP form. A likely interpretation of this observation is that the 5'-shortened U4atac snRNA is ineffective in di-snRNP formation as a result of the truncated U4atac/U6atac stem II.

The distribution of U4atac and U6atac in mock-depleted (NEΔmock) and p110-immunodepleted nuclear extract (NEΔp110) without splicing incubation (**Figure 4.15.**, NE Δmock and Δp110, panels *no splicing*) was very similar. In both extracts the U4atac/U6atac di-snRNP represents the major form of these two snRNAs, and there is very little singular U4atac snRNP.

However, after a 1.5-hour incubation, the snRNP distribution in the p110-depleted extract was dramatically altered, compared to that in the mock-depleted extract (**Figure 4.15.**, NE Δmock and Δp110, panel *post-splicing*). Most of U4atac (60 %) was converted to the singular U4atac snRNP, and most of U6atac (72 %) was shifted to the bottom of the gradient, representing the singular U6atac form; only approximately 20 % of U4atac and U6atac remained in the di-snRNP form. The most likely explanation for these changes is that the U4atac/U6atac snRNPs are consumed during assembly of the minor spliceosome and in SCN4AENH1 pre-mRNA splicing. As a result of SCN4AENH1 pre-mRNA splicing and the recycling block, post-spliceosomal free U4atac and U6atac snRNPs accumulate. Moreover, when the depleted extract was complemented with 25, 50, and 100 ng of recombinant p110 protein, these shifts in both U4atac and U6atac distributions were reversed in a concentration-dependent manner (**Figure 4.15.**, panel post-splicing). Already the addition of 50 ng p110 protein fully restored the distribution observed in mock-depleted extract.

This demonstrates that p110 functions as a U4atac/U6atac snRNP recycling factor and is responsible for the regeneration of base-paired U4atac/U6atac snRNPs from post-spliceosomal free U4atac and U6atac snRNPs.

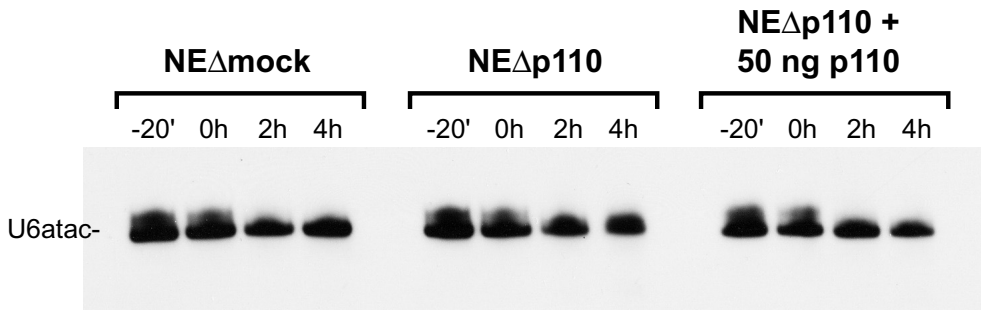


Figure 4.16. U6atac snRNA is not destabilized in p110-depleted extract.
Mock-depleted HeLa nuclear extract (panel NEΔmock), p110-depleted extract (panel NEΔp110), and p110-depleted extract supplemented with 50 ng recombinant p110 (panel NEΔp110 + 50 ng p110) were incubated under recycling conditions. Total RNA was prepared from aliquots taken before the incubation with U2 49-27 oligonucleotide (lanes -20'), before the addition of SCN4AENH1 pre-mRNA (lanes 0h), and after incubation for 2 and 4 hours (lanes 2h and 4h). The RNA was analyzed by Northern blotting using a U6atac-specific probe.

The steady-state level of U6atac snRNA after splicing appears to be considerably lower in p110-depleted than in mock-depleted extract (**Figure 4.15.**, NE Δmock and Δp110, panel post-splicing). This suggests that the stability of U6atac snRNA might be affected by p110 depletion. To confirm this hypothesis, the levels of U6atac snRNA were analyzed without CsCl gradient fractionation: Mock-depleted, p110-depleted extract, and depleted extract, complemented with 50 ng recombinant p110, were incubated under recycling conditions as described above. The U6 snRNA levels before the preblocking of U2 (lanes -20'), before the addition of the SCN4AENH1 pre-mRNA (lanes 0h) and after 2 and 4 hours of splicing (lanes 2h and 4h) were directly compared by Northern blotting (**Figure 4.16.**). The total amount of U6atac snRNA was not reduced significantly in any of these three extracts. The levels of other spliceosomal snRNAs were also not affected considerably (data not shown). Therefore the decrease of U6atac snRNA levels observed in **Figure 4.15.** is not due to degradation, U6atac appears to be lost in form of

precipitates during the splicing reaction in p110-depleted extract.

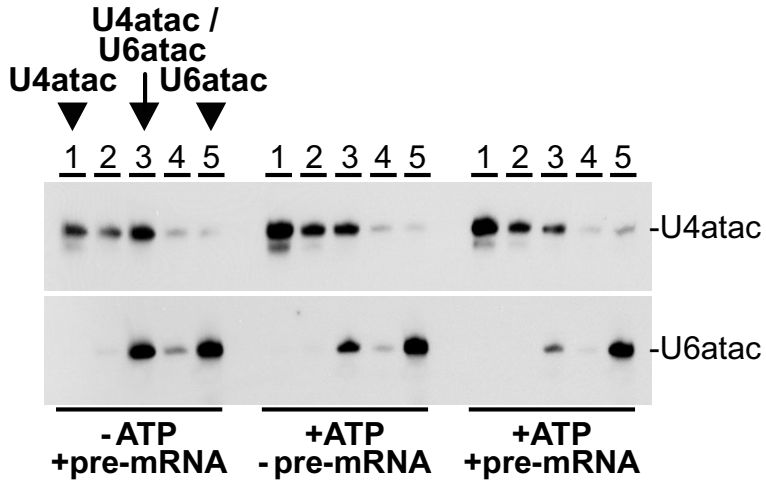


Figure 4.17. The accumulation of singular U4atac and U6atac snRNPs is enhanced by splicing *in vitro*.
p110-depleted nuclear extract (NEΔp110) was incubated for 1.5 hours under splicing conditions (panel +ATP, +pre-mRNA), as well as in the absence of ATP (panel -ATP, +pre-mRNA) or in the absence of SCN4AENH1 pre-mRNA (panel +ATP, -pre-mRNA). The distribution of the U4atac/U6atac di-snRNP and of the singular U4atac and U6atac snRNPs was determined by CsCl gradient centrifugation: RNA was prepared from fractions #1 to 5 (top to bottom; as indicated at the top) and analyzed by denaturing gel electrophoresis and Northern blotting, probing separately for U4atac (top) and U6atac (bottom). The positions of the U4atac/U6atac snRNP (fraction #3) and the singular forms of U4atac (#1) and U6atac (#5) within the CsCl gradient are given for the first panel on the left.

Interestingly, a partial disruption of U4atac/U6atac snRNP and accumulation of free U4atac and U6atac snRNPs was found in p110-depleted extract, incubated under splicing conditions, even without the addition of SCN4AENH1 pre-mRNA (**Figure 4.17**, panel +ATP/ -pre-mRNA). However, when incubated in the presence of 10 ng SCN4AENH1 pre-mRNA per 25 µl-reaction, the dissociation of the U4atac/U6atac snRNP was approximately two-fold more efficient (compare fractions #3, +ATP/- pre-mRNA and +ATP/+pre-mRNA). The disruption of the U4atac/U6atac snRNP appears to be ATP-dependent; no significant accumulation of U4atac and U6atac snRNPs, as well as no significant decrease of U4atac/U6atac snRNP was detected after incubation in the absence of ATP (panel -ATP/+pre-mRNA, compare with panels -ATP). There are two possible explanations for this phenomenon: First, endogenous pre-mRNAs containing U12-type introns might exist in the nuclear extract; second, there might be a splicing-independent disruption of U4atac/U6atac dimer in the presence of ATP. In principle, the first possibility cannot ruled out, but is much less likely than the second one, since even if these pre-mRNAs are present, they would be in extremely low levels, insufficient to promote significant levels of spliceosome assembly. The second possibility is more probable in analogy with the splicing-independent disruption of the major U4/U6 snRNP, which has been observed *in vitro*: Mock-depleted and p110-depleted HeLa nuclear extracts were incubated under U4atac/U6atac-recycling conditions as described above. Before the addition of SCN4AENH1 pre-mRNA, the U2 snRNA was blocked by incubation with U2 antisense oligonucleotide as usual. That resulted in complete inhibition of U2-dependent splicing as shown in section 4.5. These two *in vitro* splicing reactions were fractionated through CsCl gradients; ten fractions of 100 each µl plus pellet (instead of five of 200 µl each) were collected from top to bottom. Then the distribution of U4 and U6 snRNAs was analyzed by denaturing gel electrophoresis and Northern hybridization (**Figure 4.18.**). Despite the inhibition of splicing, U4/U6 snRNP was disrupted, and singular U4

and U6 snRNP - accumulated in the p110-depleted extract; in the mock-depleted extract the U4/U6 dimer was much more abundant and the free U4 snRNP – almost nonexistent.

Therefore the spliceosome cycle is not the sole pathway to disrupt U4/U6 and U4atac/U6atac heteroduplexes (discussed in section 5.4.).

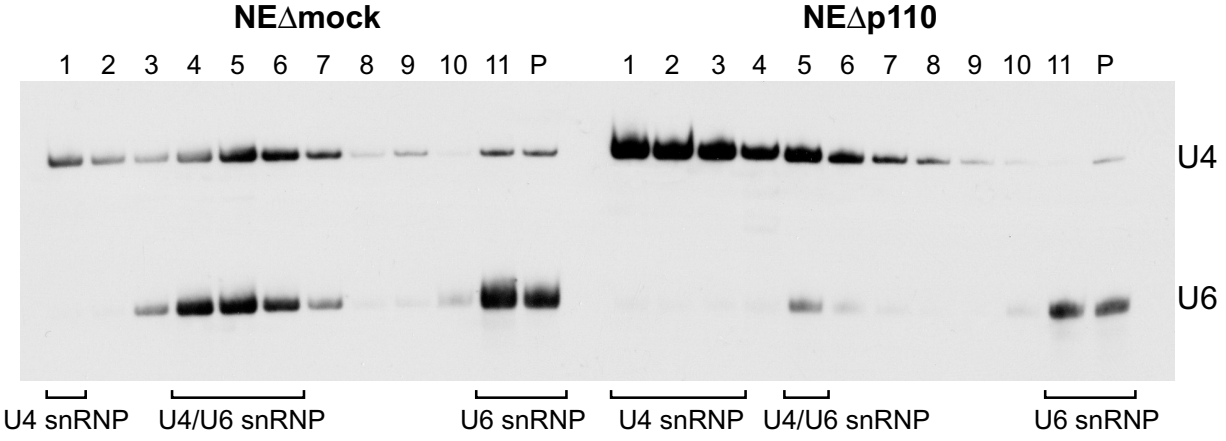


Figure 4.18. The U4/U6 snRNP is disrupted in p110-depleted extract independent of splicing.

Mock-depleted (panel NEΔmock) and p110-depleted extracts (panel NEΔp110) were incubated for 20 min with U2 49-27 DNA oligonucleotide and then for 1.5 hours under recycling conditions with SCN4AENH1 pre-mRNA. The extracts were fractionated through CsCl gradients. Ten fractions plus the pellet were collected from top to bottom and the RNA purified from these fractions was analyzed denaturing gel electrophoresis and Northern blotting with probes detecting U4 and U6 snRNA. The peaks of singular U4 and U6 snRNPs as well as of U4/U6 di-snRNP are indicated below.

4.7. Knockdown of p110 in HeLa cells by RNAi

To check the *in vivo* effect of p110 depletion on major and minor di-snRNP recycling, HeLa cells were transfected with an siRNA oligonucleotide dimer (AK001239 siRNA), targeted against the CDS of p110 mRNA. The cells were harvested 24, 48, and 72 hours after transfection and lysed in the presence of 1% Triton X-100. p110 levels in the lysates of transfected and mock-transfected cells, where no siRNA oligonucleotide dimer had been added, were analyzed by Western blotting (**Figure 4.19.**). p110 was reduced to 10 % of the control levels already 24

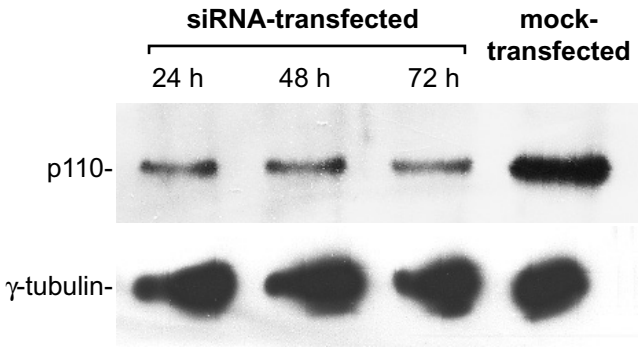


Figure 4.19. RNAi-knockdown of p110 in HeLa cells.

HeLa cells were transfected with an siRNA oligonucleotide dimer (AK001239 siRNA) targeted against p110 mRNA. The transfected cells were harvested 24, 48, and 72 hours after transfection, mock-transfected cells were harvested after 48 hours. Cell lysates were prepared and analysed by Western blotting using anti-p110 antiserum (top panel) or anti-γ-tubulin antibodies (bottom panel).

hours after transfection, while the amount of γ-tubulin remained unchanged. The transfection efficiency was approximately 90 %: Indirect *in situ* immunofluorescence analysis showed that in 60% of the cells p110 was below the detectable levels, and that in approximately 30% of the cells p110 levels were reduced in comparison to the mock-transfected cells (in collaboration with

D. Stanek and K. Neugebauer, Planck Institut für molekulare Zellbiologie und Genetik, Dresden; data not shown).

The siRNA oligonucleotide transfection resulted in inhibition of growth and cell death (data not shown). The control cells, as well as cells transfected with other siRNA oligonucleotides, also designed to target p110 mRNA, but inefficient in p110-knockdown, had no growth-inhibition phenotype (data not shown).

The base-pairing status of U4, U6, U4atac, and U6atac snRNA was studied 48 hours post-transfection. Lysates prepared from p110-knockdown cells as well as from mock-transfected cells were fractionated on CsCl gradients; ten fractions were collected from top to bottom, the pellet was recovered as well, and the snRNA distribution was determined by denaturing gel electrophoresis and Northern blot (**Figure 4.20.**). No significant difference in the relative amounts of U4/U6 snRNP between transfected and mock-transfected cells was observed (top panel). Similarly, U4atac and U6atac base-pairing status was not significantly altered by p110 knockdown (middle and lower panels). In the total cell lysates U6 (U6atac) exists in a larger excess over U4 (U4atac) snRNA, compared to the nuclear extract. Considerable fraction of U6 (U6atac) is always in the form of singular snRNP, therefore the distribution of U4 and U4atac snRNAs is more indicative for the efficiency of di-snRNP recycling. However, neither U4 nor U4atac snRNPs accumulated after p110 knockdown.

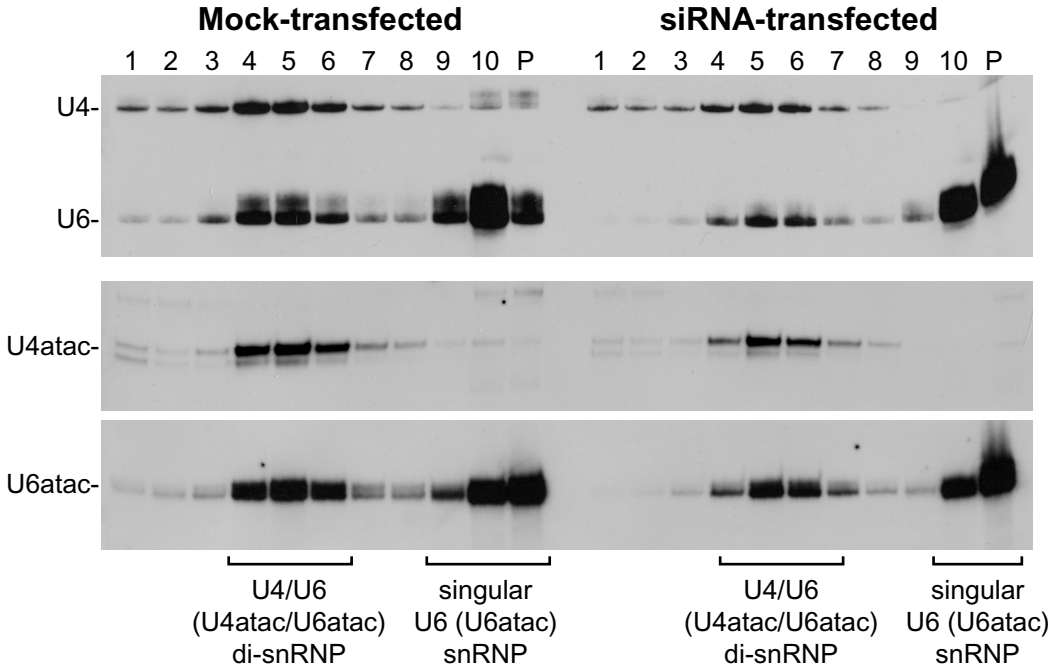


Figure 4.20. p110 knockdown does not disrupt U4/U6 and U4atac/U6atac snRNPs *in vivo*. Lysates from mock-transfected (left part) or siRNA-transfected HeLa cells were fractionated through a CsCl gradient. Ten fractions plus the pellet were collected from top to bottom; the RNA from these fractions was purified and analyzed by denaturing gel electrophoresis and Northern blotting with a mixed probe detecting U4 and U6 snRNAs (top panel), or with U4atac and U6atac-specific probes (middle and bottom panels).The peaks of the di-snRNPs and the singular U6 (U6atac) snRNP are indicated below.

This result suggests that *in vivo* there might be alternative mechanisms of di-snRNP recycling, not mediated by p110. The growth-inhibition phenotype observed after p110-knockdown suggests further that p110 may have other functions in addition to recycling.

4.8. Purification of postspliceosomal U4 snRNP

As it has been shown in section 4.6., postspliceosomal U4 and U4atac snRNPs accumulate in p110 depleted HeLa nuclear extract after incubation under splicing conditions. The protein components of these complexes have not been characterized until now.

To collect additional information about the postspliceosomal snRNPs, p110-depleted extract was incubated with an excess of Minx pre-mRNA under splicing conditions as described by (Bell *et al.*, 2002). This extract was fractionated by glycerol gradient sedimentation under native conditions. The distribution of the postspliceosomal snRNPs was analyzed by denaturing gel electrophoresis and silver staining and compared with that of snRNPs in extract that had not been incubated (**Figure 4.21.**).

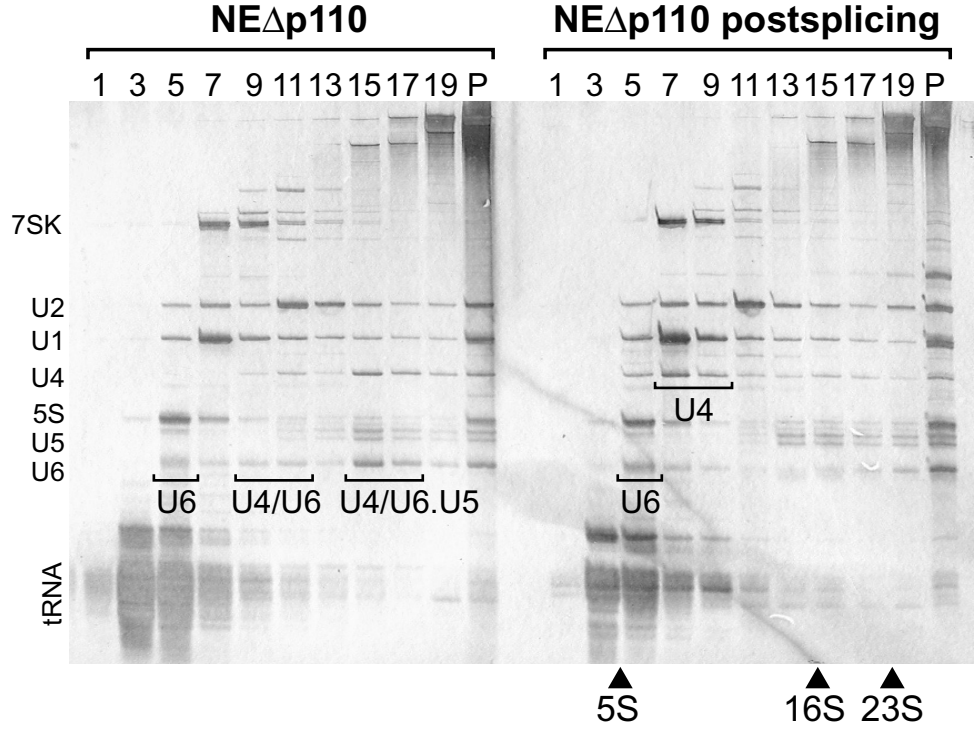


Figure 4.21. Prespliceosomal and postspliceosomal complexes of U4 and U6 snRNAs.
snRNP complexes from p110-depleted extract, incubated for one hour with Minx pre-mRNA under recycling conditions (panel NEΔp110 postsplicing), were fractionated through a 10-30 % glycerol gradient. Control fractionation of snRNPs from p110-depleted extract, not incubated under splicing conditions, was carried out in parallel (panel NEΔp110). RNA was prepared from the odd fractions plus pellet, and analyzed by denaturing gel electrophoresis and silver staining. The most abundant RNAs in the nuclear extract and the peaks of the snRNPs containing U4 or U6 snRNA are indicated. Marker 5S, 16S, and 23S rRNAs were sedimented in parallel on a control gradient; their peaks are marked by arrowheads at the bottom of the gel.

No significant differences of the sedimentation profile of U1 and U2 snRNPs were observed (compare panels NE Δp110 and NE Δp110 post-splicing). Before incubation, U6, U4/U6, and U4/U6.U5 snRNPs are typically resolved with peaks in gradient fractions 5, 11, and 15, respectively (panel NE Δp110). After incubation under splicing conditions, no peaks of di- and tri-snRNP complexes were detected anymore; U4 snRNA peaked in fraction 7, U6 snRNA in fraction 5. Therefore, U4 and U6 snRNPs have remained predominantly in their singular states. U5 peak shifted from fraction 15 before the incubation to fraction 13 after the incubation,

probably reflecting the disruption of the tri-snRNP and the accumulation of the 20S U5 snRNP. The minor spliceosome complexes were not detected by silver staining since they are much less abundant.

The postspliceosomal U4 snRNP was initially purified from glycerol gradient fractions 6 to 9 by H20 immunoaffinity chromatography (Schneider *et al.*, 2002). The H20 monoclonal antibody specifically recognizes the hypermethylated cap of the spliceosomal snRNAs. To improve the purification efficiency, the material derived from the gradient was passed twice through the column. The immobilized snRNPs were eluted by a large excess of m^7G cap analogue, and the RNA prepared from the eluate was analyzed by denaturing gel electrophoresis followed by silver staining (**Figure 4.22.**). U1 and U2 snRNPs, which also contain 2,2,7-methyl guanosine cap co-purified with U4 snRNP as expected. U6 snRNA, however, was not present in the eluted material, indicating that the postspliceosomal U4 snRNP is not associated with U6.

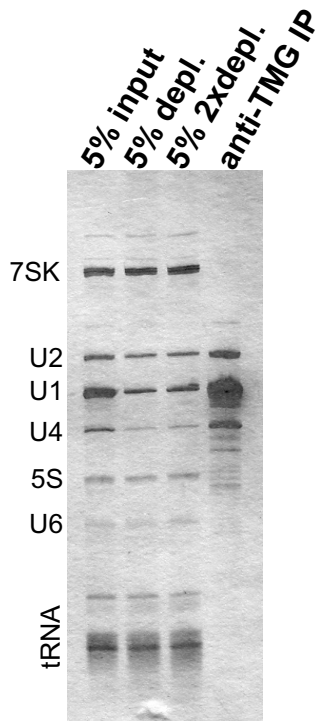
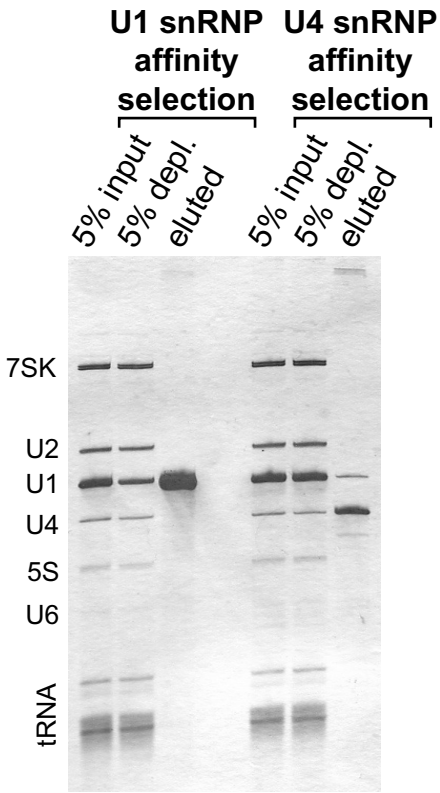


Figure 4.22. Immunoaffinity purification of snRNPs containing hypermethylated snRNAs.
p110-depleted nuclear extract, incubated under recycling conditions with Minx pre-mRNA was fractionated through a 10-30% glycerol gradient. Gradient fractions 6 to 9, containing the singular U4 snRNP, were loaded on an H20-Sepharose column. The flow-through was loaded on the column for the second time. The immobilized snRNPs were eluted by m^7G cap analogue. RNA was prepared from 5% of the input (lane 5% input), the material passed through the column once (lane 5% depl.) or twice (lane 5% 2xdepl.), as well as from the eluate (lane anti-TMG IP), separated in denaturing gel and silver-stained. The electrophoretic mobilities of the major nuclear RNAs are indicated on the left side.

Figure 4.23. Affinity selection of U1 and U4 snRNPs.
p110-depleted nuclear extract, incubated under recycling conditions with Minx pre-mRNA was fractionated through a 10-30% glycerol gradient. Gradient fractions 6 to 9, containing the singular U4 snRNP, were pooled and used as an input material. U1 and U4 snRNPs were selected by biotinylated antisense 2'-*O*-methylated RNA oligonucleotides, immobilized on neutravidine-agarose beads, washed and eluted under denaturing conditions. RNA was prepared from 5% of the pooled fractions before (lanes 5% input) and after (lanes 5% depl.) the snRNPs were immobilized, and from the eluates (lanes eluted). The RNA was analyzed by denaturing gel electrophoresis and silver staining.



U4 snRNP was also affinity-selected from the same gradient fractions using a biotinylated antisense 2'-*O*-methylated RNA oligonucleotide. It anneals to U4 snRNA nucleotides 69 to 84, a region not involved in extensive secondary structure (see **Figure 1.7.** for secondary structure models of U4 and U4/U6 snRNA) nor in binding of any of the known U4/U6-

specific proteins (Nottrott *et al.*, 2002). The selected snRNPs were immobilized on neutravidin-agarose beads, the RNA was eluted, separated on denaturing gel and silver-stained (**Figure 4.23**, right part). The purified complex(es) contained U4 snRNA as well as substoichiometric amount of U1 snRNA. This might be explained by U1 and U4 being released from the B complex in the form of a U1/U4 di-snRNP. For that reason a control affinity selection of U1 snRNP was carried out in parallel, using a biotinylated U1-antisense 2'-*O*-methylated RNA oligonucleotide. The latter selection contained only U1, but not U4 snRNA (**Figure 4.23**, left part). Therefore the U1 snRNP was selected independently of the postspliceosomal U4 snRNP; a postspliceosomal U1/U4 snRNP is not likely to exist.

The attempts to identify the protein components of the purified U4 snRNP so far were unsuccessful. The affinity selection contained a high background of proteins that bind directly to the U4-antisense oligonucleotide (data not shown). The material purified on the H20 column was not sufficient for detecting any of the known U4/U6-specific proteins by Western blotting (in collaboration with S. Nottrott and R. Lührmann, Max-Planck-Institut für biophysikalische Chemie, Göttingen, data not shown).

Further experiments are required to characterize the postspliceosomal U4 snRNP.

4.9. Purification and analysis of U6 snRNP from HeLa S100 extract

As shown in section 4.7., U6 occurs in a large excess over U4 snRNA in total cell lysates. In nuclear extracts, however, most of U6 is complexed with U4 snRNA. A large fraction of singular U6 snRNPs leaks out of the nuclei during extract preparation. This leaked fraction of the U6 snRNP might be present in cytoplasmic extracts, such as S100.

The U6 snRNP from S100 extract was affinity-selected by a biotinylated RNA oligonucleotide (U6 bio-101-82) that binds close to the 3' end of U6 snRNA. That region is part of a large irregular stem-loop, and therefore the annealing of the antisense oligonucleotide should disrupt it at least partially. There is also a small amount of U4/U6 snRNP in HeLa S100 extract. The antisense oligonucleotide should then be able to co-select U4/U6 snRNP, moreover the corresponding region of U6 snRNA occurs in a single-stranded form according to the U4/U6 secondary structure model (see **Figure 4.24.** for a schematic representation of the affinity selection procedure). To avoid U4/U6 contamination of affinity-selected U6 snRNP, U4 snRNA was subjected to antisense oligonucleotide-directed RNase H cleavage before the annealing of the biotinylated oligonucleotide.

DNA oligonucleotides that target U4 in the regions of stem I (S1 anti-U4) and II (S2 anti-U4) had been chosen to disrupt the U4/U6 snRNP. The efficiency of U4 snRNA cleavage was tested in the presence of 1.5 μM DNA oligonucleotides directed against U4/U6 stem I (**Figure 4.25.**, lane S1), stem II (lane S2), or both of them (lane S1+S2). The U4 snRNA was quantitatively cleaved in the presence of S1 anti-U4 oligonucleotide, and partially in the presence of S1 anti-U4 oligonucleotide. RNase H cleavage of U4 snRNA, directed by both oligonucleotides was routinely carried out prior to U6 snRNP affinity selection.

Untreated extract (S100) and extract, in which U4 had been digested as described (S100ΔU4)

were used for the affinity selection of U6 snRNP (**Figure 4.26.**). Control selections (mock) in the absence of the biotinylated oligonucleotide were prepared in parallel. The immobilized complexes were eluted, and the RNA components analyzed by denaturing gel electrophoresis and silver staining (**Figure 4.26., panel A,** lanes E), in parallel with RNA prepared from the extracts before (lanes 1 and 2) and after the affinity-selection (lanes FT). Complexes, containing only U6 snRNA were purified from S100ΔU4 extract (lane 5). Those selected from S100 extract contained a small amount of U4 in addition to U6 snRNA (lane 9), suggesting that in this case U6 snRNP was contaminated by U4/U6 snRNP. The mock-selected material contained no U6 snRNA in both cases (lanes 3 and 7).

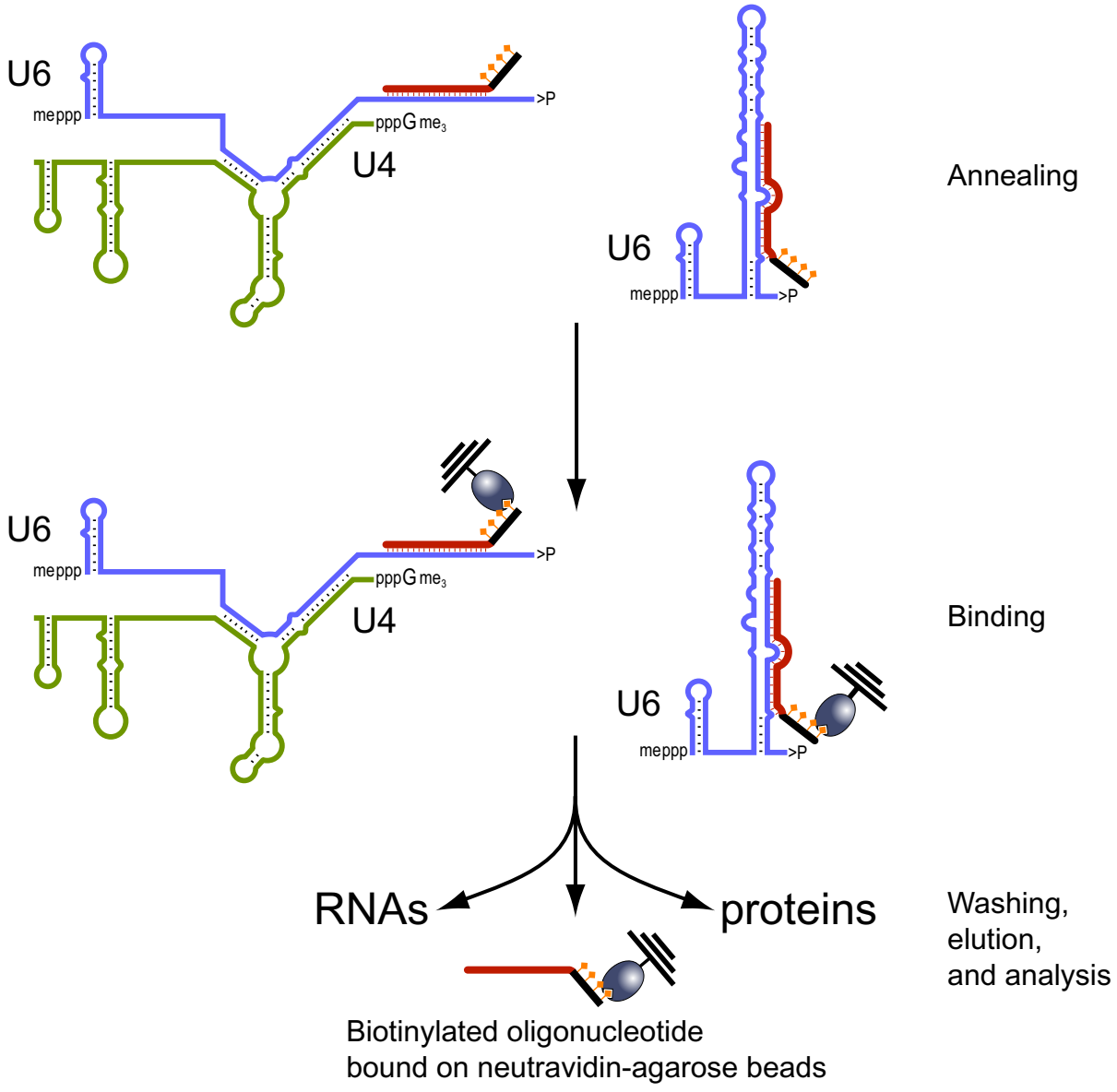


Figure 4.24. Schematic representation of U6 snRNP affinity selection.
U6 as well as U4/U6 snRNP are recognized by antisense 2'-*O*-methylated RNA oligonucleotide (shown in red). The oligonucleotide has four biotin groups (represented as orange squares) at its 5' end. The selected snRNPs are immobilized on neutravidin-agarose beads through interaction of the avidin (the blue oval) with the biotin groups. After washing, protein and RNA components are eluted under denaturing conditions. Due to the high-affinity biotin-neutravidin interaction, the antisense oligonucleotide remains immobilized.

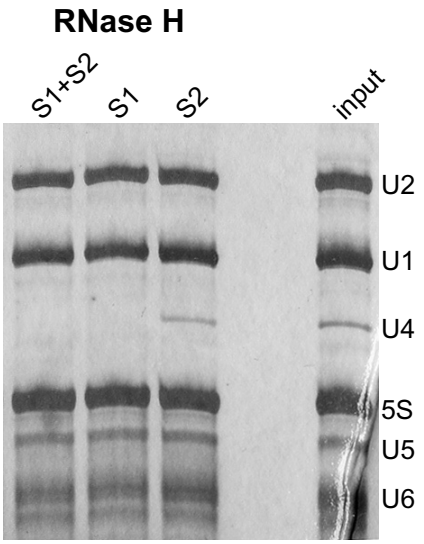


Figure 4.25. DNA oligonucleotide-directed RNase H cleavage of U4 snRNA.
HeLa S100 extract was incubated with RNase H in the presence of S1 anti-U4 oligonucleotide recognizing U4 snRNA in the region of U4/U6 stem I (lane S1), in the presence of S2 anti-U4 oligonucleotide that anneals to U4 snRNA in the region of U4/U6 stem II (lane S2), or in the presence of both of them (lane S1+S2). RNA from these reactions as well as from S100 extract (lane input) was prepared, separated by denaturing gel electrophoresis and silver-stained.

The protein components of the selected snRNPs were separated by SDS-PAGE and stained by Coomassie brilliant blue (**Figure 4.26., panel B**). The protein profiles of the selected snRNPs from S100ΔU4 and S100 extracts were essentially identical (compare lanes 3 and 5). Two protein bands appeared specifically in the selected, but not in the mock-selected material (compare lanes 2 and 3; 4 and 5). The lower one (indicated by an open arrowhead) ran at approximately 50 kDa was identified as La protein (see below), the upper band (marked by a black arrowhead) represented a protein with apparent molecular weight of 100 kDa. A protein of that size has not been reported as a component of U6 or U4/U6 snRNPs; therefore it might be a novel U6-binding protein. This band was gel-purified and analyzed by mass-spectrometry (see below).

Surprisingly, the other known U6-specific proteins, the Lsm complex and p110, were not detectable by Coomassie staining in the selected U6 snRNP. To confirm the absence of p110 and the identity of the 50 kDa-band, the protein eluates were analyzed by Western blot with anti-p110 antiserum and anti-La autoimmune serum (**Figure 4.27.**). The La protein was confirmed as a component of U6 snRNPs, selected both from S100ΔU4 and S100 extracts. It was not detected in mock-selected material (**panel A**). p110 protein was not found in U6 snRNPs selected from S100ΔU4 as well as from S100 extracts (**panel B**). The Y12 monoclonal antibody, which recognizes primarily the Sm B/B' and less efficiently Sm D proteins gave a weak positive signal only with the snRNP selected from S100, but not from S100ΔU4 extract (**panel C**). The Sm complex is a known component of U4/U6 snRNP. The latter finding therefore correlates with the observed coselection of U4/U6 snRNP from S100 extract.

A cDNA (GenBank accession AK001239), cloned and sequenced from human teratocarcinoma cells, contains an open reading frame consisting of 759 amino acids (Ota *et al.*, 2004). Four peptides, identified by mass-spectrometric analysis of the 100 kDa-band, mapped to the hypothetical protein encoded by this reading frame (**Figure 4.28.**). The hypothetical protein (called p85 in the following) has a calculated molecular weight of 85 kDa. It is a basic protein with predicted isoelectric point of 9.69, which is consistent with the apparent molecular weight of 100 kDa in SDS gels. Analysis of the domain structure of p85 revealed that it contains four RRMs and thus constitute a *bona fide* RNA-binding protein (**Figure 4.28.**). p85 is weakly similar to the yeast NOP4, a protein that also contains four RRMs and is localized to the nucleoli.

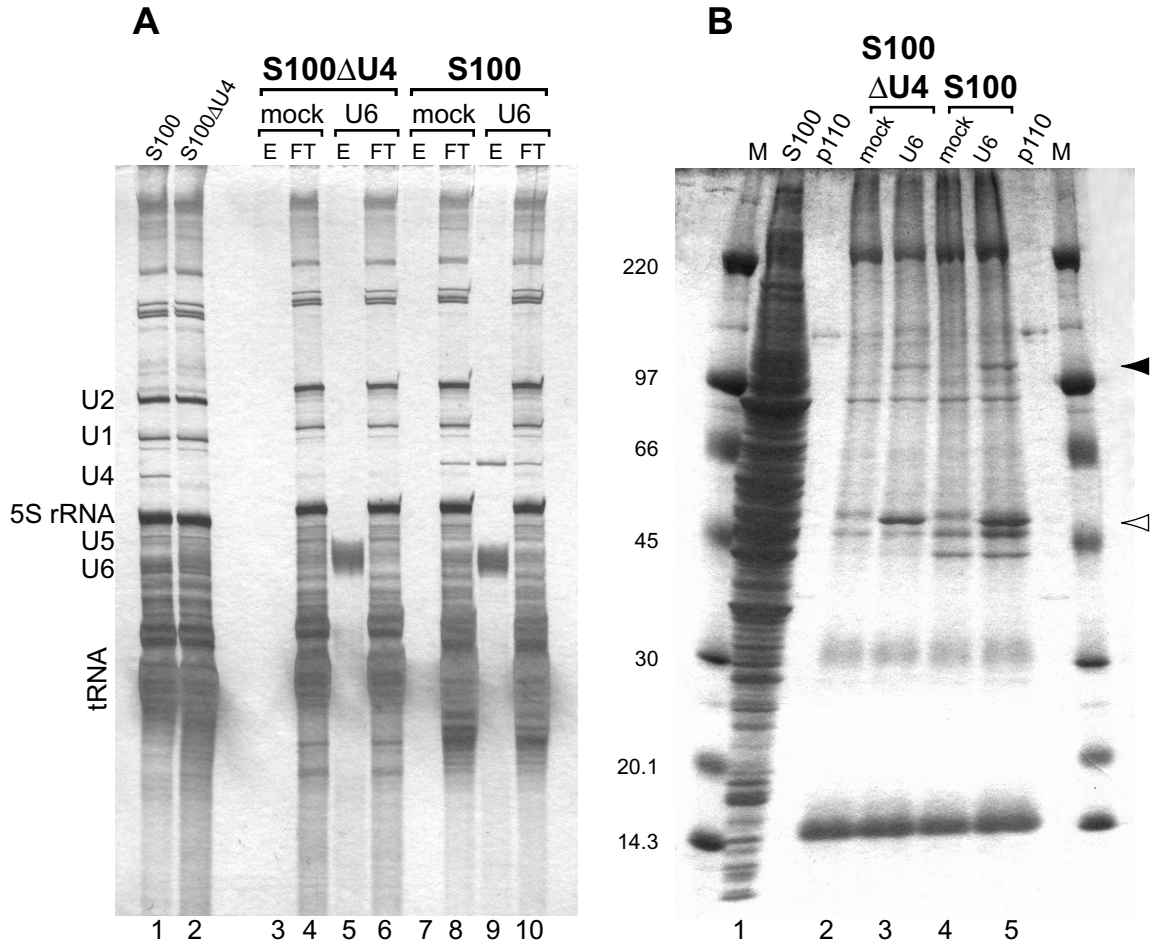


Figure 4.26. Affinity selection and analysis of U6 snRNP from S100 extract.
(A) The U6 snRNP was affinity-selected as shown on **Figure 4.24.** from HeLa S100 extract (S100, lanes 1, 7-10) or from S100 extract, preincubated with RNase H and oligonucleotides S1 anti-U4 and S2 anti-U4 (S100ΔU4, lanes 2, 3-6). The selection was carried out in the presence of biotinylated antisense 2'-*O*-methylated RNA oligonucleotide U6 bio-101-82 (U6, lanes 5-6 and 9-10). Mock-selections in the absence of the biotinylated oligonucleotide (mock, lanes 3-4 and 7-8) were done in parallel. RNA was prepared from 2 % of the extract before (lanes 1-2) or after binding of the U6 snRNP (FT, flow-through, lanes 4, 6, 8, and 10), as well as from the eluates (E, lanes 3, 5, 7, and 9). The RNA was analyzed by denaturing gel electrophoresis and silver staining. The prominent RNA bands are designated on the left side.
(B) The protein components of the eluted snRNPs were separated by SDS-PAGE with a 6-20% polyacrylamide gradient gel and Coomassie-stained. Lanes 2, 3, 4, and 5 correspond to lanes 3, 5, 7, and 9 shown in **(A)**. Input S100 extract (S100, lane 1), molecular weight marker (lanes M), and recombinant p110 (lanes p110) were also separated in parallel. The La protein (see **Figure 4.27.**) is indicated by an open arrowhead; The 100 kDa-protein band, pointed by a black arrowhead was identified by mass-spectrometric analysis as the protein encoded by AK001239 mRNA.

NOP4 is required for the normal production of rRNA and of the large ribosomal subunits (Sun and Woolford, 1994; Sun and Woolford, 1997). Conserved homologues of p85 were found by database search in mice, rat, and other vertebrates, however, none of them has been functionally studied.

The localization and the RNA binding preferences of p85 were studied by expression of GST-tagged p85 in HEK293 cells. The CDS of p85 mRNA was cloned into pHIL vector, which bears a CMV promoter and an N-terminal GST-tag. The GST-p85 fusion protein was expressed by transient transfection in HEK293 cells. A control transfection with pHIL vector, encoding

only the GST protein, was also carried out. The subcellular localization of these proteins was determined 24 hours after transfection by indirect *in situ* immunofluorescence with anti-GST antibodies (in collaboration with P. Stoilov and S. Stamm, Friedrich-Alexander-Universität, Erlangen). The localization of the endogenous SR protein SC35 was checked in parallel (**Figure 4.29.**). The GST protein was found in the cytoplasm (top-left panel). GST-p85 was a localized to the nucleus, but it did not show the characteristic nuclear speckled pattern of SC35 (bottom panels; the nuclear speckles are the sites where pre-mRNA splicing occurs). Surprisingly, GST-p85 is not a nucleolar protein like the yeast NOP4; instead, it was found in the nucleoplasm and was enriched at the immediate surrounding of the nucleoli (top-right panel).

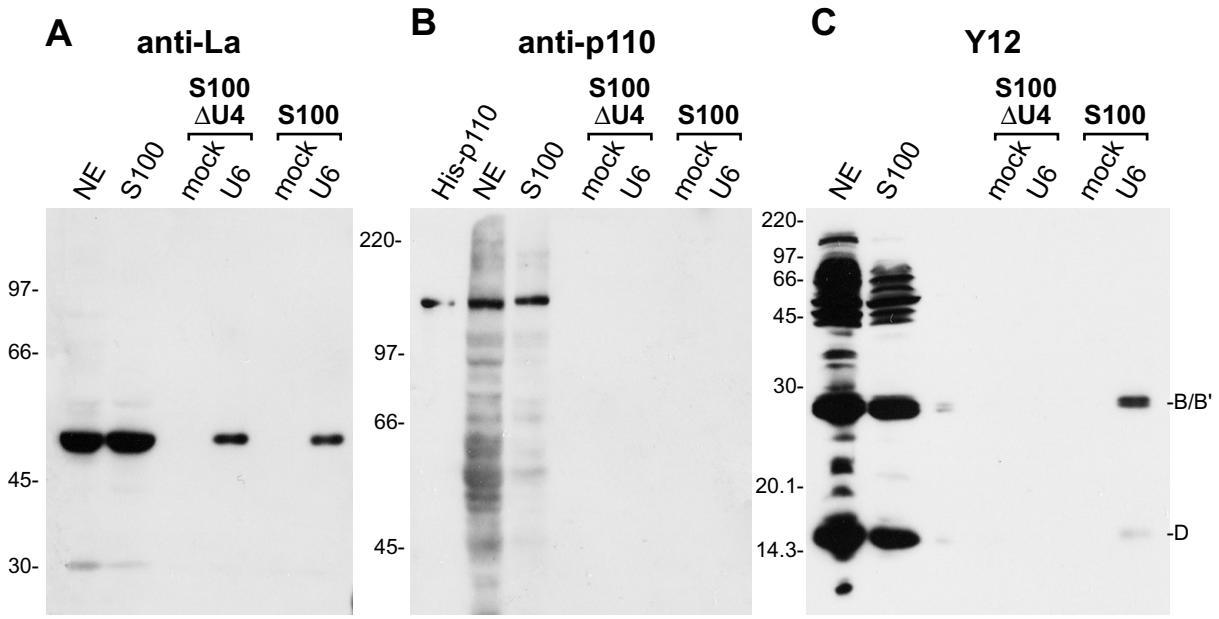


Figure 4.27. La protein, but not p110, is associated with the selected U6 snRNP.
The protein components of the U6snRNP (lanes U6), selected from S100ΔU4 and S100 extracts (as shown on **Figure 4.26, panel B**) were analyzed by Western blotting with anti-La autoimmune serum (**A**), anti-p110 antiserum (**B**), and Y12 monoclonal antibody (**C**). Control selections, carried out in the absence of the U6 bio-101-82 oligonucleotide (lanes mock) were tested in parallel. 2% of the input S100 extract (lanes S100), an equal volume of HeLa nuclear extract (lanes NE), and 100 ng of recombinant p110 (lane His-p110) were loaded as controls.

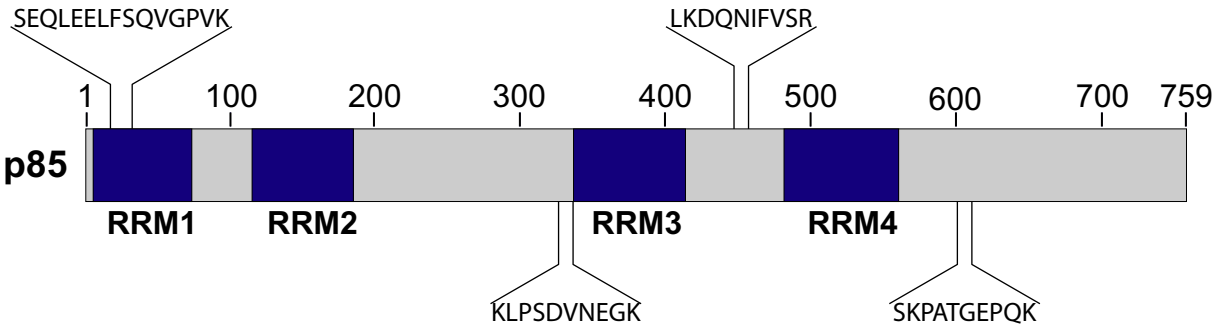


Figure 4.28. Domain organization of p85.
Four RNA-recognition motifs in p85, the translated product of AK01239 mRNA, were predicted by NCBI conserved domain search (http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi).
The internal peptides shown were identified by mass-spectrometric analysis of the 100 kDa-protein, co-selected with U6 snRNP. The numbers above indicate the positions of the amino acids. The RRMs and the positions of the peptides are drawn to scale.

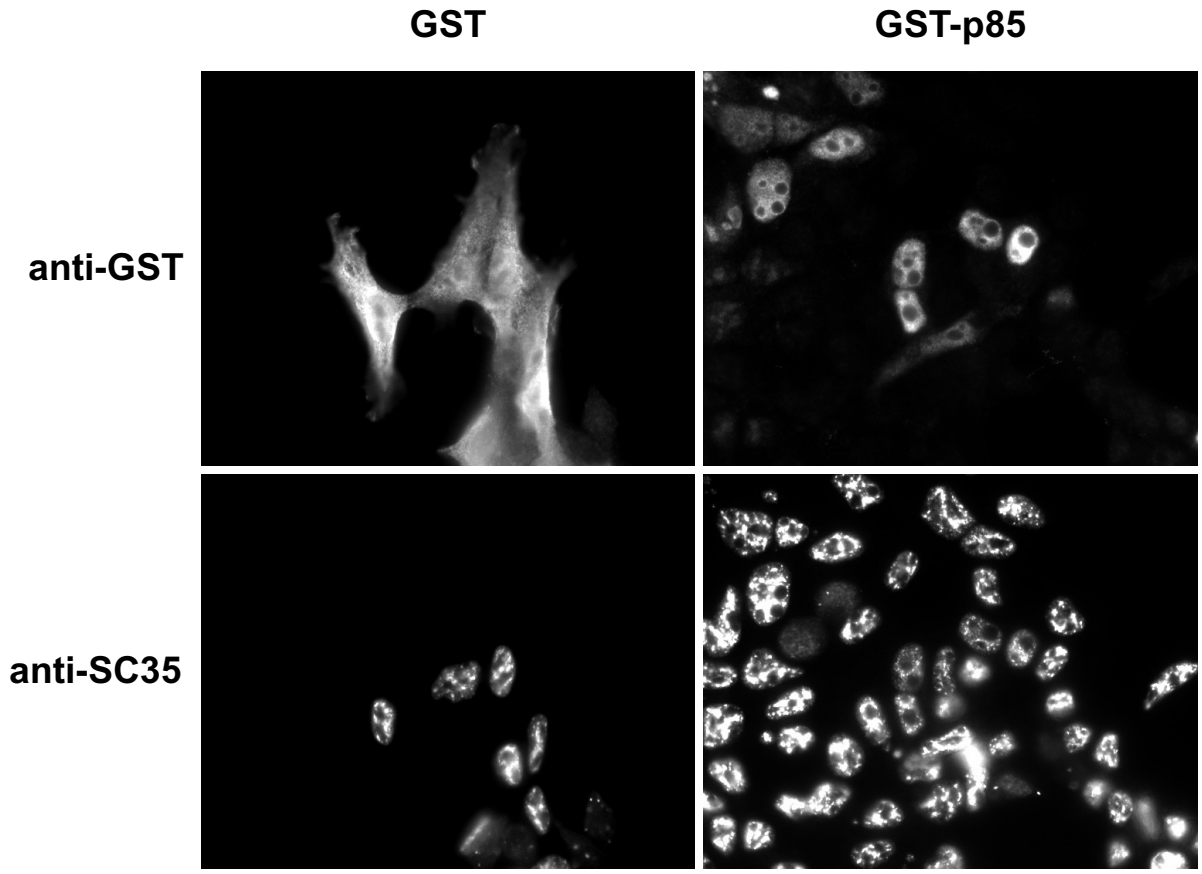


Figure 4.29. GST-p85 fusion protein, expressed in HEK293 cells is localized to the nucleoplasm and enriched at the nucleolar border.
HEK293 cells were transiently transfected with plasmids, encoding GST (left column), or GST-p85 (right column). The localization of these proteins was determined 24 hours after transfection by indirect *in situ* immunofluorescence using anti-GST antibodies (top row). The localization of the endogenous SC35 was determined in parallel by anti-SC35 antibodies (bottom row).

The RNA-binding specificity of GST-p85 was determined by a GST pull-down assay: HEK293 cells were transfected with expression vectors encoding GST or GST-p85 proteins, respectively, and harvested after 36 hours. Cellular extracts, containing both cytoplasmic and nuclear material were prepared and incubated with glutathione-Sepharose beads. The beads were washed with buffers containing 100, 250, or 500 mM KCl. The protein and RNA components of the immobilized complexes were eluted and analyzed by SDS-PAGE and by denaturing gel electrophoresis, followed by silver staining and Northern blotting, respectively (**Figure 4.30.**). The eluted proteins are shown on **panel A**. GST (marked by a black arrowhead) was expressed in much higher levels than GST-p85 (open arrowhead). The apparent molecular mass of GST-p85 is 130 kDa, consistent with the size of the GST-tag (29 kDa) and the apparent molecular mass of p85 (approximately 100 kDa; see **Figure 4.26., panel B**). RNAs co-precipitating with GST and GST-p85 were separated by denaturing gel electrophoresis and silver-stained (**Figure 4.30., panel B**). Total RNA, prepared from the cellular extracts (lanes 3-4 and 8-9) and from HeLa nuclear extract (lanes 1-2) was analyzed in parallel. Only minor amounts of some high-molecular weight RNAs, often found to precipitate non-specifically, were detected in the GST eluate (lanes 5-7). GST-p85, in contrast, was associated with many RNAs, primarily with 5S and 5.8S rRNAs (lanes 10-12).

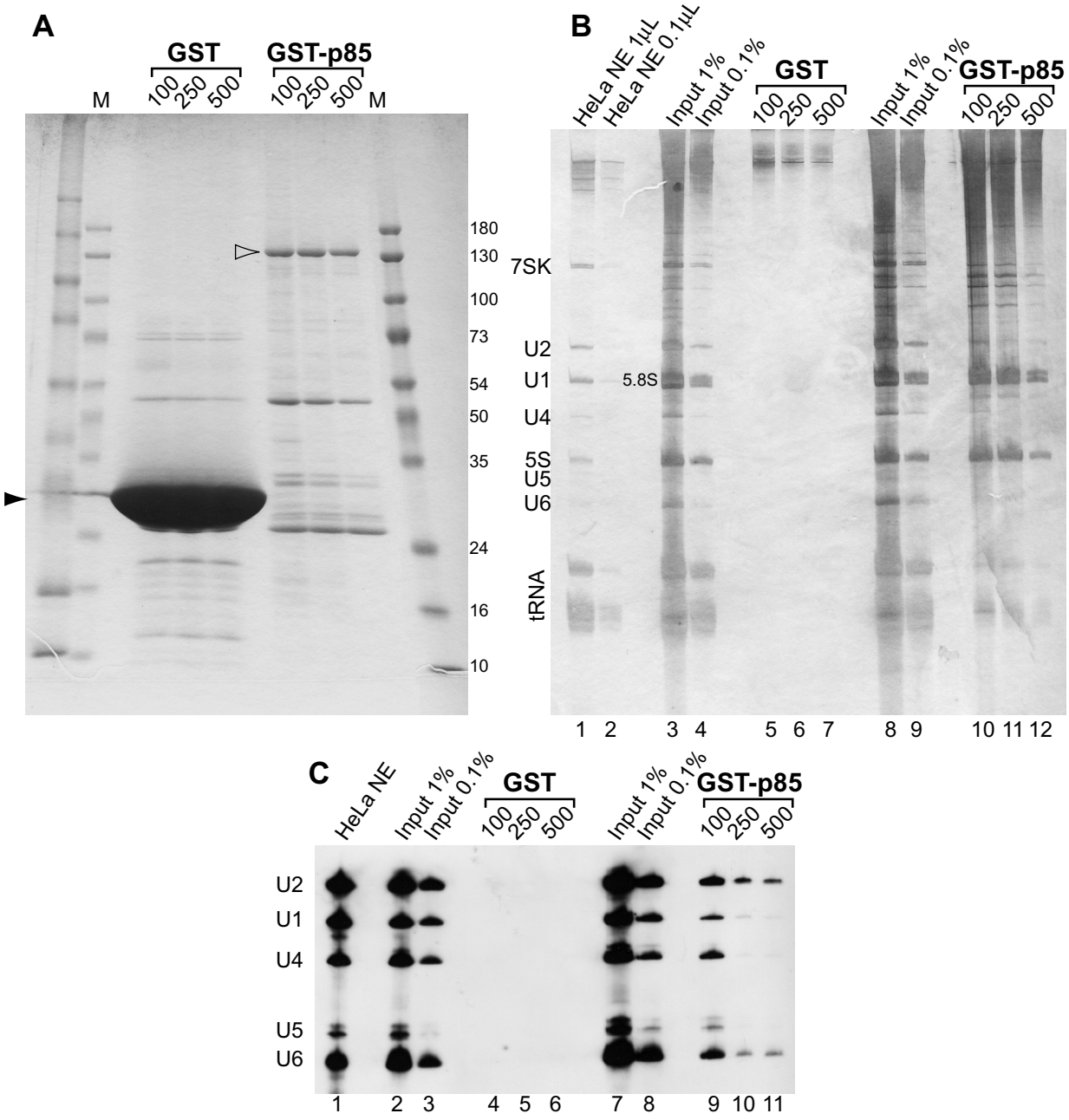


Figure 4.30. GST-p85 is associated primarily with 5S and 5.8S rRNAs, as well as with U2 and U6 snRNAs.

HEK293 cells were transfected with plasmids, encoding GST or GST-p85. The cells were harvested after 36 hour and total extracts were prepared. The complexes of GST (lanes GST) and GST-p85 (lanes GST-p85) were immobilized on glutathione-Sepharose, washed by buffers containing 100, 250, and 500 mM KCl, and eluted under denaturing conditions (lanes 100, 200, and 500, respectively).

(A) The protein components of the eluted complexes were analyzed by SDS-PAGE in a 6-20 % gradient gel, followed by Roti-Blue staining. GST is indicated by a black arrowhead, GST-p85 by an open arrowhead. The molecular weight marker (lanes M) sizes (10, 16, 24, 35, 50, 54, 73, 100, 130, and 180 kDa) are shown on the right side.

(B) RNA was prepared from the eluted GST and GST-p85 complexes (lanes 5-7 and 10-12, respectively), from 1 % or 0.1 % of the input extracts (GST, lanes 3-4 and GST-p85, lanes 8-9), as well as from 1 µl or 0.1 µl HeLa nuclear extract (lanes 1-2). The RNA samples were analyzed by electrophoresis in 8% denaturing gel and silver staining. The electrophoretic mobilities of 7SK, 5S and 5.8S rRNAs, tRNA, U1, U2, U4, U5, and U6 snRNAs are shown on the left.

(C) Aliquots from the same RNA samples, shown in (B), were analyzed by Northern blot using mixed probe detecting the major spliceosomal snRNAs (indicated on the left).

The major spliceosomal snRNAs were detected by Northern hybridization (**panel C**). GST protein did not bind to any of them (lanes 4-6). In 100 mM KCl, GST-p85 was complexed with U1, U2, U4, U6, and U6 snRNAs (lane 9). At more stringent conditions however, it remained associated only with U2 and U6 snRNAs (lanes 10 and 11).

Therefore p85 appears to be a potent RNA-binding protein, interacting preferentially with 5S and 5.8S rRNAs, but also with U2 and U6 snRNAs.

5. Discussion

5.1. Specificity and RNA-binding properties of U4/U6 recycling factors Prp24 and p110

The discovery that the human p110 protein represents the U4/U6 snRNP recycling factor (Bell *et al.*, 2002) led to the identification of a putative orthologs from several other species (see **Figure 1.7**). Unlike most of these proteins, Prp24 lacks the N-terminal HAT motifs and is only remotely similar to p110. In contrast to yeast and human recycling factors, U6 is the most highly conserved spliceosomal snRNA. There is an extensive sequence homology between human and yeast U6 snRNA (**Figure 5.1., panel B**); yeast U6 can even be partially aligned with human U6atac snRNA. (**panel C**). The secondary structures of human and east U4/U6 heteroduplexes, as well as of human U4atac/U6atac are conserved; the structures of their singular forms, however, are more variable (**panel A**).

Despite their differences, both Prp24 and p110 interact exclusively with U6 and U4/U6 snRNAs; p110 also binds to their snRNA counterparts of the minor spliceosome. The RNA sequence requirements of Prp24- and p110-binding are difficult to compare since they were studied by different approaches: footprinting and genetic mutational analysis in case of Prp24 (see below), and *in vitro* binding of shortened derivatives of U6 and U6atac snRNA in case of p110.

As shown by *in vitro* binding with shortened RNA derivatives, p110 interacts with human U6 between nucleotides 38 and 57 and with U6atac snRNA between nucleotides 10 and 30 (**Figures 4.3, 4.4., and 4.5.**). These two sequences represent the most highly conserved part between U6 and U6atac and are functionally related: They contain both the conserved ACAGAG box (or its corresponding atac version) and the nucleotides involved in stem I formation (**see Figure 4.6)**.

Hydroxyl radical footprinting revealed that Prp24 protects U6 snRNA between nucleotides 30 to 65 (Ghetti *et al.*, 1995). This footprinting analysis, however, did not show directly the RNA region required for Prp24 binding. The protected U6 stretch was discontinuous; the hydroxyl radical cleavage at some nucleotides within the region of nucleotides 30-65 was not affected at all by addition of Prp24, others were protected to a stronger or weaker extent. Therefore this protection effect could be due not only to decreased accessibility of the nucleotides involved in RNA-protein interaction, but also to structural alterations inflicted by the binding of Prp24.

The mutational analysis of yeast U6 snRNA also did not provide conclusive information about the minimal sequence requirements of Prp24-U6 interaction. U6 mutants within the region of nucleotides 36-43 were inefficient in interaction with Prp24 or defective in tri-snRNP formation (Jandrositz and Guthrie, 1995; Ryan *et al.*, 2002). Notably these mutations map within the RNA region, protected by Prp24.

The accumulated data are not sufficient to define the minimal RNA sequence motifs, required for interaction with Prp24 and p110. These motifs are narrowed down to 20 nucleotides for human U6 and U6atac snRNAs, and to 35 nucleotides for yeast U6 snRNA. Shorter sequence elements may be sufficient for binding. However, Prp24 has two RRMs more than p110 and this might explain the more extended sequence element it can protect. Nevertheless, there is a clear

similarity between p110 and Prp24 binding sites: The ACAGAG box and stem I nucleotides are reqired for RNA-protein interaction (**Figure 5.1., panel B**). Yeast U6 and human U6atac snRNA are too diverged, so that the identified sequence elements could be aligned only partially (**Figure 5.1., panel C**). The above data indicate that Prp24 and p110 may interact with their recycling substrates in an analogous way. Another observation supports this hypothesis: Mutations in the 3' part of the telestem or the 3' stem-loop that lead to disruption of base pairing increase the stability of the U6-Prp24 complex (Ryan *et al.*, 2002). Likewise, destabilization of the large 3'

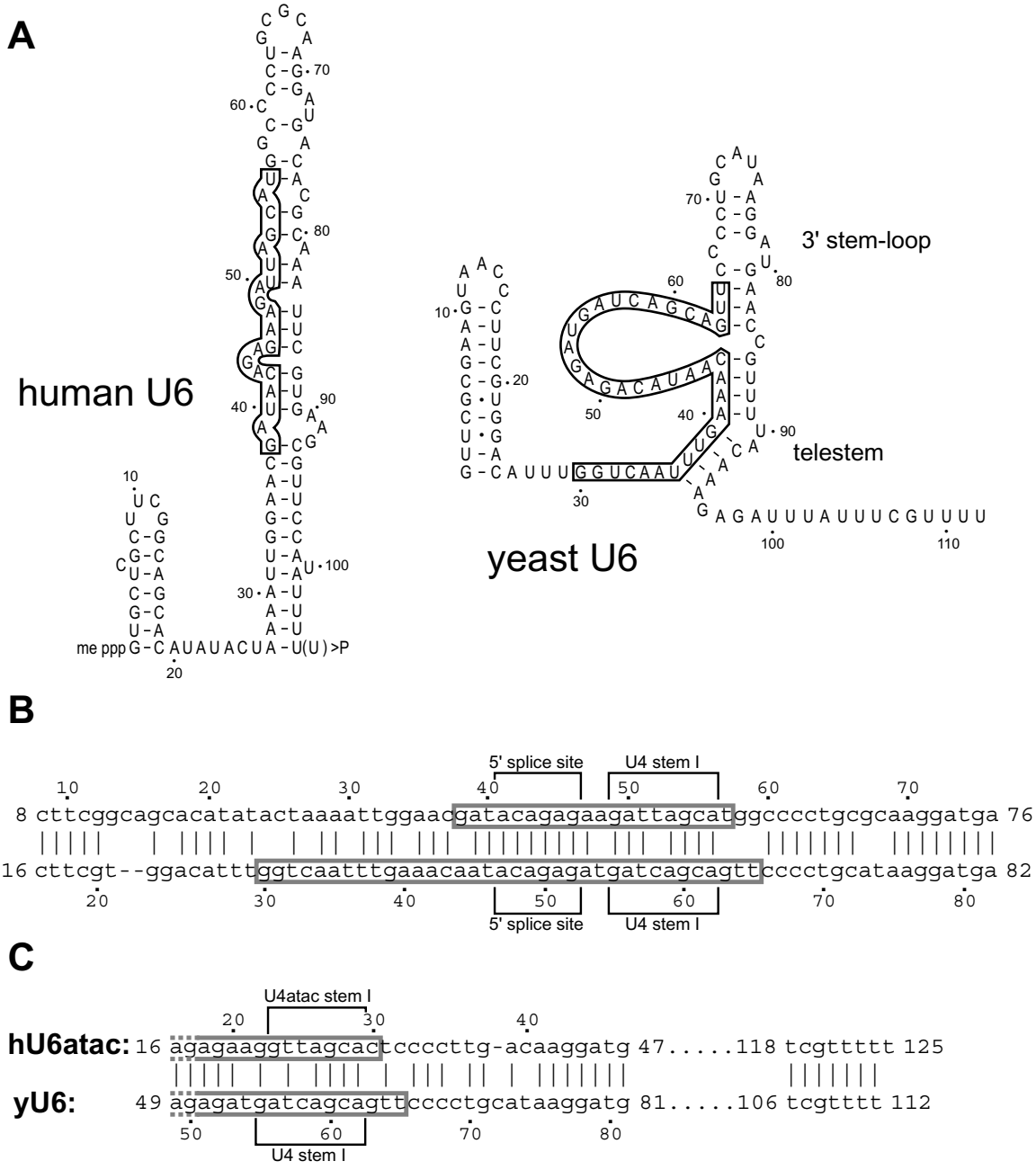


Figure 5.1. Prp24 and p110 share similar sequence requirements for U6 snRNA binding.
(A) Secondary structures of yeast and human U6 snRNA. The 3' stem-loop and the telestem are indicated in the structure of yeast U6.
(B) Alignment of human U6 (nucleotides 8 to 76) and yeast U6 (nucleotides 16 to 82).
(C) Alignment of human U6atac and yeast U6 snRNA. Only the parts sharing significant similarity are shown.
Prp24 and p110 binding regions are boxed on all panels. The sequences important for 5' splice site interaction and U4 base pairing are indicated on panels B and C.

terminal stem-loop of human U6 by annealing an antisense oligonucleotide to its 3'-terminal part greatly increases the binding affinity of p110 (compare **panels A and B on Figure 4.7.**). Therefore it might be hypothesized that Prp24 and p110 bind to the 5' part of these structures, resulting in destabilization and partial opening of the double-helical region. This is probably an essential prerequisite for the subsequent association of U4 (U4atac).

The sequence elements recognized by Prp24 and p110 do not participate in conserved RNA structure: p110 binds to a part of irregular stem in U6 and part of the 5' stem-loop of U6atac snRNA. Prp24 binds to the telestem and the large loop of yeast U6 snRNA. These three RNA elements have little in common besides their sequences and this is consistent with the theory that the protein stabilizes a more open structure of the RNA upon binding.

After heteroduplex formation, the secondary structures of human U4/U6, U4atac/U6atac, and yeast U4/U6 clearly resemble each other: A characteristic Y-shaped structure of two intermolecular stems and the 5'stem-loop of U4 (U4atac) snRNA is formed.

p110 interacts with 2- to 3-fold lower affinity with U4/U6 and U4atac/U6atac duplexes compared to their corresponding singular snRNAs (**Figure 4.7., panels B and C**). On the contrary, it has been found that the affinity of Prp24 for yeast U4/U6 heteroduplex is 5-fold higher than for the singular U6 snRNA (Ghetti *et al.*, 1995). The same authors also reported that the Prp24 footprinting patterns on U6 and U4/U6 snRNAs are different: When interacting with Prp24, the U4/U6 hybrid is protected in the region 39-57 of U6 and in stem II of both U6 (nucleotides 64-79), and U4 snRNA. The nucleotides of U6 participating in stem I formation were no longer protected by Prp24. Unfortunately, no information is available about the p110 binding sites on human U4/U6 and U4atac/U6atac heteroduplexes.

5.2. Dynamics of di-snRNP recycling: p110 is associated with U6, U6atac, and U4/U6, but not with U4atac/U6atac snRNP.

Immunoprecipitation studies using HeLa nuclear and S100 extracts revealed that p110 is a protein component of both U6 and U4/U6 snRNP. Regarding the minor spliceosome complexes, it was found almost exclusively in the singular U6atac snRNP; U4atac/U6atac snRNP contained background levels of p110 (**Figures 4.1 and 4.2.**). This is a striking difference: p110 is involved in recycling of both major and minor di-snRNPs, however, in case of the minor spliceosomal complexes it is associated with U6atac - the substrate of the recycling reaction; in case of the major complexes, with both the substrate and the recycled product, U4/U6 snRNP.

The differences described above are not due to low-affinity interaction of p110 with the U4atac/U6atac heteroduplex as shown on **Figure 4.7., panel B and C**: The apparent dissociation constants of p110 complexes with U4/U6 and U4atac/U6atac RNA hybrids were very similar: 12.1 ± 3.0 nM and 10.7 ± 2.3 nM, respectively. p110 has a two- to three-fold higher preference to bind U6 or U6atac snRNA than their corresponding hybrids with U4 or U4atac. Interestingly, the yeast recycling factor Prp24 is found primarily in complex with singular U6, but not with U4/U6 snRNP (Shannon and Guthrie, 1991), although it has an approximately five-fold greater affinity to U4/U6 heteroduplex than to U6 snRNA (Ghetti *et al.*, 1995). Obviously the stabilities

of these RNA-protein complexes do not have a critical role for the distribution of p110 and Prp24 between the substrates and the products of the recycling reaction.

A possible explanation for this phenomenon could be the efficiency of the recycling reaction: If the restoration of the U4atac/U6atac is inefficient compared to that of the major U4/U6 dimer, the recycling factor p110 is expected to accumulate in U6atac snRNP. However, this hypothesis is not supported by experimental evidence. On the contrary, both major and minor snRNPs appear to be recycled very efficiently, since the vast majority of U4 (U4atac) is associated with U6 (U6atac) snRNA; singular U4 and U4atac snRNPs are virtually undetectable in nuclear and S100 extracts as shown by fractionation through glycerol and CsCl gradients (see **Figures 4.21. and 4.15.**).

The restoration of the di-snRNPs is only one step of the recycling pathway; it is followed by another event, in which the di-snRNPs combine with 20S U5 snRNP to reform the tri-snRNPs. U5 snRNP is in very large excess over U4atac/U6atac, but not over U4/U6 snRNP, thus U5 may associate efficiently with the minor di-snRNP, reducing its steady-state levels. To test that possibility, the relative amounts of singular, di- and tri-snRNPs of the major and minor snRNAs were determined by glycerol gradient sedimentation under native conditions (**Figure 4.9.**). Surprisingly, U4atac/U6atac di-snRNP comprised the predominant part of the U4atac and U6atac complexes, while U4atac/U6atac.U5 tri-snRNP accounted for less than 10%. In contrast, the U4/U6.U5 tri-snRNP was more abundant than the U4/U6 snRNP. This result shows that the low levels of p110-U4atac/U6atac complex are also not due to its consumption in the following recycling of the minor tri-snRNP.

The findings discussed above indicate a clear difference in the dynamics of major and minor tri-snRNP recycling (see schematic **Figure 5.2.**): p110 binds to both U6 and U6atac snRNPs. It is unknown whether that occurs independently of LSm complex reassociation or these processes are mutually enhanced. Next, the di-snRNPs are restored via a mechanism not yet clarified, involving association of U4 and U4atac snRNPs, respectively, and formation of intermolecular stems I and II. The tri-snRNPs are subsequently built up from di-snRNPs and U5 snRNP. Anti-p110 immunoprecipitation from gradient fractions revealed that a considerable portion of the U4/U6 snRNP co-precipitated in fraction 11 (**Figure 4.10.**). It could thus be concluded that most of the U4/U6 heteroduplex is bound either by p110 or by the U5 snRNP, which favours the possibility that p110 is displaced from the U4/U6 si-snRNP by U5, when the tri-snRNP is formed. In contrast to that, p110 leaves the U4atac/U6atac complex shortly after recycling. The formation of the minor tri-snRNP is not efficient (**see Figure 4.9.**) and as a result, U4atac/U6atac snRNP accumulates. It is not clear whether the absence of p110 from the minor di-snRNP contributes to the inefficient formation of U4atac/U6atac.U5 snRNP.

The latter findings are in accordance with the hypothesis that U12-type splicing could be a rate-limiting step of pre-mRNA processing (Patel *et al.*, 2002). Specifically, the inefficient minor tri-snRNP recycling may slow down the excision of U12-dependent introns.

At present it is unclear why p110 leaves U4atac/U6atac, but remains associated with U4/U6 until it is joined by U5 snRNP. None of the possible reasons discussed above could be experimentally

confirmed. The early dissociation of p110 from U4atac/U6atac snRNP then is likely to be due to subtle structural differences between the major and the minor di-snRNPs. Unfortunately, the U4atac/U6atac snRNP is poorly characterized in terms of its protein complexes. There is only indirect evidence that the U4/U6-specific proteins are also present in atac di-snRNPs (Schneider *et al.*, 2002). Whether these proteins are associated in analogy to the major U4/U6 snRNP, and whether there are any atac di-snRNP-specific factors, is unknown.

Notably, the minor di- and tri-snRNPs did not co-fractionate with their corresponding major snRNPs; the peaks of the minor snRNPs were shifted two fractions towards the top of the gradient (**Figure 4.9.**). This has not been observed previously (Schneider *et al.*, 2002), however, that study has been carried out with anti-TMG affinity-purified U snRNPs, but not directly with nuclear extract. Therefore these snRNPs may have been changed structurally or may have lost protein components during the purification process.

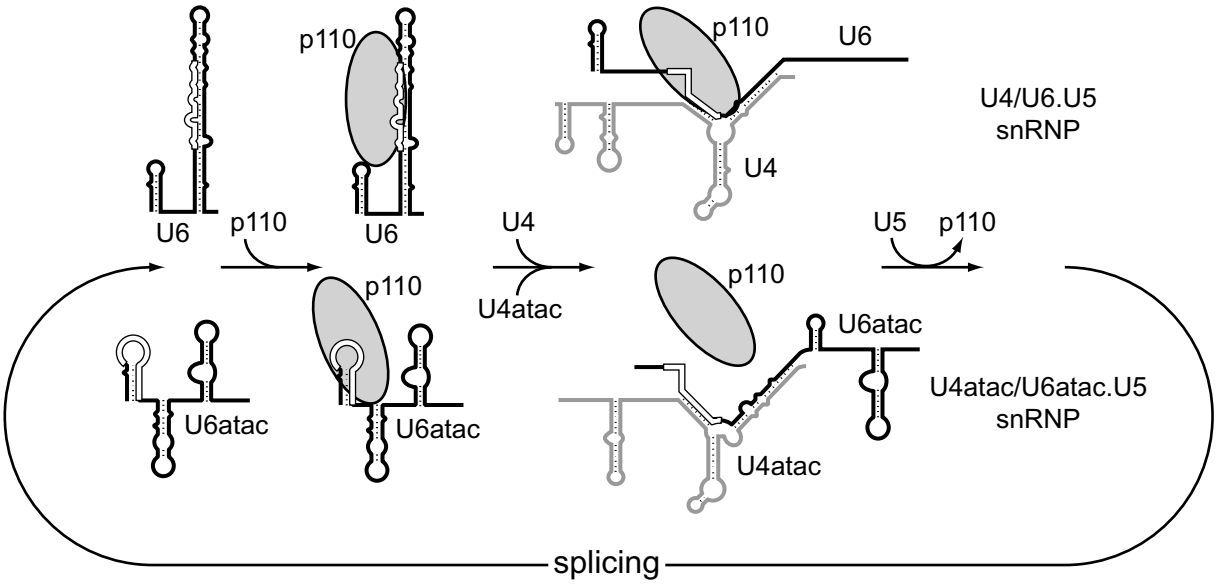


Figure 5.2. Model of U4/U6 and U4atac/U6atac snRNP recycling.
Binding and release of p110 is schematically represented during the post-spliceosomal transitions from the singular to the di-snRNP structures and further to the tri-snRNP complexes. The p110 binding regions in U6 and U6atac snRNA are indicated by open lines for both the singular and duplex snRNA structures.

5.3. Di-snRNP recycling *in vitro* and *in vivo*

The role of p110 for U4atac/U6atac snRNP recycling was studied under splicing conditions *in vitro*. First, the effect of p110 depletion on U12-dependent splicing *in vitro* was tested. At normal substrate concentration (2 ng per 25 µl-reaction), there was little difference between SCN4AENH1 pre-mRNA splicing efficiency in p110-depleted extract and depleted extract, complemented with recombinant p110 (**Figure 4.13.**, left side). However, when the amount of pre-mRNA was increased up to 10 ng per 25-µl reaction, a significant two-fold stimulation of splicing was observed (**Figure 4.13.**, right side). This result can be interpreted as p110 being important for splicing only with an excess of pre-mRNA, presumably because the recycling stage of the spliceosome cycle becomes limiting for U12-dependent splicing. Even under these recycling conditions, U4atac snRNA (the spliceosomal snRNA of lowest abundance) was in

large excess (approximately 70-fold) over the U12-dependent splicing substrate. Therefore it is likely that not all spliceosomal snRNPs are functional in splicing *in vitro*.

The effect of p110 depletion on U2-dependent splicing *in vitro* was similar (Medenbach *et al.*, unpublished results), although under the established recycling conditions (100 ng splicing substrate per 25-µl reaction) the stimulation by Minx pre-mRNA splicing by p110 was higher. This probably reflects the generally low splicing efficiency of SCN4AENH1 pre-mRNA.

The experiment discussed above indicates that p110 acts at some step during the recycling phase of minor spliceosome cycle. Taking into account that p110 binds to U6atac snRNA and that it has previously been found to recycle the major U4/U6 snRNP (Bell *et al.*, 2002), the most likely function of that protein in U12-dependent splicing is to recycle the U4atac/U6atac snRNP. This hypothesis was tested by U4atac/U6atac recycling assays, established the same way as for U4/U6 recycling (see schematic **Figure 4.14.**): The snRNPs accumulated after incubation of p110-depleted or mock-depleted nuclear extract under recycling conditions (see above) were fractionated through CsCl gradient and analyzed by Northern hybridization. The result of the experiment proved that p110 is also involved in minor di-snRNP recycling (**Figure 4.15.**); compared to U4/U6 recycling, lower concentrations of p110 protein were sufficient to restore U4atac/U6atac snRNP.

Under the high-salt conditions in the CsCl gradient, loosely bound proteins dissociate from the snRNP complexes, but the RNA-RNA interactions are not affected. An empirical formula has been derived to estimate the protein content of the fractionated snRNPs:

Protein % = (1.85 - ρ) / 0.006, where ρ is the buoyant density of the snRNP, expressed in g/ml (Spirin, 1969).

Using this equation, the protein components of U4atac snRNP found in gradient fraction 3 were estimated to approximately 130 kDa. Assuming that the protein composition of the U4atac/U6atac snRNP is similar to that of U4/U6 snRNP, the most likely interpretation of this result is that only the seven canonical Sm polypeptides (~115 kDa) and possibly the 15.5K protein remained bound to U4atac snRNA during CsCl density centrifugation. U6atac is probably stripped of the LSm proteins; the 61K protein as well as the 20K, 60K, and 90K proteins, which form a stable heteromeric complex (Horowitz *et al.*, 1997), probably did not remain in a complex with U4atac/U6atac. According to this model, U4atac and U6atac snRNAs in gradient fraction 3 do not represent an intermediate complex, in which they might be associated by protein-protein interactions. More likely, U4atac and U6atac are base-paired as in the fully recycled di-snRNP.

The levels of U6atac snRNA in p110-depleted extract after incubation under recycling conditions appeared to be much less than in the extracts containing endogenous or recombinant p110 (**Figure 4.15.**). Therefore the possibility that p110 stabilizes U6atac was considered. However, a control experiment (**Figure 4.16.**) showed that the decrease of U6atac levels in the absence of p110 was not due to degradation, but rather to loss in the CsCl gradient. Due to low splicing efficiency, the U4atac/U6atac recycling assay was carried out in the presence of PVA, which enhance splicing *in vitro* due to its volume-exclusion effect. PVA was not present in U4/U6 snRNP recycling reactions, and remarkably, the levels of U6 snRNA after centrifugation in CsCl gradient have not

been decreased (Bell *et al.*, 2002). When the splicing reaction was mixed with the CsCl solution, the PVA tended to coat the walls of the centrifuge tube and to form an upper layer on top of the gradient. As a result, U6atac snRNA has likely been lost partially due to absorption by the PVA coating.

To investigate the role of p110 *in vivo*, this protein was depleted in HeLa cells by RNAi knockdown. p110 was reduced to 10% of the control levels (**Figure 4.19.**), which resulted in specific inhibition of cell growth and decreased cell viability 72 hours after transfection. However, the distribution of singular U4 and U6 snRNPs, of U4/U6 di-snRNP as well as of their minor spliceosomal counterparts was not detectably altered after p110-knockdown (**Figure 4.20.**). This result is difficult to interpret, since gene expression, pre-mRNA processing, and mRNA export are closely coupled *in vivo* as discussed in section 1.3. Redundant activities in addition to p110 may be involved in di-snRNP restoration; alternatively, p110 may have another function essential for growth and cell survival, in addition to di-snRNP recycling.

It has been shown that postspliceosomal U4 (Bell *et al.*, 2002) and U4atac snRNPs (**Figure 4.15.**) accumulated in p110-depleted extract, incubated under U2- and U12-dependent recycling conditions, respectively. In extract containing endogenous levels of p110, these snRNPs were efficiently converted back to di-snRNPs and were never available for purification and characterization until now. U4atac occurs in a much lower abundance than U4 snRNA, therefore the purification of the postspliceosomal U4atac snRNP is not a realistic goal at the moment. Therefore the attention was focused on the U4 snRNP. At first, the sedimentation of postspliceosomal U4 snRNP was determined by glycerol gradient fractionation under native conditions: It peaked in fraction 7 as a 7-10S complex (**Figure 4.21.**, panel NEΔp110 postsplicing). Prior to splicing, U4 snRNA was present only in U4/U6 and U4/U6.U5 snRNPs with peaks in fractions 11 and 15, respectively (**Figure 4.21.**, panel NEΔp110); fraction 7 did not contain detectable levels of U4 snRNA. The singular U6 snRNP accumulated in fraction 5, either with or without incubation under recycling conditions. Therefore the postspliceosomal U4 snRNP is a slightly larger complex than the singular U6 snRNP.

Initial attempts were made to purify the U4 snRNP from p110-depleted extracts by anti-TMG immunoaffinity chromatography (**Figure 4.22.**) and by affinity selection, using an antisense biotinylated 2'-*O*-methylated RNA oligonucleotide (**Figure 4.23.**). However, until now the protein components of the purified postspliceosomal U4 snRNP are not identified; further experiments are required to shed light on the composition of this interesting complex.

5.4. "The pseudo-spliceosome cycle" : Disruption of U4/U6 and U4atac/ U6atac heteroduplexes independent of splicing

It has been shown that U2-dependent splicing was fully inhibited in extracts, incubated with U2-antisense DNA oligonucleotide (**Figure 4.11.**, panel Minx pre-mRNA). The oligonucleotide targeted the region of U2 snRNA, responsible for recognition of the branch site. It may be assumed that the spliceosome assembly was inhibited at A complex stage, before the tri-snRNP is able to join and form the B complex. Therefore splicing-dependent disruption of U4/U6 duplex is

not expected. Interestingly, when such extract, also depleted of p110, was incubated under U12-dependent recycling conditions, singular U4 and U6 snRNPs accumulated in addition of their atac versions (**Figure 4.18.**). This is a clear indication that U4/U6 heteroduplex can disintegrate independently of splicing. Previous studies revealed that both human and yeast tri-snRNP complexes are unstable in the presence of ATP . The release of U4 within the spliceosomal B complex is promoted by the ATP-dependent RNA helicase U5-200K (yeast Brr2; (Laggerbauer *et al.*, 1998; Raghunathan and Guthrie, 1998a), which is a component of U5 and U4/U6.U5 snRNPs. It is possible then that the same helicase can work independently of splicing to unwind the U4/U6 duplex within the tri-snRNP.

The equivalent test for splicing independent U4atac/U6atac disruption - after preincubation with anti-U12 oligonucleotide - was not performed. However, it is likely that an analogous process also occurs with that complex: When p110-depeted extract was incubated under recycling conditions in the absence of SCN4AENH1 pre-mRNA, a partial accumulation of singular U4atac and U6atac snRNPs was observed (**Figure 4.17.**, compare middle and right parts). It is not likely that this accumulation is due to splicing of endogenous U12-type pre-mRNA, as discussed in section 4.6., which leaves only the possibility that it occurs through a splicing-independent mechanism. Disintegration of the atac heteroduplex was ATP-dependent (**Figure 4.17.**, left part), strengthening the analogy with the disruption of U4/U6.

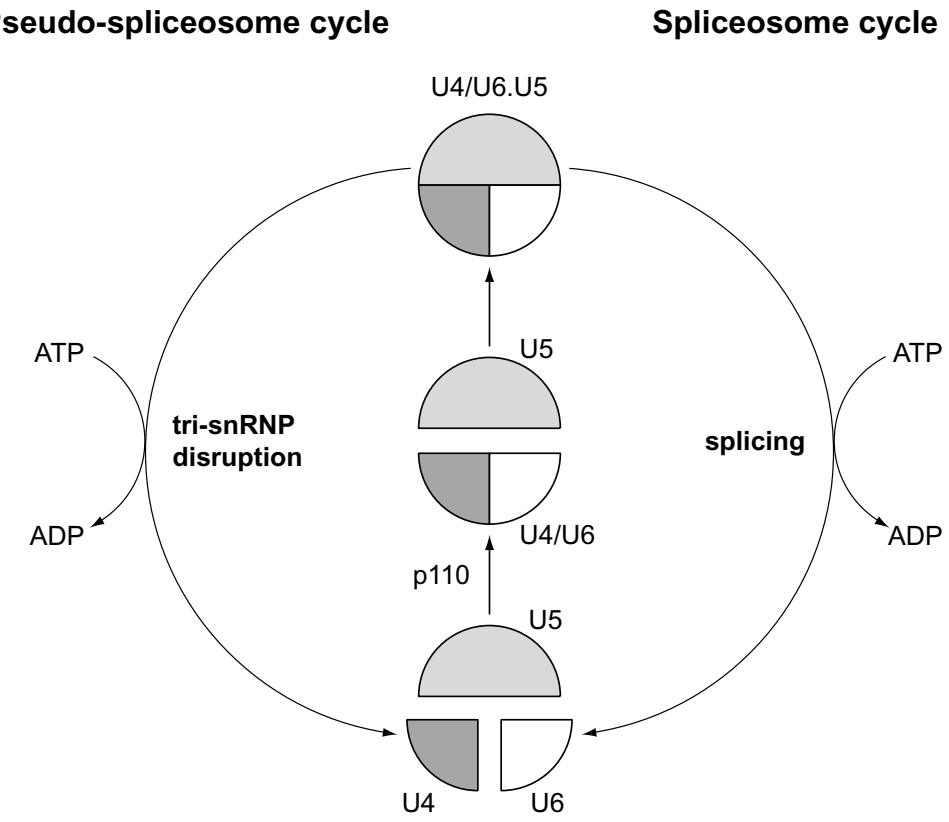


Figure 5.3. Splicing-dependent and splicing-independent disruption of tri-snRNP complexes.

The U4/U6.U5 snRNP is converted to singular U4, U5, and U6 snRNPs via two distinct pathways: spliceosome assembly and catalysis (right side), and ATP-dependent disruption (left side). The singular snRNPs produced either by splicing-dependent or splicing-independent mechanism are recycled back to tri-snRNPs (in the middle). The analogous cycles of the U4atac/U6atac.U5 snRNP are not shown.

In conclusion, the accumulation of singular snRNPs in p110-depleted extract, incubated under recycling conditions, appear to be due to two distinct ATP-dependent processes (**Figure 5.3.**). The first one is U2- or U12-dependent splicing, and second is tri-snRNP disintegration. The later process mimics the assembly and catalysis parts of the spliceosomal cycle as it produces singular U4(atac) and U6(atac) snRNPs, as well as U5 snRNP, probably in its 20S form. It can be considered as a shortcut that generates postspliceosomal snRNPs directly from the splicing-competent tri-snRNPs.

p110-dependent recycling of major and minor di-snRNP must therefore compensate for both splicing and tri-snRNP disruption. A second cycle, shorter than the spliceosome cycle, breaks down and restores the tri-snRNPs. Currently the role of this "pseudo-spliceosome" cycle is not clear, but it may provide a mechanism to control the splicing efficiency at recycling level.

5.5. Purification of the singular U6 snRNP from S100 extract: an intermediate in U6 snRNA maturation?

U6 occurs in excess of U4 snRNA, and for that reason considerable amounts of singular U6 snRNP are typically present in human cells as well as in cellular extracts. Yet the singular U6 is the least well-characterized spliceosomal snRNP. A complex of LSm proteins 2 to 8 (Achsel *et al.*, 1999), La protein (Pannone *et al.*, 1998) and p110 (Bell *et al.*, 2002) have previously been shown to associate with U6 snRNA, generating a heterogeneous population of U6 snRNPs.

It is believed that U6, unlike the other spliceosomal snRNPs, remains exclusively in the nucleus. However, a large fraction of U6 snRNP is found in the cytoplasmic S100 extract from HeLa cells, which probably is due to leakage during extract preparation. Therefore HeLa S100 could be a useful source for purification of the singular U6 snRNP, since there it is more abundant than the U4/U6 and U4/U6.U5 snRNPs.

Affinity selection using an antisense 2'-*O*-methylated biotinylated RNA oligonucleotide was employed to purify and characterize the U6 snRNP from S100 extract. Several oligonucleotides were tested, and one that anneals close to the 3' end of U6 was chosen, since it selected U6 snRNP efficiently (see schematic **Figure 4.24.**), with only minor contamination of U4/U6 snRNP (**Figure 4.26., panel A**). The co-selected U4/U6 snRNP was eliminated by antisense DNA oligonucleotide-mediated RNase H cleavage of U4 (**Figure 4.25.**). Interestingly, the selection from HeLa nuclear extract, using the same biotinylated RNA oligonucleotide, was very inefficient: Predominantly U4/U6 snRNP was selected (data not shown). This suggests that the singular U6 snRNPs in nuclear and S100 extracts differ not only in abundance, but also in composition and/or structure.

Neither p110 nor the LSm proteins, factors associated with U6 during the spliceosome cycle, were co-selected with U6 from S100 extracts. The absence of p110 was confirmed by Western blot (**Figure 4.27.**, middle panel); unfortunately, antibodies against the human LSm proteins were not available at that time. Two other proteins with apparent molecular masses of 50 and 100 kDa were specifically eluted with U6 snRNA (**Figure 4.26., panel B**): The first one was confirmed by Western blotting as the La protein (**Figure 4.27.**, left panel); the second one was identified by

mass-spectrometric analysis as p85, the translated product of AK001239 mRNA. La protein is a chaperone of the nascent polymerase III transcripts; it binds to their 3' polyuridinylated ends and increases their stability. p85 is a protein of unknown function. It is weakly similar to the yeast NOP4, a nucleolar protein involved in ribosomal RNA production and formation of the large ribosome subunits (Sun and Woolford, 1994; Sun and Woolford, 1997). The domain analysis of p85 predicted four RRM motifs, characteristic for RNA-binding proteins (**Figure 4.28.**).

The affinity selection did not provide definitive evidence that p85 was associated with U6, since it could have been bound also by the biotinylated RNA oligonucleotide, which was added in excess over U6 snRNA.

A GST-p85 fusion protein, transiently expressed in HEK293 cells, was localized to the nucleoplasm. Therefore p85 is not a nucleolar protein like the yeast NOP4, although it was enriched at the perinucleolar space (**Figure 4.29.**). The GST-pull down assay revealed that p85 is a potent RNA-binding protein, associated predominantly with 5S and 5.8S rRNAs (**Figure 4.30., panel B**), but also with U2 and U6 snRNAs (**Figure 4.30., panel C**). In sum, p85 represents a novel U6 snRNP protein. However, in contrast to p110 protein, it is not exclusively U6-specific, but interacts also with other RNAs.

As the presence of La protein indicates, the affinity-selected U6 complex is probably not functional in splicing - it is an immature U6 snRNP. p85 is concentrated around the nucleoli and shares limited homology with yeast NOP4, a putative rRNA processing factor. Interestingly, nascent U6 transcripts are transiently located to the nucleoli, where they are posttranscriptionally modified (Ganot *et al.*, 1999; Tycowski *et al.*, 1998). U2 is the most extensively pseudouridinylated and 2'-*O*-methylated snRNA, and the modifications at its 5' part are essential for splicing (Yu *et al.*, 1998). p85 could therefore be involved in modification of U2 and U6 snRNA or be nucleolar transport factor.

5.6. Perspectives

As discussed above, p110 was identified as the first human snRNP recycling factor involved in di-snRNP restoration. However, the detailed mechanism of di-snRNP recycling remains unclear: The recruitment of U4 (U4atac) snRNP to p110-U6 (U6atac) complex and the intermediate steps of RNA duplex formation are not characterized. It will be therefore important to reproduce the reaction in a purified system, which should further help to understand the mechanistic differences between major and minor snRNP recycling. For that purpose it will be necessary to identify the protein components of the postspliceosomal U4 snRNP.

Next, the function of the "pseudo-spliceosomal" cycle needs to be clarified, as it may represent an important part of the regulation of gene expression.

Regarding the singular U6 from HeLa S100 extract, it is unlikely that La and p85 are part of distinct U6 snRNPs that were co-selected. More likely, they are associated within one complex, which remains to be experimentally confirmed. Finally, the sequence requirements of p85 RNA binding, and the function of the p85-RNA complex will be investigated.

6. References

Abelson, J., Trotta, C.R. and Li, H. (1998) tRNA splicing. *J. Biol. Chem.*, **273**, 12685-12688.

Achsel, T., Ahrens, K., Brahms, H., Teigelkamp, S. and Luhrmann, R. (1998) The human U5-220kD protein (hPrp8) forms a stable RNA-free complex with several U5-specific proteins, including an RNA unwindase, a homologue of ribosomal elongation factor EF-2, and a novel WD-40 protein. *Mol. Cell. Biol.*, **18**, 6756-6766.

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.

Arenas, J.E. and Abelson, J.N. (1997) Prp43: An RNA helicase-like factor involved in spliceosome disassembly. *Proc. Natl. Acad. Sci. USA*, **94**, 11798-11802.

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

Bell, M., Schreiner, S., Damianov, A., Reddy, R. and Bindereif, A. (2002) p110, a novel human U6 snRNP protein and U4/U6 snRNP recycling factor. *EMBO J.*, **21**, 2724-2735.

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

Berglund, J.A., Chua, K., Abovich, N., Reed, R. and Rosbash, M. (1997) The splicing factor BBP interacts specifically with the pre-mRNA branchpoint sequence UACUAAC. *Cell*, **89**, 781-787.

Bindereif, A., Wolff, T. and Green, M.R. (1990) Discrete domains of human U6 snRNA required for the assembly of U4/U6 snRNP and splicing complexes. *EMBO J.*, **9**, 251-255.

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

Blatch, G.L. and Lassle, M. (1999) The tetratricopeptide repeat: a structural motif mediating protein-protein interactions. *Bioessays*, **21**, 932-939.

Blencowe, B.J., Bauren, G., Eldridge, A.G., Issner, R., Nickerson, J.A., Rosonina, E. and Sharp, P.A. (2000) The SRm160/300 splicing coactivator subunits. *RNA*, **6**, 111-120.

Booth, B.L., Jr. and Pugh, B.F. (1997) Identification and characterization of a nuclease specific for the 3' end of the U6 small nuclear RNA. *J. Biol. Chem.*, **272**, 984-991.

Burge, C.B., Padgett, R.A. and Sharp, P.A. (1998) Evolutionary fates and origins of U12-type introns. *Mol. Cell*, **2**, 773-785.

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.

Celotto, A.M. and Graveley, B.R. (2001) Alternative splicing of the Drosophila Dscam pre-mRNA is both temporally and spatially regulated. *Genetics*, **159**, 599-608.

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.

Chapman, K.B. and Boeke, J.D. (1991) Isolation and characterization of the gene encoding yeast debranching enzyme. *Cell*, **65**, 483-492.

Chech, T.R. and Golden, B.L. (1999) Building a catalytic active active site using only RNA. *The RNA world, second edition, Cold Spring Harbour Laboratoy Press, New York*, 321-349.

Chen, C. and Okayama, H. (1987) High-efficiency transformation of mammalian cells by plasmid DNA. *Mol. Cell. Biol.*, **7**, 2745-2752.

Cheng, S.C. and Abelson, J. (1987) Spliceosome assembly in yeast. *Genes Dev.*, **1**, 1014-1027.

Chiara, M.D., Palandjian, L., Feld Kramer, R. and Reed, R. (1997) Evidence that U5 snRNP recognizes the 3' splice site for catalytic step II in mammals. *EMBO J.*, **16**, 4746-4759.

Collins, C.A. and Guthrie, C. (1999) Allele-specific genetic interactions between Prp8 and RNA active site residues suggest a function for Prp8 at the catalytic core of the spliceosome. *Genes Dev.*, **13**, 1970-1982.

Collins, C.A. and Guthrie, C. (2000) The question remains: is the spliceosome a ribozyme? *Nat. Struct. Biol.*, **7**, 850-854.

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.

Cooper, T.A. (1999) In vivo SELEX in vertebrate cells. *Methods Mol. Biol.*, **118**, 405-417.

Coulter, L.R., Landree, M.A. and Cooper, T.A. (1997) Identification of a new class of exonic splicing enhancers by in vivo selection. *Mol. Cell. Biol.*, **17**, 2143-2150.

Cunningham, T.P., Hagan, J.P. and Grabowski, P.J. (1995) Reconstitution of exon-bridging activity with purified U2AF and U1 snRNP components. *Nucleic Acids Symp. Ser.*, 218-219.

Czaplinski, K., Ruiz-Echevarria, M.J., Paushkin, S.V., Han, X., Weng, Y., Perlick, H.A., Dietz, H.C., Ter-Avanesyan, M.D. and Peltz, S.W. (1998) The surveillance complex interacts with the translation release factors to enhance termination and degrade aberrant mRNAs. *Genes Dev.*, **12**, 1665-1677.

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

de la Mata, M., Alonso, C.R., Kadener, S., Fededa, J.P., Blaustein, M., Pelisch, F., Cramer, P., Bentley, D. and Kornblihtt, A.R. (2003) A slow RNA polymerase II affects alternative splicing in vivo. *Mol. Cell*, **12**, 525-532.

Derry, J.J., Richard, S., Valderrama Carvajal, H., Ye, X., Vasioukhin, V., Cochrane, A.W., Chen, T. and Tyner, A.L. (2000) Sik (BRK) phosphorylates Sam68 in the nucleus and negatively regulates its RNA binding ability. *Mol. Cell. Biol.*, **20**, 6114-6126.

Deutsch, M. and Long, M. (1999) Intron-exon structures of eukaryotic model organisms. *Nucleic Acids Res.*, **27**, 3219-3228.

Dietrich, R.C., Incorvaia, R. and Padgett, R.A. (1997) Terminal intron dinucleotide sequences do not distinguish between U2- and U12-dependent introns. *Mol. Cell*, **1**, 151-160.

Dostie, J. and Dreyfuss, G. (2002) Translation is required to remove Y14 from mRNAs in the cytoplasm. *Curr. Biol.*, **12**, 1060-1067.

Du, C., McGuffin, M.E., Dauwalder, B., Rabinow, L. and Mattox, W. (1998) Protein phosphorylation plays an essential role in the regulation of alternative splicing and sex determination in Drosophila. *Mol. Cell*, **2**, 741-750.

Du, H. and Rosbash, M. (2002) The U1 snRNP protein U1C recognizes the 5' splice site in the absence of base pairing. *Nature*, **419**, 86-90.

Elliott, D.J., Bourgeois, C.F., Klink, A., Stevenin, J. and Cooke, H.J. (2000) A mammalian germ cell-specific RNA-binding protein interacts with ubiquitously expressed proteins involved in splice site selection. *Proc. Natl. Acad. Sci. U S A*, **97**, 5717-5722.

Eperon, I.C., Ireland, D.C., Smith, R.A., Mayeda, A. and Krainer, A.R. (1993) Pathways for selection of 5' splice sites by U1 snRNPs and SF2/ASF. *EMBO J.*, **12**, 3607-3617.

Fairbrother, W.G. and Chasin, L.A. (2000) Human genomic sequences that inhibit splicing. *Mol. Cell. Biol.*, **20**, 6816-6825.

Fairbrother, W.G., Yeh, R.F., Sharp, P.A. and Burge, C.B. (2002) Predictive identification of exonic splicing enhancers in human genes. *Science*, **297**, 1007-1013.

Fischer, U., Liu, Q. and Dreyfuss, G. (1997) The SMN-SIP1 complex has an essential role in spliceosomal snRNP biogenesis. *Cell*, **90**, 1023-1029.

Fischer, U., Sumpter, V., Sekine, M., Satoh, T. and Luhrmann, R. (1993) Nucleo-cytoplasmic

transport of U snRNPs: definition of a nuclear location signal in the Sm core domain that binds a transport receptor independently of the m3G cap. *EMBO J.*, **12**, 573-583.

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

Fong, Y.W. and Zhou, Q. (2001) Stimulatory effect of splicing factors on transcriptional elongation. *Nature*, **414**, 929-933.

Friesen, W.J., Massenet, S., Paushkin, S., Wyce, A. and Dreyfuss, G. (2001a) SMN, the product of the spinal muscular atrophy gene, binds preferentially to dimethylarginine-containing protein targets. *Mol. Cell*, **7**, 1111-1117.

Friesen, W.J., Paushkin, S., Wyce, A., Massenet, S., Pesiridis, G.S., Van Duyne, G., Rappsilber, J., Mann, M. and Dreyfuss, G. (2001b) The methylosome, a 20S complex containing JBP1 and pICln, produces dimethylarginine-modified Sm proteins. *Mol. Cell. Biol.*, **21**, 8289-8300.

Fromont-Racine, M., Mayes, A.E., Brunet-Simon, A., Rain, J.C., Colley, A., Dix, I., Decourty, L., Joly, N., Ricard, F., Beggs, J.D. and Legrain, P. (2000) Genome-wide protein interaction screens reveal functional networks involving Sm-like proteins. *Yeast*, **17**, 95-110.

Fu, X.D. and Maniatis, T. (1992) The 35-kDa mammalian splicing factor SC35 mediates specific interactions between U1 and U2 small nuclear ribonucleoprotein particles at the 3' splice site. *Proc. Natl. Acad. Sci. USA*, **89**, 1725-1729.

Ganot, P., Jady, B.E., Bortolin, M.L., Darzacq, X. and Kiss, T. (1999) Nucleolar factors direct the 2'-O-ribose methylation and pseudouridylation of U6 spliceosomal RNA. *Mol. Cell. Biol.*, **19**, 6906-6917.

Ghetti, A., Company, M. and Abelson, J. (1995) Specificity of Prp24 binding to RNA: a role for Prp24 in the dynamic interaction of U4 and U6 snRNAs. *RNA*, **1**, 132-145.

Golas, M.M., Sander, B., Will, C.L., Luhrmann, R. and Stark, H. (2003) Molecular architecture of the multiprotein splicing factor SF3b. *Science*, **300**, 980-984.

Gondran, P. and Dautry, F. (1999) Regulation of mRNA splicing and transport by the tyrosine kinase activity of src. *Oncogene*, **18**, 2547-2555.

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.

Gu, J., Shimba, S., Nomura, N. and Reddy, R. (1998) Isolation and characterization of a new 110 kDa human nuclear RNA-binding protein (p110nrb). *Biochim. Biophys. Acta*, **1399**, 1-9.

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

Hanamura, A., Caceres, J.F., Mayeda, A., Franza, B.R., Jr. and Krainer, A.R. (1998) Regulated tissue-specific expression of antagonistic pre-mRNA splicing factors. *RNA*, **4**, 430-444.

Hartmann, A.M., Nayler, O., Schwaiger, F.W., Obermeier, A. and Stamm, S. (1999) The interaction and colocalization of Sam68 with the splicing-associated factor YT521-B in nuclear dots is regulated by the Src family kinase p59(fyn). *Mol. Biol. Cell*, **10**, 3909-3926.

Hartmuth, K., Urlaub, H., Vornlocher, H.P., Will, C.L., Gentzel, M., Wilm, M. and Luhrmann, R. (2002) Protein composition of human prespliceosomes isolated by a tobramycin affinity-selection method. *Proc. Natl. Acad. Sci. USA*, **99**, 16719-16724.

Hertel, K.J. and Maniatis, T. (1998) The function of multisite splicing enhancers. *Mol. Cell*, **1**, 449-455.

Hide, W.A., Babenko, V.N., van Heusden, P.A., Seoighe, C. and Kelso, J.F. (2001) The

contribution of exon-skipping events on chromosome 22 to protein coding diversity. *Genome Res.*, **11**, 1848-1853.

Hoffman, B.E. and Grabowski, P.J. (1992) U1 snRNP targets an essential splicing factor, U2AF65, to the 3' splice site by a network of interactions spanning the exon. *Genes Dev.*, **6**, 2554-2568.

Horowitz, D.S., Kobayashi, R. and Krainer, A.R. (1997) A new cyclophilin and the human homologues of yeast Prp3 and Prp4 form a complex associated with U4/U6 snRNPs. *RNA*, **3**, 1374-1387.

Huang, T., Vilardell, J. and Query, C.C. (2002) Pre-spliceosome formation in S.pombe requires a stable complex of SF1-U2AF(59)-U2AF(23). *EMBO J.*, **21**, 5516-5526.

Huber, J., Cronshagen, U., Kadokura, M., Marshallsay, C., Wada, T., Sekine, M. and Luhrmann, R. (1998) Snurportin1, an m3G-cap-specific nuclear import receptor with a novel domain structure. *EMBO J.*, **17**, 4114-4126.

Hui, J., Stangl, K., Lane, W.S. and Bindereif, A. (2003) HnRNP L stimulates splicing of the eNOS gene by binding to variable-length CA repeats. *Nat. Struct. Biol.*, **10**, 33-37.

Jamison, S.F., Pasman, Z., Wang, J., Will, C., Luhrmann, R., Manley, J.L. and Garcia-Blanco, M.A. (1995) U1 snRNP-ASF/SF2 interaction and 5' splice site recognition: characterization of required elements. *Nucleic Acids Res.*, **23**, 3260-3267.

Jandrositz, A. and Guthrie, C. (1995) Evidence for a Prp24 binding site in U6 snRNA and in a putative intermediate in the annealing of U6 and U4 snRNAs. *EMBO J.*, **14**, 820-832.

Jankowsky, E., Gross, C.H., Shuman, S. and Pyle, A.M. (2001) Active disruption of an RNA-protein interaction by a DExH/D RNA helicase. *Science*, **291**, 121-125.

Jensen, K.B., Dredge, B.K., Stefani, G., Zhong, R., Buckanovich, R.J., Okano, H.J., Yang, Y.Y. and Darnell, R.B. (2000) Nova-1 regulates neuron-specific alternative splicing and is essential for neuronal viability. *Neuron*, **25**, 359-371.

Jumaa, H. and Nielsen, P.J. (1997) The splicing factor SRp20 modifies splicing of its own mRNA and ASF/SF2 antagonizes this regulation. *EMBO J.*, **16**, 5077-5085.

Kim, D.H. and Rossi, J.J. (1999) The first ATPase domain of the yeast 246-kDa protein is required for in vivo unwinding of the U4/U6 duplex. *RNA*, **5**, 959-971.

Kim, S.H. and Lin, R.J. (1996) Spliceosome activation by PRP2 ATPase prior to the first transesterification reaction of pre-mRNA splicing. *Mol. Cell. Biol.*, **16**, 6810-6819.

Kleinschmidt, A.M., Patton, J.R. and Pederson, T. (1989) U2 small nuclear RNP assembly in vitro. *Nucleic Acids Res.*, **17**, 4817-4828.

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

Konarska, M.M. and Sharp, P.A. (1987) Interactions between small nuclear ribonucleoprotein particles in formation of spliceosomes. *Cell*, **49**, 763-774.

Kramer, A. (1996) The structure and function of proteins involved in mammalian pre-mRNA splicing. *Ann. Rev. Biochem.*, **65**, 367-409.

Krecic, A.M. and Swanson, M.S. (1999) hnRNP complexes: composition, structure, and function. *Curr. Opin. Cell Biol.*, **11**, 363-371.

Ladd, A.N. and Cooper, T.A. (2002) Finding signals that regulate alternative splicing in the post-genomic era. *Genome Biol.*, **3(11)**, Review 8.

Laggerbauer, B., Achsel, T. and Luhrmann, R. (1998) The human U5-200kD DEXH-box protein unwinds U4/U6 RNA duplices in vitro. *Proc. Natl. Acad. Sci. USA*, **95**, 4188-4192.

Lander, E.S., Linton, L.M., Birren, B., Nusbaum, C., Zody, M.C., Baldwin, J. *et al.* (2001) Initial sequencing and analysis of the human genome. *Nature*, **409**, 860-921.

Le Hir, H., Izaurralde, E., Maquat, L.E. and Moore, M.J. (2000) The spliceosome deposits

multiple proteins 20-24 nucleotides upstream of mRNA exon-exon junctions. *EMBO J.*, **19**, 6860-6869.

Lelay-Taha, M.N., Reveillaud, I., Sri-Widada, J., Brunel, C. and Jeanteur, P. (1986) RNA-protein organization of U1, U5 and U4-U6 small nuclear ribonucleoproteins in HeLa cells. *J. Mol. Biol.*, **189**, 519-532.

Levine, A. and Durbin, R. (2001) A computational scan for U12-dependent introns in the human genome sequence. *Nucleic Acids Res.*, **29**, 4006-4013.

Lewis, J.D., Izaurralde, E., Jarmolowski, A., McGuigan, C. and Mattaj, I.W. (1996) A nuclear cap-binding complex facilitates association of U1 snRNP with the cap-proximal 5' splice site. *Genes Dev.*, **10**, 1683-1698.

Lim, L.P. and Burge, C.B. (2001) A computational analysis of sequence features involved in recognition of short introns. *Proc. Natl. Acad. Sci. USA*, **98**, 11193-11198.

Linder, P., Tanner, N.K. and Banroques, J. (2001) From RNA helicases to RNPases. *Trends Biochem. Sci.*, **26**, 339-341.

Liu, H.X., Chew, S.L., Cartegni, L., Zhang, M.Q. and Krainer, A.R. (2000) Exonic splicing enhancer motif recognized by human SC35 under splicing conditions. *Mol. Cell. Biol.*, **20**, 1063-1071.

Liu, H.X., Zhang, M. and Krainer, A.R. (1998) Identification of functional exonic splicing enhancer motifs recognized by individual SR proteins. *Genes Dev.*, **12**, 1998-2012.

Liu, Q., Fischer, U., Wang, F. and Dreyfuss, G. (1997) The spinal muscular atrophy disease gene product, SMN, and its associated protein SIP1 are in a complex with spliceosomal snRNP proteins. *Cell*, **90**, 1013-1021.

Lund, E. and Dahlberg, J.E. (1992) Cyclic 2',3'-phosphates and nontemplated nucleotides at the 3' end of spliceosomal U6 small nuclear RNA's. *Science*, **255**, 327-330.

Lynch, K.W. and Weiss, A. (2000) A model system for activation-induced alternative splicing of CD45 pre-mRNA in T cells implicates protein kinase C and Ras. *Mol. Cell. Biol.*, **20**, 70-80.

Makarov, E.M., Makarova, O.V., Urlaub, H., Gentzel, M., Will, C.L., Wilm, M. and Luhrmann, R. (2002) Small nuclear ribonucleoprotein remodeling during catalytic activation of the spliceosome. *Science*, **298**, 2205-2208.

Makarova, O.V., Makarov, E.M., Liu, S., Vornlocher, H.P. and Luhrmann, R. (2002) Protein 61K, encoded by a gene (PRPF31) linked to autosomal dominant retinitis pigmentosa, is required for U4/U6*U5 tri-snRNP formation and pre-mRNA splicing. *EMBO J.*, **21**, 1148-1157.

Malca, H., Shomron, N. and Ast, G. (2003) The U1 snRNP base pairs with the 5' splice site within a penta-snRNP complex. *Mol. Cell. Biol.*, **23**, 3442-3455.

Manley, J.L. (2002) Nuclear coupling: RNA processing reaches back to transcription. *Nat. Struct. Biol.*, **9**, 790-791.

Maroney, P.A., Romfo, C.M. and Nilsen, T.W. (2000) Functional recognition of 5' splice site by U4/U6.U5 tri-snRNP defines a novel ATP-dependent step in early spliceosome assembly. *Mol. Cell*, **6**, 317-328.

Massenet, S. and Branlant, C. (1999) A limited number of pseudouridine residues in the human atac spliceosomal UsnRNAs as compared to human major spliceosomal UsnRNAs. *RNA*, **5**, 1495-1503.

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

McCracken, S., Fong, N., Rosonina, E., Yankulov, K., Brothers, G., Siderovski, D., Hessel, A., Foster, S., Shuman, S. and Bentley, D.L. (1997a) 5'-Capping enzymes are targeted to pre-mRNA by binding to the phosphorylated carboxy-terminal domain of RNA polymerase II. *Genes Dev.*, **11**, 3306-3318.

McCracken, S., Fong, N., Yankulov, K., Ballantyne, S., Pan, G., Greenblatt, J., Patterson, S.D., Wickens, M. and Bentley, D.L. (1997b) The C-terminal domain of RNA polymerase II couples mRNA processing to transcription. *Nature*, **385**, 357-361.

Meshorer, E., Erb, C., Gazit, R., Pavlovsky, L., Kaufer, D., Friedman, A., Glick, D., Ben-Arie, N. and Soreq, H. (2002) Alternative splicing and neuritic mRNA translocation under long-term neuronal hypersensitivity. *Science*, **295**, 508-512.

Minvielle-Sebastia, L. and Keller, W. (1999) mRNA polyadenylation and its coupling to other RNA processing reactions and to transcription. *Curr. Opin. Cell Biol.*, **11**, 352-357.

Mougin, A., Gottschalk, A., Fabrizio, P., Luhrmann, R. and Branlant, C. (2002) Direct probing of RNA structure and RNA-protein interactions in purified HeLa cell's and yeast spliceosomal U4/U6.U5 tri-snRNP particles. *J. Mol. Biol.*, **317**, 631-649.

Myers, L.C., Lacomis, L., Erdjument-Bromage, H. and Tempst, P. (2002) The yeast capping enzyme represses RNA polymerase II transcription. *Mol. Cell*, **10**, 883-894.

Niwa, M. and Berget, S.M. (1991) Mutation of the AAUAAA polyadenylation signal depresses in vitro splicing of proximal but not distal introns. *Genes Dev.*, **5**, 2086-2095.

Niwa, M., Rose, S.D. and Berget, S.M. (1990) In vitro polyadenylation is stimulated by the presence of an upstream intron. *Genes Dev.*, **4**, 1552-1559.

Nobile, C., Marchi, J., Nigro, V., Roberts, R.G. and Danieli, G.A. (1997) Exon-intron organization of the human dystrophin gene. *Genomics*, **45**, 421-424.

Nottrott, S., Urlaub, H. and Luhrmann, R. (2002) Hierarchical, clustered protein interactions with U4/U6 snRNA: a biochemical role for U4/U6 proteins. *EMBO J.*, **21**, 5527-5538.

Ogg, S.C. and Lamond, A.I. (2002) Cajal bodies and coilin--moving towards function. *J. Cell Biol.*, **159**, 17-21.

Ohno, M., Segref, A., Bachi, A., Wilm, M. and Mattaj, I.W. (2000) PHAX, a mediator of U snRNA nuclear export whose activity is regulated by phosphorylation. *Cell*, **101**, 187-198.

Ortlepp, D., Laggerbauer, B., Mullner, S., Achsel, T., Kirschbaum, B. and Luhrmann, R. (1998) The mammalian homologue of Prp16p is overexpressed in a cell line tolerant to Leflunomide, a new immunoregulatory drug effective against rheumatoid arthritis. *RNA*, **4**, 1007-1018.

Ota, T., Suzuki, Y., Nishikawa, T., Otsuki, T., Sugiyama, T., Irie, R. *et al.* (2004) Complete sequencing and characterization of 21,243 full-length human cDNAs. *Nat. Genet.*, **36**, 40-45.

Otake, L.R., Scamborova, P., Hashimoto, C. and Steitz, J.A. (2002) The divergent U12-type spliceosome is required for pre-mRNA splicing and is essential for development in Drosophila. *Mol. Cell*, **9**, 439-446.

Padgett, R.A., Konarska, M.M., Grabowski, P.J., Hardy, S.F. and Sharp, P.A. (1984) Lariat RNA's as intermediates and products in the splicing of messenger RNA precursors. *Science*, **225**, 898-903.

Pannone, B.K. and Wolin, S.L. (2000) Sm-like proteins wRING the neck of mRNA. *Curr. Biol.*, **10**, R478-481.

Pannone, B.K., Xue, D. and Wolin, S.L. (1998) A role for the yeast La protein in U6 snRNP assembly: evidence that the La protein is a molecular chaperone for RNA polymerase III transcripts. *EMBO J.*, **17**, 7442-7453.

Patel, A.A., McCarthy, M. and Steitz, J.A. (2002) The splicing of U12-type introns can be a rate-limiting step in gene expression. *EMBO J.*, **21**, 3804-3815.

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

Patel, N.A., Chalfant, C.E., Watson, J.E., Wyatt, J.R., Dean, N.M., Eichler, D.C. and Cooper, D.R. (2001) Insulin regulates alternative splicing of protein kinase C beta II through a

phosphatidylinositol 3-kinase-dependent pathway involving the nuclear serine/arginine-rich splicing factor, SRp40, in skeletal muscle cells. *J. Biol. Chem.*, **276**, 22648-22654.

Patton, J.R., Habets, W., van Venrooij, W.J. and Pederson, T. (1989) U1 small nuclear ribonucleoprotein particle-specific proteins interact with the first and second stem-loops of U1 RNA, with the A protein binding directly to the RNA independently of the 70K and Sm proteins. *Mol. Cell. Biol.*, **9**, 3360-3368.

Polydorides, A.D., Okano, H.J., Yang, Y.Y., Stefani, G. and Darnell, R.B. (2000) A brain-enriched polypyrimidine tract-binding protein antagonizes the ability of Nova to regulate neuron-specific alternative splicing. *Proc. Natl. Acad. Sci. USA*, **97**, 6350-6355.

Prasad, J., Colwill, K., Pawson, T. and Manley, J.L. (1999) The protein kinase Clk/Sty directly modulates SR protein activity: both hyper- and hypophosphorylation inhibit splicing. *Mol. Cell. Biol.*, **19**, 6991-7000.

Preker, P.J. and Keller, W. (1998) The HAT helix, a repetitive motif implicated in RNA processing. *Trends Biochem. Sci.*, **23**, 15-16.

Rader, S.D. and Guthrie, C. (2002) A conserved Lsm-interaction motif in Prp24 required for efficient U4/U6 di-snRNP formation. *RNA*, **8**, 1378-1392.

Raghunathan, P.L. and Guthrie, C. (1998) A spliceosomal recycling factor that reanneals U4 and U6 small nuclear ribonucleoprotein particles. *Science*, **279**, 857-860.

Raghunathan, P.L. and Guthrie, C. (1998a) RNA unwinding in U4/U6 snRNPs requires ATP hydrolysis and the DEIH-box splicing factor Brr2. *Curr. Biol.*, **8**, 847-855.

Raker, V.A., Plessel, G. and Luhrmann, R. (1996) The snRNP core assembly pathway: identification of stable core protein heteromeric complexes and an snRNP subcore particle in vitro. *EMBO J.*, **15**, 2256-2269.

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. *J. Biol. Chem.*, **276**, 31142-31150.

Reed, R. and Palandjian, L. (1997) Spliceosome assembly. *In Eucaryotic mRNA processing, A. Krainer, ed. New York, Oxford University Press Inc., New York*, 103-129.

Rodrigues, J.P., Rode, M., Gatfield, D., Blencowe, B.J., Carmo-Fonseca, M. and Izaurralde, E. (2001) REF proteins mediate the export of spliced and unspliced mRNAs from the nucleus. *Proc. Natl. Acad. Sci. USA*, **98**, 1030-1035.

Rosonina, E. and Blencowe, B.J. (2002) Gene expression: the close coupling of transcription and splicing. *Curr. Biol.*, **12**, R319-321.

Rossi, F., Forne, T., Antoine, E., Tazi, J., Brunel, C. and Cathala, G. (1996) Involvement of U1 small nuclear ribonucleoproteins (snRNP) in 5' splice site-U1 snRNP interaction. *J. Biol. Chem.*, **271**, 23985-23991.

Roy, J., Kim, K., Maddock, J.R., Anthony, J.G. and Woolford, J.L., Jr. (1995) The final stages of spliceosome maturation require Spp2p that can interact with the DEAH box protein Prp2p and promote step 1 of splicing. *RNA*, **1**, 375-390.

Ruskin, B. and Green, M.R. (1985) An RNA processing activity that debranches RNA lariats. *Science*, **229**, 135-140.

Ruskin, B., Krainer, A.R., Maniatis, T. and Green, M.R. (1984) Excision of an intact intron as a novel lariat structure during pre-mRNA splicing in vitro. *Cell*, **38**, 317-331.

Ryan, D.E., Stevens, S.W. and Abelson, J. (2002) The 5' and 3' domains of yeast U6 snRNA: Lsm proteins facilitate binding of Prp24 protein to the U6 telestem region. *RNA*, **8**, 1011-1033.

Sanford, J.R. and Bruzik, J.P. (1999) Developmental regulation of SR protein phosphorylation and activity. *Genes Dev.*, **13**, 1513-1518.

Schaal, T.D. and Maniatis, T. (1999) Selection and characterization of pre-mRNA splicing

enhancers: identification of novel SR protein-specific enhancer sequences. *Mol. Cell. Biol.*, **19**, 1705-1719.

Schaffert, N., Ingelfinger, D., Hossbach, M., Achsel, T. and Luhrmann, R. (2003) RNAi knockdown of the 61K or 102K protein leads to U4/U6.U5 tri-snRNP dissociation and accumulation of U4/U6 and U5 in Cajal bodies. *Eukaryotic mRNA processing.* Cold Spring Harbour Laboratory, N.Y., Cold Spring Harbour, p. 84.

Schneider, C., Will, C.L., Makarova, O.V., Makarov, E.M. and Luhrmann, R. (2002) Human U4/U6.U5 and U4atac/U6atac.U5 tri-snRNPs exhibit similar protein compositions. *Mol. Cell. Biol.*, **22**, 3219-3229.

Schwer, B. and Guthrie, C. (1991) PRP16 is an RNA-dependent ATPase that interacts transiently with the spliceosome. *Nature*, **349**, 494-499.

Schwer, B. and Meszaros, T. (2000) RNA helicase dynamics in pre-mRNA splicing. *EMBO J.*, **19**, 6582-6591.

Seipelt, R.L., Zheng, B., Asuru, A. and Rymond, B.C. (1999) U1 snRNA is cleaved by RNase III and processed through an Sm site-dependent pathway. *Nucleic Acids Res.*, **27**, 587-595.

Sell, S.M., Reese, D. and Ossowski, V.M. (1994) Insulin-inducible changes in insulin receptor mRNA splice variants. *J. Biol. Chem.*, **269**, 30769-30772.

Shannon, K.W. and Guthrie, C. (1991) Suppressors of a U4 snRNA mutation define a novel U6 snRNP protein with RNA-binding motifs. *Genes Dev.*, **5**, 773-785.

Sharp, P.A. and Burge, C.B. (1997) Classification of introns: U2-type or U12-type. *Cell*, **91**, 875-879.

Sharp, P.A., Konarska, M.M., Grabowski, P.J., Lamond, A.I., Marciniak, R. and Seiler, S.R. (1987) Splicing of messenger RNA precursors. *Cold Spring Harbor Symp. Quant. Biol.*, Cold Spring Harbor N.Y., Vol. LII, p. 277–285.

Shatkin, A.J. and Manley, J.L. (2000) The ends of the affair: capping and polyadenylation. *Nat. Struct. Biol.*, **7**, 838-842.

Shuman, S. and Schwer, B. (1995) RNA capping enzyme and DNA ligase: a superfamily of covalent nucleotidyl transferases. *Mol. Microbiol.*, **17**, 405-410.

Silverman, E., Edwalds-Gilbert, G. and Lin, R.J. (2003) DExD/H-box proteins and their partners: helping RNA helicases unwind. *Gene*, **312**, 1-16.

Singh, G. and Lykke-Andersen, J. (2003) New insights into the formation of active nonsense-mediated decay complexes. *Trends Biochem. Sci.*, **28**, 464-466.

Singh, R. and Reddy, R. (1989) Gamma-monomethyl phosphate: a cap structure in spliceosomal U6 small nuclear RNA. *Proc. Natl. Acad. Sci. USA*, **86**, 8280-8283.

Smith, C.W. and Valcarcel, J. (2000) Alternative pre-mRNA splicing: the logic of combinatorial control. *Trends Biochem. Sci.*, **25**, 381-388.

Spafford, J.D., Spencer, A.N. and Gallin, W.J. (1998) A putative voltage-gated sodium channel alpha subunit (PpSCN1) from the hydrozoan jellyfish, Polyorchis penicillatus: structural comparisons and evolutionary considerations. *Biochem. Biophys. Res. Commun.*, **244**, 772-780.

Spirin, A.S. (1969) The second Sir Hans Krebs Lecture. Informosomes. *Eur. J. Biochem.*, **10**, 20-35.

Staley, J.P. and Guthrie, C. (1999) An RNA switch at the 5' splice site requires ATP and the DEAD box protein Prp28p. *Mol. Cell*, **3**, 55-64.

Stamm, S., Zhang, M.Q., Marr, T.G. and Helfman, D.M. (1994) A sequence compilation and comparison of exons that are alternatively spliced in neurons. *Nucleic Acids Res.*, **22**, 1515-1526.

Stevens, S.W., Ryan, D.E., Ge, H.Y., Moore, R.E., Young, M.K., Lee, T.D. and Abelson, J. (2002) Composition and functional characterization of the yeast spliceosomal penta-

snRNP. *Mol. Cell*, **9**, 31-44.
Stojdl, D.F. and Bell, J.C. (1999) SR protein kinases: the splice of life. *Biochem. Cell Biol.*, **77**, 293-298.
Strasser, K. and Hurt, E. (2001) Splicing factor Sub2p is required for nuclear mRNA export through its interaction with Yra1p. *Nature*, **413**, 648-652.
Stutz, F., Bachi, A., Doerks, T., Braun, I.C., Seraphin, B., Wilm, M., Bork, P. and Izaurralde, E. (2000) REF, an evolutionary conserved family of hnRNP-like proteins, interacts with TAP/Mex67p and participates in mRNA nuclear export. *RNA*, **6**, 638-650.
Sun, C. and Woolford, J.L., Jr. (1994) The yeast NOP4 gene product is an essential nucleolar protein required for pre-rRNA processing and accumulation of 60S ribosomal subunits. *EMBO J.*, **13**, 3127-3135.
Sun, C. and Woolford, J.L., Jr. (1997) The yeast nucleolar protein Nop4p contains four RNA recognition motifs necessary for ribosome biogenesis. *J. Biol. Chem.*, **272**, 25345-25352.
Tanner, N.K. and Linder, P. (2001) DExD/H box RNA helicases: from generic motors to specific dissociation functions. *Mol. Cell*, **8**, 251-262.
Tarn, W.Y. and Steitz, J.A. (1995) Modulation of 5' splice site choice in pre-messenger RNA by two distinct steps. *Proc. Natl. Acad. Sci. USA*, **92**, 2504-2508.
Tarn, W.Y. and Steitz, J.A. (1996) Highly diverged U4 and U6 small nuclear RNAs required for splicing rare AT-AC introns. *Science*, **273**, 1824-1832.
Tarn, W.Y. and Steitz, J.A. (1996a) A novel spliceosome containing U11, U12, and U5 snRNPs excises a minor class (AT-AC) intron in vitro. *Cell*, **84**, 801-811.
Tarn, W.Y. and Steitz, J.A. (1997) Pre-mRNA splicing: the discovery of a new spliceosome doubles the challenge. *Trends Biochem. Sci.*, **22**, 132-137.
Tazi, J., Forne, T., Jeanteur, P., Cathala, G. and Brunel, C. (1993) Mammalian U6 small nuclear RNA undergoes 3' end modifications within the spliceosome. *Mol. Cell. Biol.*, **13**, 1641-1650.
Terns, M.P., Lund, E. and Dahlberg, J.E. (1992) 3'-end-dependent formation of U6 small nuclear ribonucleoprotein particles in Xenopus laevis oocyte nuclei. *Mol. Cell. Biol.*, **12**, 3032-3040.
Thanaraj, T.A. and Clark, F. (2001) Human GC-AG alternative intron isoforms with weak donor sites show enhanced consensus at acceptor exon positions. *Nucleic Acids Res.*, **29**, 2581-2593.
Trippe, R., Sandrock, B. and Benecke, B.J. (1998) A highly specific terminal uridylyl transferase modifies the 3'-end of U6 small nuclear RNA. *Nucleic Acids Res.*, **26**, 3119-3126.
Tycowski, K.T., You, Z.H., Graham, P.J. and Steitz, J.A. (1998) Modification of U6 spliceosomal RNA is guided by other small RNAs. *Mol. Cell*, **2**, 629-638.
Valcarcel, J. and Green, M.R. (1996) The SR protein family: pleiotropic functions in pre-mRNA splicing. *Trends Biochem. Sci.*, **21**, 296-301.
van der Houven van Oordt, W., Diaz-Meco, M.T., Lozano, J., Krainer, A.R., Moscat, J. and Caceres, J.F. (2000) The MKK(3/6)-p38-signaling cascade alters the subcellular distribution of hnRNP A1 and modulates alternative splicing regulation. *J. Cell Biol.*, **149**, 307-316.
Vidal, V.P., Verdone, L., Mayes, A.E. and Beggs, J.D. (1999) Characterization of U6 snRNA-protein interactions. *RNA*, **5**, 1470-1481.
Vidaver, R.M., Fortner, D.M., Loos-Austin, L.S. and Brow, D.A. (1999) Multiple functions of Saccharomyces cerevisiae splicing protein Prp24 in U6 RNA structural rearrangements. *Genetics*, **153**, 1205-1218.
Wang, Z., Hoffmann, H.M. and Grabowski, P.J. (1995) Intrinsic U2AF binding is modulated by

exon enhancer signals in parallel with changes in splicing activity. *RNA*, **1**, 21-35.

Wassarman, K.M. and Steitz, J.A. (1992) The low-abundance U11 and U12 small nuclear ribonucleoproteins (snRNPs) interact to form a two-snRNP complex. *Mol. Cell. Biol.*, **12**, 1276-1285.

Weg-Remers, S., Ponta, H., Herrlich, P. and Konig, H. (2001) Regulation of alternative pre-mRNA splicing by the ERK MAP-kinase pathway. *EMBO J.*, **20**, 4194-4203.

Wersig, C. and Bindereif, A. (1990) Conserved domains of human U4 snRNA required for snRNP and spliceosome assembly. *Nucleic Acids Res.*, **18**, 6223-6229.

Will, C.L. and Luhrmann, R. (2001) Spliceosomal UsnRNP biogenesis, structure and function. *Curr. Opin. Cell Biol.*, **13**, 290-301.

Will, C.L., Schneider, C., Benecke, H., Urlaub, H., Hossbach, M. and Luhrmann, R. (2003) Characterization of human U11/U12 snRNP proteins potentially involved in U12-type intron bridging and the identification of a highly divirgent drosophyla U11 snRNA. *Eukaryotic mRNA processing.* Cold Spring Harbour Laboratory, N.Y., Cold Spring Harbour, p. 224.

Will, C.L., Schneider, C., Reed, R. and Luhrmann, R. (1999) Identification of both shared and distinct proteins in the major and minor spliceosomes. *Science*, **284**, 2003-2005.

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

Woerfel, G. and Bindereif, A. (2001) In vitro selection of exonic splicing enhancer sequences: identification of novel CD44 enhancers. *Nucleic Acids Res.*, **29**, 3204-3211.

Wolff, T. and Bindereif, A. (1992) Reconstituted mammalian U4/U6 snRNP complements splicing: a mutational analysis. *EMBO J.*, **11**, 345-359.

Wolff, T. and Bindereif, A. (1993) Conformational changes of U6 RNA during the spliceosome cycle: an intramolecular helix is essential both for initiating the U4-U6 interaction and for the first step of slicing. *Genes Dev.*, **7**, 1377-1389.

Wu, J.Y. and Maniatis, T. (1993) Specific interactions between proteins implicated in splice site selection and regulated alternative splicing. *Cell*, **75**, 1061-1070.

Wu, Q. and Krainer, A.R. (1998) Purine-rich enhancers function in the AT-AC pre-mRNA splicing pathway and do so independently of intact U1 snRNP. *RNA*, **4**, 1664-1673.

Xiao, S.H. and Manley, J.L. (1998) Phosphorylation-dephosphorylation differentially affects activities of splicing factor ASF/SF2. *EMBO J.*, **17**, 6359-6367.

Xie, J. and Black, D.L. (2001) A CaMK IV responsive RNA element mediates depolarization-induced alternative splicing of ion channels. *Nature*, **410**, 936-939.

Xie, J. and McCobb, D.P. (1998) Control of alternative splicing of potassium channels by stress hormones. *Science*, **280**, 443-446.

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.

Yang, D., Nakao, M., Shichijo, S., Sasatomi, T., Takasu, H., Matsumoto, H., Mori, K., Hayashi, A., Yamana, H., Shirouzu, K. and Itoh, K. (1999) Identification of a gene coding for a protein possessing shared tumor epitopes capable of inducing HLA-A24-restricted cytotoxic T lymphocytes in cancer patients. *Cancer Res.*, **59**, 4056-4063.

Yeakley, J.M., Tronchere, H., Olesen, J., Dyck, J.A., Wang, H.Y. and Fu, X.D. (1999) Phosphorylation regulates in vivo interaction and molecular targeting of serine/arginine-rich pre-mRNA splicing factors. *J. Cell Biol.*, **145**, 447-455.

Yu, Y.-T., Scharl, E.C., Smith, C.M. and Steitz, J.A. (1999) The growing world of small nuclear ribonucleoproteins. In Atkins, J.F. (ed.), *The RNA world*. Cold Spring Harbour Press, Cold Spring Harbour, N.Y., pp. 487-524.

Yu, Y.T., Shu, M.D. and Steitz, J.A. (1998) Modifications of U2 snRNA are required for snRNP assembly and pre-mRNA splicing. *EMBO J.*, **17**, 5783-5795.

Yuryev, A., Patturajan, M., Litingtung, Y., Joshi, R.V., Gentile, C., Gebara, M. and Corden, J.L. (1996) The C-terminal domain of the largest subunit of RNA polymerase II interacts with a novel set of serine/arginine-rich proteins. *Proc. Natl. Acad. Sci. USA*, **93**, 6975-6980.

Zahler, A.M., Lane, W.S., Stolk, J.A. and Roth, M.B. (1992) SR proteins: a conserved family of pre-mRNA splicing factors. *Genes Dev.*, **6**, 837-847.

Zahler, A.M. and Roth, M.B. (1995) Distinct functions of SR proteins in recruitment of U1 small nuclear ribonucleoprotein to alternative 5' splice sites. *Proc. Natl. Acad. Sci. USA*, **92**, 2642-2646.

Zhou, Z., Licklider, L.J., Gygi, S.P. and Reed, R. (2002) Comprehensive proteomic analysis of the human spliceosome. *Nature*, **419**, 182-185.

Zhou, Z., Luo, M.J., Straesser, K., Katahira, J., Hurt, E. and Reed, R. (2000) The protein Aly links pre-messenger-RNA splicing to nuclear export in metazoans. *Nature*, **407**, 401-405.

Zhou, Z. and Reed, R. (1998) Human homologs of yeast prp16 and prp17 reveal conservation of the mechanism for catalytic step II of pre-mRNA splicing. *EMBO J.*, **17**, 2095-2106.

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.

Acknowledgements

I would like to thank my supervisor Prof. Albrecht Bindereif for the great support and help he is always willing to provide.

I am grateful also to our lab crew: Silke Schreiner, Zsofia Palfi, Jingyi Hui, Jan Medenbach, Lee-Hsueh Hung, and Pingping Wang, as well as to the former colleagues Mathias Bell, Stephan Lücke, and Gertrud Woerfel, who were at hand every time I needed assistance or advice. I am obliged to the colleagues from the Institute of Biochemistry and the "Graduiertenkolleg", especially to Prof. Alfred Pingoud, and Jamilah Michel whose warm welcome and help I much appreciate.

I wish to thank our collaborators, Peter Stoilov and Stefan Stamm from the University of Erlangen, Stephanie Nottrott and Reinhard Lührmann from Max Planck Institute of Biophysical Chemistry, Göttingen, David Stanek and Karla Neugebauer from Max Planck Institute of Molecular Cell Biology and Genetics, Dresden, for providing antibodies, for sharing their experience, and for the helpful discussion.

Special thanks also to Prof. Rainer Renkawitz for reviewing this thesis.

List of publications

Part of this work is published in the following papers:

Damianov, A., Schreiner, S. and Bindereif, A. (2004) Recycling of the U12-type spliceosome requires p110, a component of the U6atac snRNP. *Mol. Cell. Biol.*, **24,** 1700-1708.

Bell, M., Schreiner, S., Damianov, A., Reddy, R. and Bindereif, A. (2002) p110, a novel human U6 snRNP protein and U4/U6 snRNP recycling factor. *EMBO J.*, **21**, 2724-2735.

Erklärung

Hiermit erkläre ich, daß ich die vorliegende Arbeit selbständig durchgeführt und keine anderen als die angegebenen Quellen und Hilfsmittel verwendet habe.

Gießen, den 18.03.2004.
Andrey Damianov