Univerzita Karlova v Praze

Přírodovědecká fakulta

Studijní program: Biologie

Studijní obor: Buněčná a vývojová biologie


Bc. Klára Klimešová

# Mapping of SART3 interactions with spliceosomal snRNPs

# Mapování interakcí SART3 se sestřihovými snRNP částicemi

Diplomová práce

Školitel: doc. Mgr. David Staněk, Ph.D.

Praha, 2015

Prohlašuji, že jsem závěrečnou práci zpracovala samostatně a že jsem uvedla všechny použité informační zdroje a literaturu. Tato práce ani její podstatná část nebyla předložena k získání jiného nebo stejného akademického titulu.

V Praze, 21. 4. 2015

Klára Klimešová

Chtěla bych zde poděkovat především svému školiteli Davidu Staňkovi za jeho podporu, cenné rady a za ochotu kdykoli pomoci. Dále bych chtěla poděkovat také všem ostatním členům Laboratoře biologie RNA, ÚMG AV ČR, za vytváření skvělé pracovní atmosféry.

# Abstract

The splicing of pre-mRNA transcripts is catalyzed by a huge and dynamic machinery called spliceosome. The spliceosomal complex consists of five small nuclear ribonucleoprotein (snRNP) particles and hundreds of non-snRNP proteins. Biogenesis of spliceosomal snRNPs is a multi-step process, the final steps of which take place in a specialized sub-nuclear compartment, the Cajal body. However, molecular details of snRNP targeting to the Cajal body remain mostly unclear. Our previous results revealed that SART3 protein is important for accumulation of U4, U5 and U6 snRNPs in Cajal bodies, but how SART3 binds snRNP particles is elusive. SART3 has been identified as a U6 snRNP interaction partner and U4/U6 di-snRNP assembly factor. Here, we show that SART3 interacts with U2 snRNP as well, and that it binds specifically immature U2 particles. Next, we provide evidence that SART3 associates with U2 snRNP via Sm proteins, which are components of the stable snRNP core and are present in four out of five major snRNPs (i.e. in U1, U2, U4 and U5). We propose that the interaction between SART3 and Sm proteins represents a general SART3-snRNP binding mechanism, how SART3 recognizes immature snRNPs and quality controls the snRNP assembly process in Cajal bodies.

## Keywords

pre-mRNA splicing, spliceosomal snRNP, SART3, TPR domain

# Abstrakt

Sestřih pre-mRNA je katalyzován obrovským a velmi dynamickým sestřihovým komplexem, který se skládá z pěti malých jaderných ribonukleoproteinových částic (označovaných jako snRNP) a více než stovky dalších proteinů. Biogeneze sestřihových snRNP částic je komplikovaný proces, jehož závěrečné kroky se odehrávají ve specializovaných jaderných útvarech, Cajalových tělíscích. Molekulární podstata cílení snRNP částic do Cajalových tělísek však zůstává nejasná. Naše předchozí výsledky odhalily, že protein SART3 je důležitý pro akumulaci U4, U5 a U6 snRNP v Cajalových tělíscích, není ale známo, jakým způsobem SART3 tyto sestřihové částice váže. SART3 byl původně identifikován jako interakční partner U6 snRNP a faktor napomáhající složení U4/U6 di-snRNP částice. V této práci nicméně ukazujeme, že SART3 interaguje také s U2 snRNP a že specificky váže nesložené U2 částice. Dále poskytujeme důkazy, že SART3 asociuje s U2 snRNP přes Sm proteiny, které tvoří stabilní jádro čtyř z pěti hlavních snRNP částic (tzn. U1, U2, U4 a U5). Na základě našich výsledků navrhujeme, že interakce mezi SART3 a Sm proteiny představuje obecný mechanismus, jak SART3 rozpoznává nekompletní snRNP částice a kontroluje tak jejich skládání v Cajalových tělíscích.

## Klíčová slova

pre-mRNA sestřih, sestřihové snRNP částice, SART3, TPR doména

# Contents

# Abbreviations

| | |
|---|---|
| APS | ammonium persulfate |
| ARS2 | arsenite resistance 2 |
| BSA | bovine serum albumin |
| CB | Cajal body |
| CBC | cap-binding complex |
| CBP | cap-binding protein |
| cDNA | complementary DNA |
| CRM1 | chromosome region maintenance 1 |
| CT | C-terminal |
| CypH | cyclophilin H |
| DAPI | 4',6-diamidino-2-phenylindole |
| DTT | dithiothreitol |
| E domain | glutamic-acid-rich domain |
| EDTA | ethylenediaminetetraacetic acid |
| EGTA | ethyleneglycoltetraacetic acid |
| FISH | fluorescence *in situ* hybridization |
| FL | full length |
| FRET | Förster resonance energy transfer |
| GFP | green fluorescent protein |
| HAT | half-a-TPR |
| HEK | human embryonic kidney |
| HIV | human immunodeficiency virus |
| IP | immunoprecipitation |
| LSm | Like-Sm |
| $m_3^{2,2,7}G$ | 2,2,7-trimethylguanosine |
| $m^7G$ | 7-methylguanosine |
| NC | negative control |
| NLS | nuclear localization signal |
| Oct4 | octamer-binding protein 4 |
| PBS | phosphate buffered saline |
| PHAX | phosphorylated adaptor for RNA export |
| pICln | chloride ion current inducer protein |
| PIPES | piperazine-N,N'-bis(2-ethanesulfonic acid) |

| | |
|---|---|
| PRMT5 | protein arginine N-methyltransferase 5 |
| Prp | pre-mRNA processing factor |
| qPCR | quantitative PCR |
| Ran | Ras-related nuclear protein |
| RNPS1 | RNA-binding protein with a serine-rich domain |
| RRM | RNA recognition motif |
| RT | reverse transcription |
| SART3 | squamous cell carcinoma antigen recognized by T-cells 3 |
| scaRNA | small Cajal body-specific RNA |
| SDS | sodium dodecyl sulfate |
| SEM | standard error of the mean |
| SF3 | splicing factor 3 |
| siRNA | small interfering RNA |
| SMN | survival of motor neurons |
| snoRNA | small nucleolar RNA |
| snRNA | small nuclear RNA |
| snRNP | small nuclear ribonucleoprotein |
| ss | splice site |
| SSC | saline-sodium citrate |
| Tat | transactivating factor |
| TBE | Tris/borate/EDTA |
| TEMED | tetramethylethylenediamine |
| TGS1 | trimethylguanosine synthase 1 |
| Tip110 | Tat-interacting protein of 110 kDa |
| TMG | trimethylguanosine |
| TPR | tetratricopeptide repeat |
| U2AF | U2 auxiliary factor |
| Usp15 | ubiquitin specific peptidase 15 |
| WB | Western blot |
| γ-m-$P_3$ | γ-monomethyl triphosphate |

# 1 Introduction

Human genome encodes over 20 000 protein-coding genes. Most of these genes contain intervening sequences, introns, in addition to expressed sequences, exons. There are usually multiple introns within one gene and they tend to be much longer than the coding exons. A typical human protein-coding gene thus consists of seven 120 bp-long exons interrupted by 10-fold longer introns (Lander et al, 2001). This results in a huge number of intronic sequences that have to be precisely removed from nascent transcripts.

The introns are excised during the process of pre-mRNA splicing, which is catalyzed by the complex and dynamic ribonucleoprotein machinery called the spliceosome. In addition to the obligatory constitutive splicing, many precursor mRNAs can undergo the alternative splicing to generate several different mRNA forms and subsequently different proteins. The accurate intron recognition and removal is therefore a crucial step in gene expression.

The spliceosome consists of five small nuclear ribonucleoproteins (snRNPs) and numerous non-snRNP proteins. Before joining the spliceosome, snRNPs must go through a complicated maturation pathway that takes place in distinct subcellular compartments. In each compartment, snRNPs complete several biogenesis steps and then proceed to the next destination. Generally, snRNP biogenesis occurs in locations distinct from sites of their function, preventing immature snRNPs to interfere with the splicing, and the whole maturation process is strictly regulated. At several stages, the quality of snRNPs is checked and only properly assembled particles are allowed to proceed in the maturation pathway. An impairment in the snRNP assembly process can cause a severe disease, such as spinal muscular atrophy or retinitis pigmentosa.

During each splicing reaction, snRNPs are remodeled and must be therefore regenerated before the next round of splicing. These recycling steps are highly similar or identical to the late phase of snRNP *de novo* synthesis and both pathways merge in nuclear subcompartments called Cajal bodies. The final snRNP maturation steps occurring in Cajal bodies require a special protein factor, SART3, which promotes and quality controls the assembly of a subset of snRNP particles. The aim of this thesis is to reveal the mechanism of SART3 interactions with immature spliceosomal snRNPs and to elucidate its role in snRNP assembly.

# 2 Literature Review

## 2.1 Spliceosomal snRNPs: composition and structure

Small nuclear ribonucleoprotein (snRNP) particles are key components of the spliceosome. They form the catalytic core of the whole complex and perform the splicing reaction. There are five different major snRNP particles: U1, U2, U4, U5 and U6. Each of them comprises of several (at least ten) proteins and a single molecule of U-rich small nuclear RNA called U1, U2, U4, U5 or U6 snRNA, respectively.

The common characteristic of U1, U2, U4 and U5 snRNAs, which are all transcribed by RNA Polymerase II, is a unique 2,2,7-trimethylguanosine ($m_3^{2,2,7}G$, TMG) cap at their 5' end (Reddy et al, 1974). They obtain the cap post-transcriptionally, during a cytoplasmic phase of their maturation. Aside from the cap, there are two other post-transcriptional snRNA modifications, pseudouridylation and 2'-O-methylation. Both of them are present in all snRNAs but are most frequent in the U2 snRNA (for an overview of all modified snRNA nucleotides see Karijolich & Yu, 2010). Interestingly, these modifications are always concentrated in 5' part of the snRNA molecule, leaving the 3' half unmodified.

Although these snRNAs differ considerably in the primary sequence as well as in the secondary structure, they share a sequence feature consisting of a single stranded region flanked by two hairpins. This region contains a short conserved motif of Pu-A-$U_{4-6}$-G-Pu (Pu is for purine), the so-called Sm site (Branlant et al, 1982). Here, seven Sm proteins, SmB/B', SmD1, SmD2, SmD3, SmE, SmF and SmG, forming a heteroheptameric ring structure together, can be bound. SmB and SmB' are alternative splicing products of a single gene (Chu & Elkon, 1991).

All Sm proteins are members of an evolutionarily conserved protein family and possess two highly conserved Sm motifs separated by a variable spacer (Hermann et al, 1995; Weber et al, 2010). These motifs are involved in protein-protein interactions within the Sm ring proteins (Hermann et al, 1995), as well as in the snRNA binding (Weber et al, 2010). Contrary to other Sm proteins, SmB/B', SmD1 and SmD3 have an RG-rich region at the C-terminus the arginines of which are symmetrically dimethylated (Brahms et al, 2001; Friesen et al, 2001; Meister et al, 2001). More recently, crystallography studies revealed a fascinating structure of the Sm ring bound to U1 snRNA (Pomeranz

Krummel et al, 2009; Weber et al, 2010) or U4 snRNA, respectively (Leung et al, 2011). The structures clearly show that snRNA goes through the central hole of the ring with each of the seven Sm site nucleotides contacting one Sm protein (Fig. 2.1a).

U6 snRNA differs from Sm-containing snRNPs in multiple aspects. First of all, U6 snRNA is transcribed by RNA Polymerase III. It then undergoes a different maturation pathway than the other snRNAs during which it obtains a γ-monomethyl triphosphate (γ-m-$P_3$) cap at the 5' end guanosine (Singh & Reddy, 1989). However, U6 snRNA is pseudouridylated and 2'-O-methylated same as the other snRNAs (Karijolich & Yu, 2010). Although U6 snRNA does not have an Sm site sequence, there is a stable hairpin structure in the 5' portion of the molecule that resembles the 3' hairpin of Sm site flanking motifs (Branlant et al, 1982).

Instead of an Sm ring, U6 snRNA can bind a very similar complex of seven Like-Sm (LSm) proteins called LSm2-8. These proteins belong to the Sm protein family and therefore also contain two Sm motifs with the linker in between (Achsel et al, 1999; Salgado-Garrido et al, 1999). Unlike Sm proteins, LSm2-8 are able to form the stable ring structure independently of the snRNA presence and the pre-formed complex is then bound to the very 3' end comprising of an oligo-U tract (Achsel et al, 1999). Recent crystallography data showed that the manner in which LSm heteromer binds to the snRNA is also different from the other snRNAs. Four uridines at the U6 snRNA 3' end are recognized by LSm4, LSm8, LSm2 and LSm3, respectively, in the central hole and a preceding guanosine is recognized outside the hole by LSm7, as shown in Fig. 2.1b (Zhou et al, 2014). This means that U6 snRNA does not go through the LSm ring. The comparison between both rings interacting with snRNAs is shown in Fig. 2.1c.

Aside from the U6-specific LSm2-8 ring, LSm proteins can also associate into another heteroheptameric complex called LSm1-7 which is located to cytoplasmic Processing bodies and play a role in mRNA degradation (Bouveret et al, 2000). The equilibrium between LSm2-8 and LSm1-7 is provided by LSm8 concentration and changes in LSm8 expression are followed by changes in LSm2-7 localization from the nucleus to the cytoplasm (Novotný et al, 2012). Interestingly, LSm4 can be symmetrically dimethylated on arginines at the C-terminus similarly to SmB/B', SmD1 and SmD3 proteins (Brahms et al, 2001). It is not clear whether this modification can somehow influence LSm4 association with LSm1-7 or LSm2-8 rings.

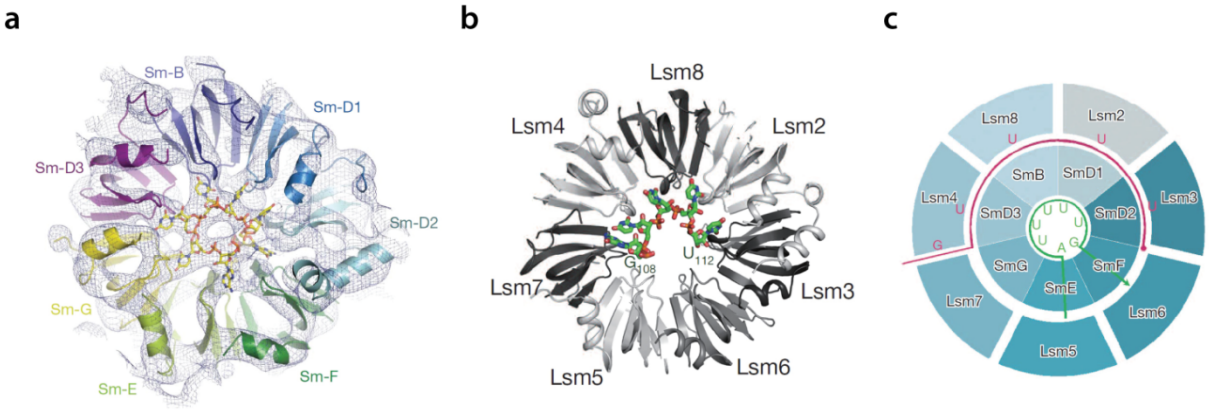


**Figure 2.1. Ring-like structures of Sm family proteins.**

(a) Crystal structure of the human Sm heptameric complex formed by SmB/D1/D2/D3/E/F/G proteins. Seven nucleotides of the U1 snRNA Sm site are in the middle, each of them contacting one Sm protein. Reprinted from Pomeranz Krummel et al, 2009.

(b) Crystal structure of the *Saccharomyces cerevisiae* LSm2-8 protein complex. Both Sm and LSm proteins exhibit the same Sm fold structure. Five nucleotides of the U6 snRNA 3'end are recognized by the central part of the LSm ring. Reprinted from Zhou et al, 2014.

(c) Comparison of snRNA recognition modes of Sm and LSm complexes. Reprinted from Zhou et al, 2014.

In addition to the snRNA molecule and Sm or LSm proteins, each snRNP particle contains a set of snRNP-specific proteins. A complete list of these proteins is represented in Fig. 2.2. U1 snRNP is the smallest of the five particles, since it comprises only three specific proteins. The major part of U1 snRNP has been crystallized so the structure of the functional core is well known (Pomeranz Krummel et al, 2009; Weber et al, 2010). Protein composition of U2 snRNP is more extensive. It contains U2A'/U2B" heterodimer, SF3a complex and SF3b complex. U2A'/U2B" bound to the 3' stem loop of U2 snRNA (Price et al, 1998) and SF3a complex (Lin & Xu, 2012) have been crystallized, while SF3b complex has been studied using electron cryomicroscopy (Golas et al, 2003), but the structure of complete U2 snRNP particle remains unsolved.

U4, U5 and U6, contrary to U1 and U2 snRNPs, have to be pre-formed into tri-snRNP particle prior to joining the spliceosome. During final steps of the maturation, U4 and U6 snRNAs base pair and form di-snRNP (Bringmann et al, 1984) which then associates with U5 snRNP through protein-protein interactions and give rise to the complete tri-snRNP particle (Black & Pinto, 1989). So far, a crystal structure is known only for 5' stem loop of U4 snRNA with hPrp31 and 15.5K proteins bound (Liu et al,

2007). However, data from electron cryomicroscopy showed the structures of separate U5 and U4/U6 snRNPs as well as the complete tri-snRNP (Sander et al, 2006).

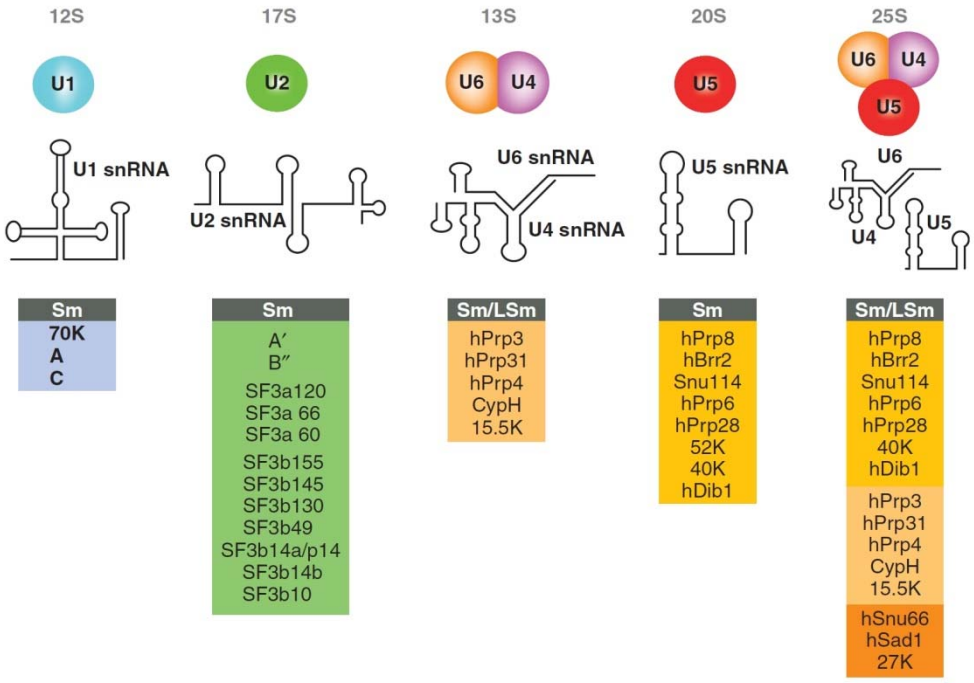


**Figure 2.2. Protein composition of individual snRNPs and secondary structure of their snRNAs.** Sm and LSm refers to Sm and LSm heptameric protein complexes. U4/U6·U5 snRNP contains two sets of Sm proteins and one set of LSm proteins. Reprinted from Will & Lührmann, 2011.

## 2.2 Early steps of snRNP biogenesis

U1, U2, U4 and U5 snRNPs, which are here referred to as Sm-class snRNPs, as well as U6 snRNP, referred to as LSm-class snRNP, undergo a complex multi-step maturation process before becoming fully functional. The mature particles function in the nucleoplasm, however, their biogenesis is restricted to the cytoplasm and several nuclear compartments including nucleolus and Cajal bodies (CBs), which makes the whole process quite complicated. It is believed that this arrangement provides an effective quality control mechanism and allows spatial isolation of unassembled snRNPs.

### 2.2.1 Sm-class snRNPs

Sm-class snRNAs are transcribed by RNA Polymerase II from special promoters. These promoters contain two characteristic features, proximal and distal sequence elements that resemble the TATA box and enhancer sequences, respectively, typical for protein-coding genes. 5' end of snRNA is capped probably co-transcriptionally like pre-mRNA (reviewed in Matera & Wang, 2014) . Surprisingly, nascent snRNAs associate with Cajal body scaffolding protein coilin and nucleate CBs in sites of snRNA transcription (Machyna et al, 2014).

Newly transcribed snRNAs have to be exported to the cytoplasm where next biogenesis steps occur (Fig. 2.3). The export is mediated by a multi-factor complex. First, 7-methylguanosine ($m^7G$) cap is recognized by cap-binding complex (CBC), a heterodimer of cap-binding protein 80 (CBP80) and CBP20 (Izaurralde et al, 1995). Then arsenite resistance 2 (ARS2) stimulates 3' end processing and binding of phosphorylated adaptor for RNA export (PHAX) to CBC (Hallais et al, 2013). PHAX must be hyperphosphorylated to function as an adaptor linking snRNA/CBC/ARS2 with chromosome region maintenance 1 (CRM1, also known as exportin 1) and RanGTP (Kitao et al, 2008; Ohno et al, 2000; Segref et al, 2001). CRM1 is an export receptor and mediates transport of the whole complex to the cytoplasm (Fornerod et al, 1997).

In the cytoplasm, PHAX is dephosphorylated, which causes disassembly of the export complex (Fig. 2.3), and then is recycled back to the nucleus (Ohno et al, 2000; Segref et al, 2001). The entire cytoplasmic phase of snRNP maturation including loading of Sm proteins, cap hypermethylation and 3' end trimming is regulated by the survival of motor neuron (SMN) complex (Massenet et al, 2002). The SMN complex consists of nine different proteins, SMN, Gemin2-8 and unrip. Gemin7, Gemin8 and SMN are located in the centre of the complex, while Gemin5 is only weakly bound on the periphery (Otter et al, 2007).

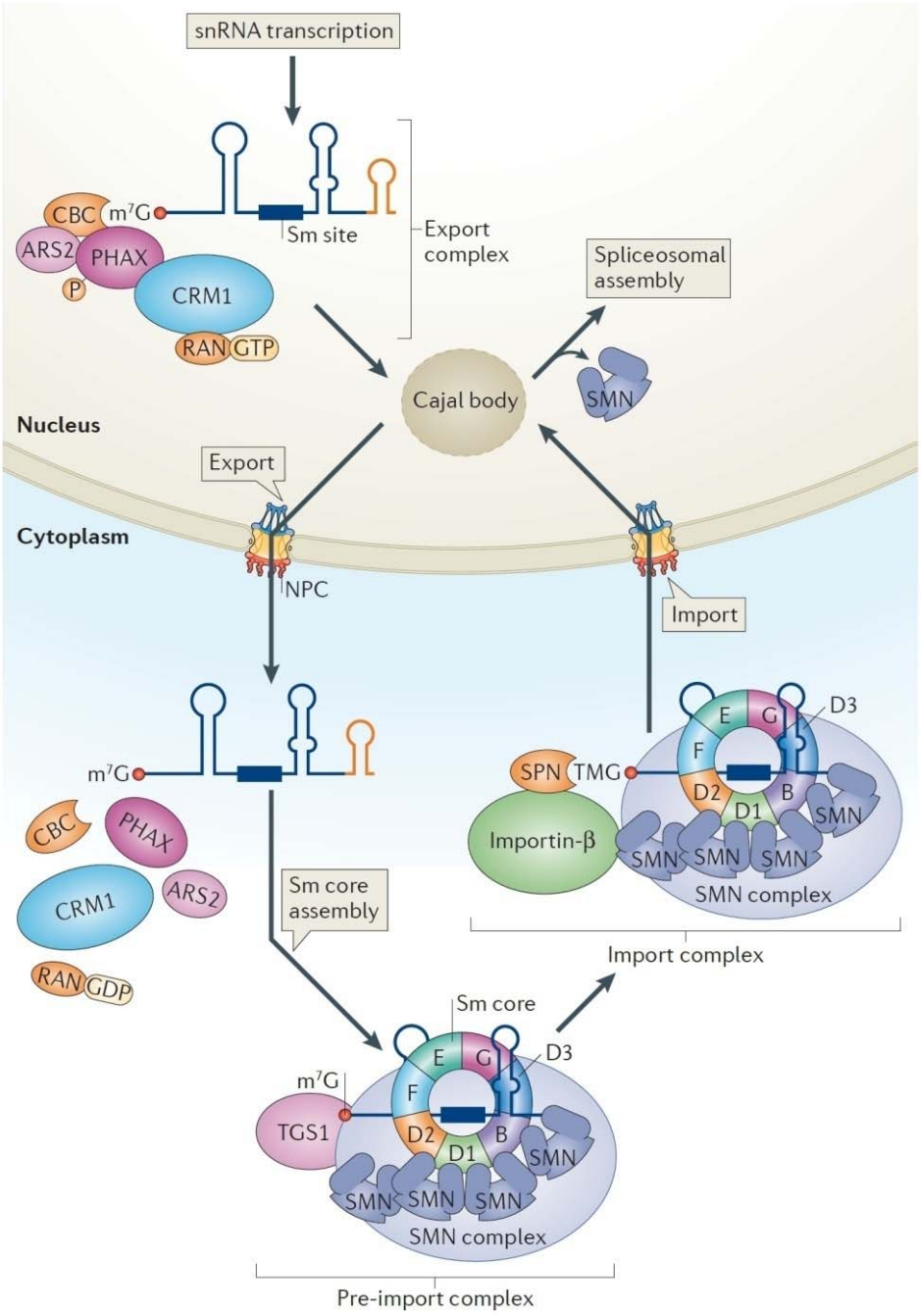


**Figure 2.3. Maturation of Sm-class snRNPs requires multiple cytoplasmic steps.**

Newly transcribed snRNA is capped and recognized by the export complex. In the cytoplasm, export factors dissociate and snRNA is bound by the SMN complex that promotes all the cytoplasmic maturation steps. The cap is hypermethylated, the 3' end trimmed and Sm proteins loaded around the Sm site. The import complex is then formed and the core particle transported back to the nucleus where it is targeted directly to Cajal bodies. Reprinted from Matera & Wang, 2014.

Sm proteins cannot be loaded on the snRNA as a complete ring, instead they pre-form into two heterodimers, B/D3 and D1/D2 and one heterotrimer, E/F/G (Raker et al, 1996). These subcomplexes are bound by protein arginine N-methyltransferase 5 (PRMT5) which symmetrically dimethylates C-terminal domains of SmB/B', SmD1 and SmD3 (Friesen et al, 2001; Meister et al, 2001). Then, a pentamer of D1/D2/E/F/G proteins interacts with pICln chaperone (Fig. 2.4) that mimics the Sm fold structure and therefore allows formation of a stable ring intermediate (Grimm et al, 2013). The pentamer is delivered to the SMN complex, bound by its Gemin2 protein (Grimm et al, 2013) and pICln is released. It possibly happens by exchanging pICln with Tudor domain of SMN that resembles Sm structure and can also interact with dimethylated C-termini of Sm proteins (Selenko et al, 2001). Gemin5 meanwhile associates with the snRNA and brings it to the SMN complex (Yong et al, 2010). The Sm ring is then enclosed by B/D3 dimer around the Sm site (Raker et al, 1996), as shown in Fig. 2.4. The SMN complex thus mediates the Sm ring formation and simultaneously ensures that right RNA substrates will be used (Pellizzoni et al, 2002).

Sm ring assembly stabilizes the snRNA and initiates further maturation steps. At 3' ends of snRNAs there are extra sequences that facilitate the snRNP core assembly but are not needed afterwards and may even interfere with snRNA function. Hence, these ends are exonucleolytically trimmed (Yong et al, 2010). The SMN complex remains bound to the core particle and recruits trimethylguanosine synthase 1 (TGS1) that hypermethylates 5' cap (Mouaikel et al, 2003). Sm ring and TMG cap both serve as nuclear localization signals. TMG cap is recognized by snurportin 1 which directly interacts with an import receptor importin β (Huber et al, 2002; Narayanan et al, 2002), the resulting import complex is depicted in Fig. 2.3.

The SMN complex is thought to be imported to the nucleus together with snRNPs (Narayanan et al, 2004) and to dissociate after snRNPs have reached their next destination, Cajal bodies. Interestingly, nuclear SMN complex slightly differs from the cytoplasmic one and lacks Gemin5 and unrip proteins (Grimmler et al, 2005; Hao le et al, 2007; Meister et al, 2000). The SMN complex released from the import complex is accumulated in special nuclear bodies. They are termed Gemini of Cajal bodies, or gems, because they are often found associated with CBs (Liu & Dreyfuss, 1996).

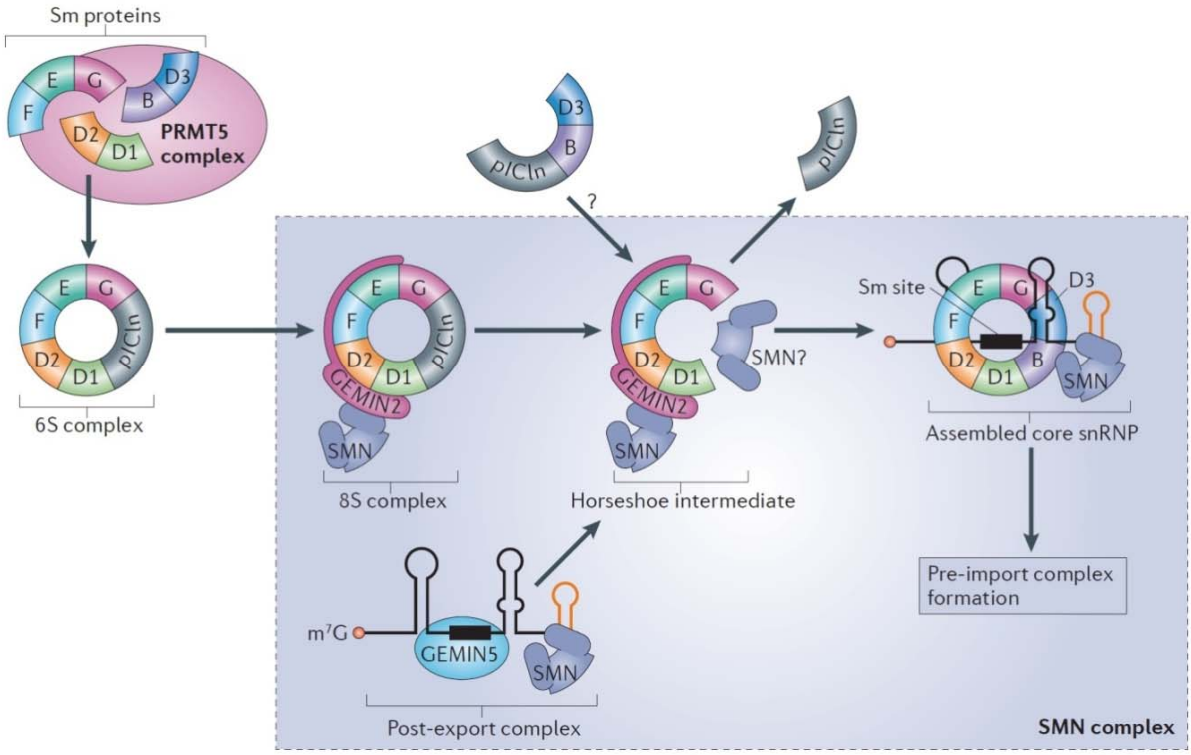


**Figure 2.4. Assembly of the Sm ring.**

Pre-formed Sm subcomplexes are methylated by the PRMT5 complex. 6S complex composed of the Sm pentamer and pICln chaperone is recruited to the SMN complex, which simultaneously binds the snRNA molecule and facilitates the Sm ring formation. pICln is released, SMN stabilizes the horseshoe intermediate and SmB/D3 closes the ring around the snRNA Sm site. Reprinted from Matera & Wang, 2014.

### 2.2.2 LSm-class snRNPs

Similarly to Sm-containing snRNPs, biogenesis of U6 snRNP is restricted from the nucleoplasm to prevent immature particles to interfere with the splicing apparatus. However, early maturation steps occur in the nucleolus instead of the cytoplasm. U6 promoters include, except for proximal and distal sequence elements, also a TATA box and the genes are transcribed by RNA Polymerase III. Like other RNA Polymerase III transcripts, U6 snRNA has a 3' end polyU sequence that serves as a transcription termination signal (reviewed in Patel & Bellini, 2008).

The 3' end of nascent snRNA is bound by La protein (Licht et al, 2008) and the complex is transported to the nucleolus. Specific enzymes then both extend (Trippe et al,

2006) and trim (Booth & Pugh, 1997) the 3' oligoU tail to the mature length. The terminal uridine is cleaved to generate a cyclic phosphate (Gu et al, 1997) that prevents the La protein from binding and enhances the affinity for LSm2-8 ring (Licht et al, 2008). Since LSm proteins are capable of forming the ring structure independently on snRNA (Achsel et al, 1999) it is likely that the LSm ring is loaded on U6 snRNA in a single-step process.

Aside from these modifications, U6 snRNA has to be methylated at the 5' end to obtain the γ-m-$P_3$ cap (Shimba & Reddy, 1994) and pseudouridylated and 2'-O-methylated which happens in reactions guided by small nucleolar RNAs (snoRNAs) (Ganot et al, 1999). After completing all the modifications, U6 snRNA can be released from the nucleolus and targeted to the Cajal body.

## 2.3 Late steps of snRNP biogenesis in Cajal bodies

### 2.3.1 Cajal bodies and scaRNAs

The final steps of snRNP maturation take place in small subnuclear compartments, Cajal bodies (CBs). They were firstly discovered by Spanish neurobiologist Santiago Ramón y Cajal in 1903, but then almost hundred years of rediscoveries and name changes followed until they were finally termed after Cajal (Gall, 2003). The main component of CBs is the scaffolding protein coilin, which is usually used as their marker (Raška et al, 1991); however, coilin is found freely diffusing in the nucleoplasm as well (Carmo-Fonseca et al, 1993).

Despite their important role in the spliceosomal snRNP generation, CBs are present only in a subset of cells with high proliferation or metabolic rate (e.g. neurons or cancer cells). Number of CBs in these cells is very variable, but there are usually around three or four bodies found per a nucleus. They are round and measure around 0.5 – 1 μm in diameter. Cajal bodies are highly dynamic structures with all the components, including coilin, constantly entering and leaving the bodies, (Dundr et al, 2004).

CB integrity depends on numerous factors, among others also on the cell cycle (Carmo-Fonseca et al, 1993; Ferreira et al, 1994). Their disappearance during the mitosis might be a consequence of stalled transcription. Several lines of evidence show that

ongoing snRNA transcription and snRNP biogenesis are the main determinants of CB nucleation and integrity (Girard et al, 2006; Kaiser et al, 2008; Lemm et al, 2006; Machyna et al, 2014; Novotný et al, 2015). CBs closely associates with SMN-containing bodies, gems, in some cell types. This association depends on coilin methylation status which modulates coilin ability to interact with SMN (Hebert et al, 2002). Both structures, though, seem to be functionally independent (Dundr et al, 2004; Lemm et al, 2006).

Having been imported from the cytoplasm, newly synthesized core snRNPs accumulate immediately in Cajal bodies (Sleeman & Lamond, 1999). Here, conserved Sm motifs of their Sm rings can interact directly with coilin (Xu et al, 2005). Maturation steps that snRNPs have to undergo in CBs can be generally divided to two events, snRNA modifications and loading of snRNP-specific proteins.

Pseudouridylation and 2'-O-methylation are provided by a special class of snoRNAs called small Cajal body-specific RNAs (scaRNAs). Contrary to classical snoRNAs, scaRNAs are localized exclusively to CBs and contain both the box C/D and box H/ACA motifs, so one molecule can direct both 2'-O-methylation and pseudouridylation (Darzacq et al, 2002; Jády et al, 2003). Importantly, recent study showed that snoRNAs traffic through CBs during their biogenesis similarly as snRNAs. Since scaRNAs contain a special CB-targeting element, CAB box, they are retained in CBs while other snoRNAs continue to nucleoli (Machyna et al, 2014).

Post-transcriptional modifications significantly increase the stability of snRNAs and might be needed for their proper function in the spliceosome (Karijolich & Yu, 2010). In case of U2 snRNP, modifications are also important for binding of U2-specific SF3 proteins and thus for the whole particle assembly (Yu et al, 1998). Modified snRNPs are captured in CBs until they are fully assembled and then are released and stored in nuclear speckles (Sleeman & Lamond, 1999). However, there is an exception; it seems that U1 snRNP undergoes a different maturation pathway and contrary to other snRNPs is assembled outside the CB (Matera & Ward, 1993; Ospina et al, 2005; Xu et al, 2005). Instead, one of the U1-specific proteins, U1-70K, was found in gems (Stejskalová & Staněk, 2014). But any details of final steps in U1 snRNP biogenesis still remain to be discovered.

### 2.3.2 U2 snRNP assembly

Assembly of the fully mature 17S U2 snRNP occurs in a step-wise process during which two intermediate particles of 12S and 15S, respectively, are generated. First, U2 snRNP core has to associate with U2A'/ U2B" heterodimer. U2A' and U2B" proteins interact already in the cytoplasm and are imported to the nucleus together, as a stable dimer, and independently on U2 snRNA (Kambach & Mattaj, 1994). Probably in Cajal bodies, U2A'/U2B" complex binds to 3' stem loop IV of U2 snRNA (Fig. 2.5a, b), both of them directly contacting the snRNA and forming the 12S particle (Boelens et al, 1991; Price et al, 1998). These intermediate snRNPs are efficiently retained within CBs (Nesic et al, 2004; Tanackovic & Krämer, 2005), though the mechanism is not known.

Next, SF3b complex consisting of eight proteins, SF3b155/145/130/125/49/14a/14b/10, is incorporated. The interaction with U2 snRNA is mediated mainly by SF3b49 protein that binds the 5' end and stem loops I and IIb (Fig. 2.5c) (Dybkov et al, 2006; Krämer et al, 1999). SF3a complex is composed of only three proteins, SF3a120/66/60, which, similarly as U2A'/U2B" and maybe also SF3b proteins, interact in the cytoplasm and are imported as a pre-formed complex (Huang et al, 2011; Nesic et al, 2004). SF3a incorporation into 15S particle is provided predominantly by SF3a60 protein that contacts nucleotides of U2 snRNA stem loops I and III (Fig. 2.5d) (Dybkov et al, 2006; Krämer et al, 1999). Interestingly, SF3b125 dissociates during 17S snRNP formation so the mature particle contains just seven SF3b proteins. This suggests a possible role of SF3b125 in facilitating 17S particle assembly (Will et al, 2002). Immediately after completing the maturation, U2 snRNPs are released from CBs and targeted to nuclear speckles (Nesic et al, 2004).

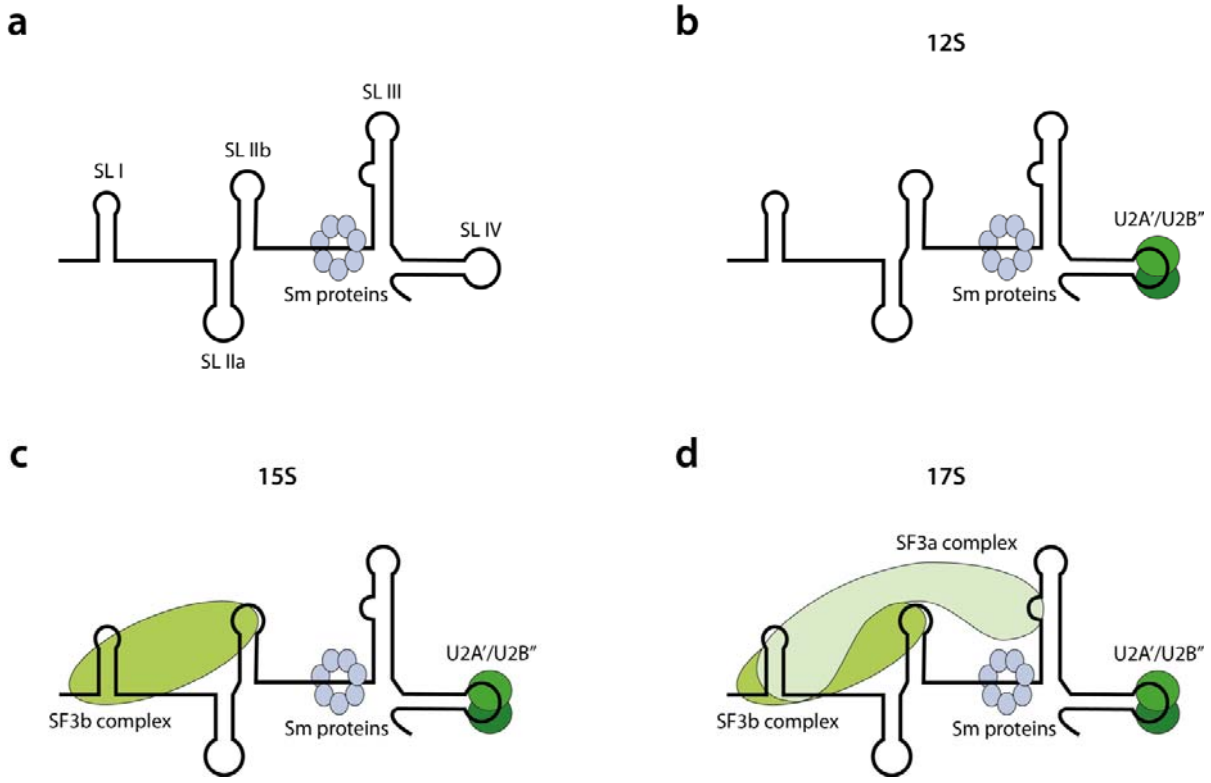


**Figure 2.5. Assembly of U2 snRNP.**

(a) Secondary structure of core U2 snRNP as transported to Cajal bodies. SL is for stem loop.

(b) U2A'/U2B" heterodimer binds to stem loop IV of U2 snRNA, forming the 12S particle.

(c) Proteins of SF3b complex interact with stem loops I and IIb and generate the 15S intermediate.

(d) Finally, SF3a complex binds to stem loops I and III and the mature 17S particle is assembled.

Based on information from Dybkov et al, 2006; Krämer et al, 1999 and Price et al, 1998.

### 2.3.3 Tri-snRNP assembly and role of SART3

Squamous cell carcinoma antigen recognized by T-cells 3 (SART3, also referred to as p110, $p110^{nrb}$ or Tip110) is a U6 and U4/U6 snRNP-specific protein that promotes di-snRNP assembly in Cajal bodies. SART3 localization in the cell is exclusively nuclear, with accumulation in CBs. Although it is expressed in most human tissues, SART3 expression is significantly increased in highly proliferative tissues such as cancer tissues, testis, hematopoietic cells or embryonic stem cells (Gu et al, 1998; Liu et al, 2013; Yang et al, 1999).

SART3 has been implicated in numerous distinct processes related to gene expression. Among others, it activates HIV-1 virus production via upregulation of HIV

transactivation factor Tat (Liu et al, 2002). SART3 also interacts with pre-mRNA splicing factor termed RNA-binding protein with a serine-rich domain (RNPS1) and in cooperation with RNPS1, SART3 was able to modulate alternative splicing of examined pre-mRNA reporter (Harada et al, 2001). Furthermore, SART3 also regulates alternative splicing of Oct4 factor in human embryonic stem cells and thus maintains their pluripotency. Expression of SART3 is regulated by c-Myc factor in this pathway (Liu et al, 2013). Recently, SART3 was suggested to have histone chaperone activities, because it recruits free ubiquitinated histones H2B to a deubiquitinating enzyme Usp15 (Long et al, 2014). This brief overview illustrates the complexity of SART3 functions within the cell. However, its full significance and possible connections between all these functions remain to be elucidated.

Human SART3 is characterized by seven half-a-tetratricopeptide repeat (TPR) motifs occupying the N-terminal two-thirds of the protein and by two RNA recognition motifs (RRM) carried by the last third (Fig. 2.6a) (Bell et al, 2002; Gu et al, 1998). The domain composition is generally conserved in eukaryotes, although number of individual motif repeats is variable. The exception is SART3-related protein Prp24 present in *Saccharomyces cerevisiae*, which contains four RRMs but entirely lacks TPR motifs (Bell et al, 2002; Rader & Guthrie, 2002). Therefore, functions provided by the C-terminal part of the protein are thought to be conserved between human and yeast, whereas TPR motifs are believed to reside in a separate protein in *S. cerevisiae*.

The RRM domain of SART3 directly binds several nucleotides in the internal region of mature U6 snRNA (Bell et al, 2002; Medenbach et al, 2004) while the C-terminal (CT) domain consisting of ten highly conserved amino acids interacts with the LSm ring (Licht et al, 2008; Rader & Guthrie, 2002; Staněk et al, 2003). Electron microscopy study revealed that the CT domain of Prp24 and possibly that of SART3 as well, is positioned on the flat side of the ring and probably contacts most or all the LSm proteins (Karaduman et al, 2008). Besides the LSm ring, SART3 is the only U6-specific protein (Fig. 2.6b). The interaction between U6 and SART3 occurs in the nucleoplasm and SART3 then targets U6 snRNP to the Cajal bodies where U4/U6 and U4/U6·U5 snRNPs are assembled (Schaffert et al, 2004; Staněk & Neugebauer, 2004; Staněk et al, 2003).

In Cajal bodies, SART3 facilitates the association of U6 with U4 snRNP, which contains Sm ring and 15.5K protein bound directly to the snRNA (Nottrott et al, 1999).

RRM domain has a suitable conformation that allows U4 snRNA to base pair with U6 snRNA, which is still bound by SART3 and promotes U4/U6 duplex formation, as showed on Prp24 crystal structure (Montemayor et al, 2014). U4/U6 snRNP then interacts with additional di-snRNP-specific proteins, hPrp31 (61K) and hPrp3/hPrp4/CypH (90K/60K/20K) heterotrimer (Nottrott et al, 2002). To further enhance di-snRNP formation and stabilize the particle, SART3 binds through its TPR domain a C-terminal region of hPrp3 protein (Medenbach et al, 2004).

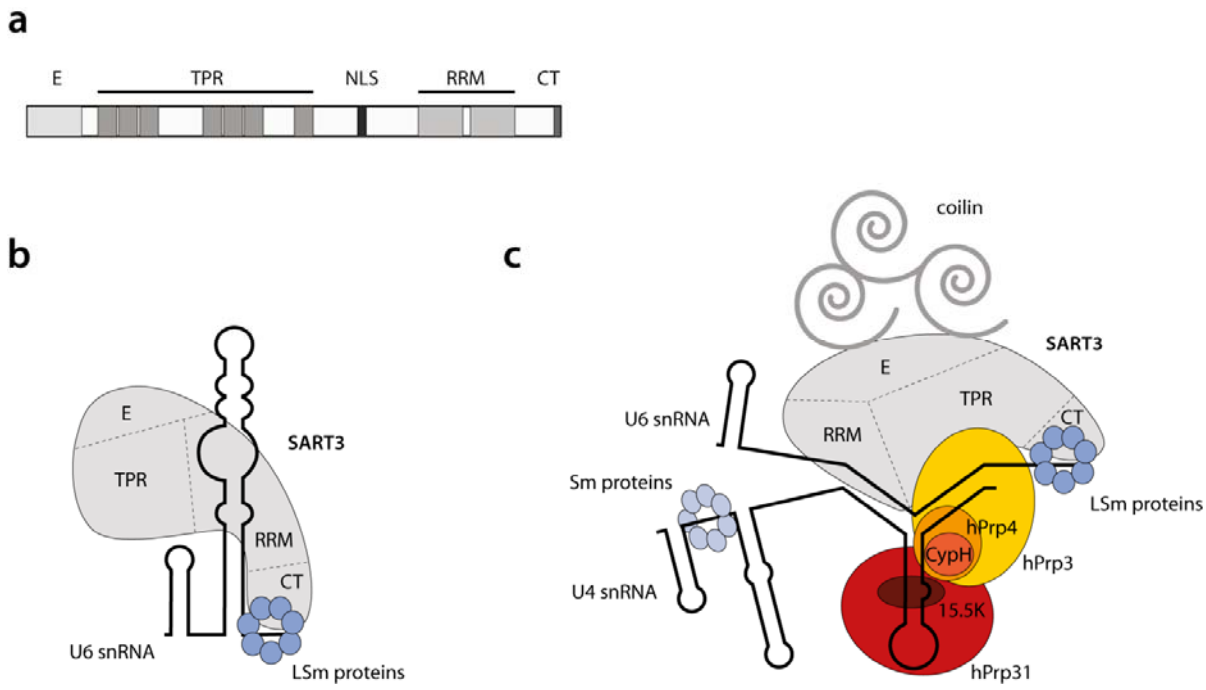


**Figure 2.6. SART3 and its interactions with U6 and U4/U6 snRNPs.**

(a) Conserved motifs of SART3 protein. E, glutamic-acid-rich domain; TPR, tetratricopeptide repeat domain; NLS, nuclear localization signal; RRM, RNA recognition motifs; CT, C-terminal domain.

(b) U6 snRNP consists of U6 snRNA, LSm ring and SART3 protein.

(c) Fully assembled di-snRNP particle. The assembly is promoted by SART3, which also anchors U4/U6 snRNP within the Cajal body.

Based on information from Bell et al, 2002; Liu et al, 2006; Medenbach et al, 2004; Novotný et al, 2015 and Rader & Guthrie, 2002.

U4/U6 di-snRNPs present immature intermediate particles that must be converted to tri-snRNPs before the release to the nucleoplasm where the splicing occurs. They are therefore efficiently sequestered in CBs (Schaffert et al, 2004; Staněk & Neugebauer,

2004). This function is also provided by SART3, namely by its N-terminal glutamic-acid-rich region, E domain, which directly interacts with CB scaffolding protein coilin (Novotný et al, 2011; Novotný et al, 2015). In addition to di-snRNP chaperone function, SART3 thus serves also as an anchor holding immature particles inside CBs and away from spliceosomes. Complete CB-retained U4/U6 particle is schematically depicted in Fig. 2.6c.

U5 snRNP is incorporated into the tri-snRNP particle only through protein-protein interactions, neither U4 nor U6 snRNAs base pair with U5 snRNA. U5 is the largest spliceosomal snRNP, it consists of an Sm ring, snRNA and eight specific proteins that together form the 20S particle (Fig. 2.2). The central component of the U5 particle is hPrp8 protein which is necessary for other specific proteins to associate and generate mature snRNP (Liu et al, 2006; Novotný et al, 2015). U5 snRNP is attached to U4/U6 snRNP via its hPrp6 protein that contacts di-snRNP-specific hPrp31 and hPrp3 (Liu et al, 2006; Schaffert et al, 2004). This reaction is probably promoted by the U5-specific 52K protein, the only protein released from the particle prior to the tri-snRNP formation. It seems that 52K functions as a chaperone for hPrp6 and protects it from unspecific interactions (Laggerbauer et al, 2005). U4/U6·U5 snRNP is then stabilized by integration of three additional tri-snRNP-specific proteins (Fig. 2.2) (Liu et al, 2006).

During the tri-snRNP generation, SART3 is evicted from the particle, so the CB-retaining anchor is releaved, and the mature U4/U6·U5 snRNP is immediately released to the nucleoplasm. Strikingly, hPrp6 protein contains several half-a-TPR and TPR motifs that interact with hPrp3 (Liu et al, 2006). The short C-terminal region of hPrp3 bound by hPrp6 perfectly overlaps with that bound by SART3 TPR domain (Liu et al, 2006; Medenbach et al, 2004), raising the interesting possibility of a system of hPrp6 and SART3 competing for hPrp3 binding site. SART3 would be displaced by hPrp6 in this model, resulting in the tri-snRNP detachment.

As already mentioned, Cajal bodies are detectable only in a subset of cells, however, coilin and SART3 are expressed even in the cells that lack CBs. Di- and tri-snRNP assembly in these cells occurs normally, but it takes place in the nucleoplasm (Staněk & Neugebauer, 2004). Interestingly, inhibition of the snRNP assembly pathway causing increased concentration of intermediate particles can trigger nucleation of new CBs (Novotný et al, 2015). This suggests that the importance of Cajal bodies does not lay in promoting the reaction itself but rather in enhancing its efficiency and kinetics. In fact,

snRNA concentration in CBs is approximately 20-fold higher than in the surrounding nucleoplasm and the presence of four CBs in the nucleus can increase tri-snRNP assembly rate 10-fold (Klingauf et al, 2006; Novotný et al, 2011). That is probably why the rapidly proliferating cells take advantage of Cajal bodies while the others do not need them.

## 2.4 Pre-mRNA splicing

Splicing of nascent mRNAs is driven by snRNAs that base pair with short conserved motifs within pre-mRNA molecules. These motifs are 5' splice site (5'ss) and 3' splice site (3'ss) located at the beginning and end, respectively, of an intron; and branch point which is usually located 15 – 50 nucleotides upstream of the 3'ss. Branch point is represented by a single adenosine and there is usually a pyrimidine-rich region between the branch point and the 3'ss. Despite their obvious importance, splice site sequences are very short and much less conserved than one would expect. The mechanism of proper intron recognition is thus still under the discussion.

The intron removal from pre-mRNA happens as a two-step event. First, branch point adenosine performs via its 2' OH group a nucleophilic attack on the 5'ss, causing a lariat structure formation. Then, 3'ss is attacked by the free 3' OH group of the upstream exon and it results in ligation of both exons and release of lariat intron (Fig. 2.7a). The splicing reaction is believed to be catalyzed by snRNAs. However, extensive remodeling of the whole complex is required to generate activated spliceosome with snRNAs in suitable conformations for the splicing catalysis. These rearrangements are promoted by eight spliceosomal RNA helicases that belong to DExD/H-box family (reviewed in Cordin & Beggs, 2013).

### 2.4.1 Role and remodeling of snRNPs within the spliceosome

The spliceosome is huge and dynamic RNP machinery. For each splicing reaction, it is assembled *de novo* on the pre-mRNA substrate, in most cases co-transcriptionally. Active spliceosomes are hence located in regions of decompacted chromatin, usually at the periphery of nuclear speckles in which mature snRNPs are

stored (Girard et al, 2012). As the first step, 5'ss is recognized by the U1 snRNP, snRNA of which base pairs with the pre-mRNA, and then U2 auxiliary factors (U2AF) 65 and 35 recognize the branch point and the 3'ss. U1 snRNP and the weakly bound U2 snRNP together form the spliceosomal complex E, which surprisingly contains also SMN complex (Makarov et al, 2012; Pellizzoni et al, 1998).

In the next step, U2 snRNP association with the complex is stabilized. U2-specific SF3b14a contacts the branch point (Will et al, 2001) and U2 snRNA interacts with the U1 snRNP and the region around the branch point, causing the branch adenosine to bulge out. Resulting complex includes another 50 non-snRNP proteins but not the SMN complex, and is termed the complex A (Behzadnia et al, 2007). Afterwards, U4/U6·U5 tri-snRNP is integrated, forming the pre-catalytic complex B (Fig. 2.7b) which is, with its more than 110 proteins, the largest spliceosomal complex (Deckert et al, 2006).

However, complex B is still not catalytically active and must undergo compositional and conformational remodeling driven by multiple RNA helicases. U4 and U6 snRNAs are unwound during the process and U6 base pairs with U2 snRNA, while U1 and U4 snRNPs together with all the di- and tri-snRNP-specific proteins are released (Makarov et al, 2002). In activated complex B*, branch point and 5'ss are in close proximity due to U2/U6 annealing and the complex completes the first splicing reaction (reviewed in more details in Will & Lührmann, 2011), generating complex C (Fig. 2.7b).

Most of the snRNP proteins are lost during the B to C transition. In addition to di- and tri-snRNP proteins, complex C lacks U6-specific LSm ring, U2-specific SF3a and SF3b proteins and half of the U5-specific proteins, whereas proteins of the hPrp19 complex become tightly associated with U5 snRNP (Bessonov et al, 2008; Schmidt et al, 2014). The stable catalytic core is thus restricted only to U2/U6 and U5 snRNAs, U2 and U5 Sm rings, U2A', U2B" and U5-specific hPrp8, hBrr2, hSnu114 and 40K proteins (Schmidt et al, 2014). hPrp8 has the crucial role in the complex, since it resides in the heart of the spliceosome and brings components of the active site together (Li et al, 2013).

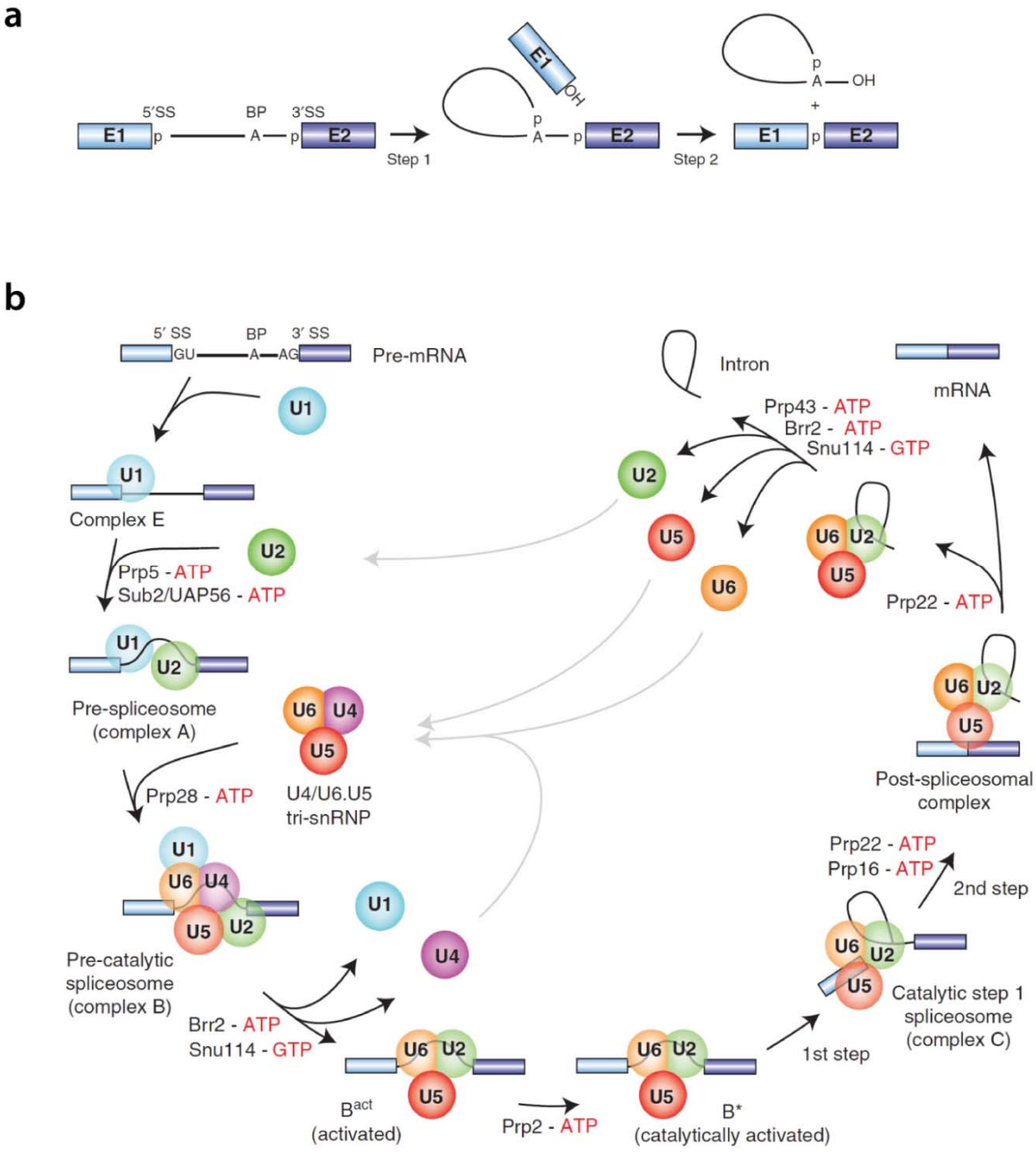


**Figure 2.7. Splicing of precursor mRNA.**

(a) The two-step mechanism of the splicing reaction. Boxes and solid line represent exons and intron, respectively; phosphate groups at the splice sites are shown. BP is for branch point.

(b) Spliceosome assembly, rearrangements and disassembly, as occurring during the splicing phase of the spliceosome cycle. For simplicity, snRNPs are represented by circles and non-snRNP proteins are not shown. Conformational changes are driven by eight DExD/H-box ATPases and hSnu114 GTPase as indicated at each step. After release from the spliceosome, snRNPs must be recycled as shown by gray arrows.

Reprinted from Will & Lührmann, 2011.

Complex C binding the splicing intermediate then goes through another rearrangement to catalyze the second splicing reaction. Spliced mRNA is released from post-spliceosomal complex (Fig. 2.7b) by hPrp22 RNA helicase that unwinds mRNA from U5 snRNA (Fourmann et al, 2013; Schwer, 2008). U2/U6 and U5 snRNPs, however, remain bound in one complex with excised lariat intron and must be liberated and recycled for another spliceosomal cycle afterwards. This step is performed by hNtr1 in cooperation with hPrp43 RNA helicase. In yeast, Ntr2 factor is also required, but no human Ntr2 homolog has been identified so far (Fourmann et al, 2013; Yoshimoto et al, 2009). Released lariat intron is linearized by a debranching enzyme hDBR1 and degraded (Yoshimoto et al, 2009).

### 2.4.2 SnRNP recycling

To complete the spliceosome cycle, post-spliceosomal snRNPs must be recycled and subsequently used for catalysis of another intron excision. SnRNP particles undergo an extensive remodeling during the splicing; however, it is believed that the Sm core remains stably bound and unaltered throughout the whole cycle. The reassembly of Sm-class snRNPs thus does not require nuclear export and cytoplasmic phase of the biogenesis, contrary to *de novo* synthesized particles.

Regeneration of snRNPs likely follows similar rules as the late nuclear biogenesis and takes place in Cajal bodies. In fact, the vast majority of snRNPs localized to CBs comprises of mature recycled particles (Staněk et al, 2008). The key players in di-snRNP reassembly process are SART3 and LSm proteins (Bell et al, 2002; Verdone et al, 2004). Since U6 snRNP loses its LSm ring during spliceosome rearrangements, it has to associate again prior to SART3 binding. These particles are then targeted to CBs where U6 and U4 snRNA reanneal and new di-snRNP forms (Staněk et al, 2008). In addition to the Sm ring, U4 snRNP likely contains also 15.5K protein, which is released from the spliceosome together with U4 snRNA. hPrp31 and hPrp3/hPrp4/CypH trimer are probably reused as well and interact with U4/U6 snRNAs to generate complete di-snRNP particle.

From all the particles, U5 snRNP undergoes the most extensive changes during the splicing catalysis. It is released as a 35S particle lacking four snRNP-specific proteins and associated with the hPrp19 complex that consists of numerous proteins (Makarov et

al, 2002; Yoshimoto et al, 2009). Thus, it needs to be regenerated to the 20S particle, but neither the mechanism nor the factors responsible for the process are known. Only after completing this step, U4/U6·U5 tri-snRNPs can be remade.

In case of U2, post-spliceosomal snRNPs lack all the SF3 proteins and are released in the form of intermediate 12S particle (Yoshimoto et al, 2009). The following reassembly seems to occur again in Cajal bodies. Like U6, U2 snRNA secondary structure changes during splicing, suggesting an existence of a U2 snRNA chaperone, though no such enzyme has been identified. Generally, U snRNAs are very stable, having a life-time of several tens of hours (Fury & Zieve, 1996), indicating that each snRNP can be recycled multiple times. This option is further supported by a low number of newly assembled snRNPs in the nucleus and their inability to maintain the splicing without a help of mature snRNPs (Staněk et al, 2008).

## 2.5 Minor spliceosome

The chapters above discuss the general mechanisms of biogenesis and function of conventional spliceosomal snRNPs. However, a small subgroup of introns is spliced by a different splicing machinery called minor or U12-type spliceosome. These introns are rather rare, less than 0.5 % of all introns represent U12-type subgroup in human genome (Levine & Durbin, 2001; Turunen et al, 2013). But at the same time, minor introns are highly evolutionarily conserved all across eukaryotic taxa (Burge et al, 1998; Turunen et al, 2013).

Minor introns are characterized by more conserved 5' ss and branch point sequences compared to major (U2-type) introns, lack of polypyrimidine region and frequent presence of AT-AC terminal dinucleotides in addition to common GT-AG termini (reviewed in Turunen et al, 2013). Due to the latter feature minor introns are also referred to as atac introns. Since the sequence of U12-type introns is far more conserved, a single point mutation is sufficient to turn it into U2-type intron but not *vice versa*. U12-type introns are thus prone to be converted into U2-type during evolution, indicating that minor introns might once have been much more common than they are now (Burge et al, 1998).

Despite their high conservation, U12-type introns are spliced more slowly than other introns (Patel et al, 2002). That is caused by lowered stability of one of the minor spliceosome components, U6atac snRNP, providing a possible regulatory mechanism for the expression of U12-type intron-containing genes (Younis et al, 2013). Such a function could give an explanation of preservation of a U12-type intron subset in evolution.

Similarly to the major spliceosome, the minor one also consists of five snRNP particles. U1, U2, U4 and U6 are replaced by their functional homologs called U11, U12, U4atac and U6atac snRNP, respectively, while U5 snRNP is shared by both spliceosomes. Although having a different sequence, minor snRNAs and their major counterparts exhibit a strikingly similar secondary structure. Protein composition is also rather conserved between minor and major snRNPs. The most obvious difference is thus the existence of pre-formed U11/U12 di-snRNP, the association of which is protein mediated (Wassarman & Steitz, 1992). U11/U12 shares with U1 and U2 snRNPs only Sm proteins and SF3b complex and contains seven additional U11/U12-specific proteins (Will et al, 2004). The other three snRNPs then form U4atac/U6atac·U5 particle which has identical protein composition as the major tri-snRNP (Schneider et al, 2002).

Both U11/U12 and U4atac/U6atac·U5 snRNPs assembly probably occurs in Cajal bodies, as supported by several evidences. First, U11 and U12 snRNAs are accumulated in CBs (Matera & Ward, 1993). Second, di-snRNP-specific proteins bind to U4atac/U6atac snRNAs in the same hierarchical manner as in the case of U4/U6 snRNAs (Nottrott et al, 1999; Nottrott et al, 2002). And finally, U4atac/U6atac snRNP assembly is assisted by SART3, even though SART3 interaction with U4atac/U6atac seems to be weaker than with U4/U6 snRNP (Damianov et al, 2004). Interestingly, sequences of U6atac and U6 snRNAs recognized by SART3 are the most conserved regions shared by both snRNAs (Bell et al, 2002; Damianov et al, 2004).

Major snRNPs and their minor counterparts play the same roles during the splicing catalysis. Since U11 and U12 snRNPs interact prior to pre-mRNA binding, the stage of complex E is skipped and spliceosome assembly starts with complex A formation. SF3b proteins contact the pre-mRNA branch point same way as in the major spliceosome (Golas et al, 2005). During the spliceosome activation, U4atac/U6atac snRNA duplex is unwound, U6atac base pairs with U12 snRNA and U11 together with U4atac snRNP is released from the spliceosome. The catalytic core is therefore formed by U12/U6atac and U5 snRNAs (reviewed in Turunen et al, 2013) . After the splicing,

minor snRNPs have to be recycled. In the case of U4atac/U6atac·U5 snRNP, the recycling is driven by SART3 protein (Damianov et al, 2004); however, the mechanism of U11/U12 recycling remains unknown.

# 3 Aims of the Thesis

SART3 protein is a crucial di-snRNP assembly factor. Accumulating in nuclear Cajal bodies, SART3 facilitates U4/U6 snRNP assembly and tethers these incomplete particles to coilin, the scaffolding protein of Cajal bodies. In the nucleoplasm, SART3 specifically binds U6 snRNPs and targets them to CBs to promote their efficient incorporation into di- and tri-snRNP particles. This interaction has been well studied, revealing that it is the C-terminal part of the protein which is responsible for U6 snRNP binding. While the RRM domain associates directly with U6 snRNA (Bell et al, 2002), the CT domain recognizes U6-specific LSm proteins (Rader & Guthrie, 2002).

We recently showed that SART3 functions as an anchor sequestering incomplete snRNPs in Cajal bodies. Interestingly, not only U6 but also U4 and U5 snRNPs accumulation in CBs is SART3 dependent (Novotný et al, 2015). The mechanism of the accumulation is not known though. The aim of this thesis is to reveal how SART3 interacts with spliceosomal snRNPs. In particular, we aim to elucidate which of the SART3 domains are used for the interactions and which snRNP components they interact with. We further want to investigate possible roles of SART3-snRNP interactions in snRNP assembly and quality control processes.

# 4 Material and Methods

## 4.1 Material

### 4.1.1 Primers

| Primer | Sequence (5' to 3') | Used for |
|---|---|---|
| Random hexamers | NNNNNN | Reverse transcription |
| U2 snRNA_F | CTCGGCCTTTTGGCTAAGAT | qPCR |
| U2 snRNA_R | CGTTCCTGGAGGTACTGCAA | qPCR |
| U4 snRNA_F | TGGCAGTATCGTAGCCAATG | qPCR |
| U4 snRNA_R | CTGTCAAAAATTGCCAGTGC | qPCR |
| U5 snRNA_F | CTCTGGTTTCTCTTCAGATC | qPCR |
| U5 snRNA_R | TGTTCCTCTCCACGGAAATC | qPCR |
| U6 snRNA_F | CGCTTCGGCAGCACATATAC | qPCR |
| U6 snRNA_R | AAAATATGGAACGCTTCACGA | qPCR |
| attB1-SART3_F | GGGGACAAGTTTGTACAAAAAAGCA GGCTCGATGGCGACTGCGGCCGAA | Gateway cloning |
| attB2-STOP-SART3_R | GGGGACCACTTTGTACAAGAAAGCTG GGTTATTACAGGGAGGCTGCCTTCTC | Gateway cloning |

**Table 4.1. List of primers.**

### 4.1.2 Primary antibodies

| Antibody | Origin | Producer | Used for |
|---|---|---|---|
| anti-coilin | Mouse | M. Carmo-Fonseca (Almeida et al, 1998) | Immunofluorescence |
| anti-coilin | Rabbit | Santa Cruz Biotechnology | WB |
| anti-FLAG | Mouse | Sigma-Aldrich | WB |
| anti-GFP | Mouse | Santa Cruz Biotechnology | WB |
| anti-hPrp31 | Rabbit | R. Lührmann (Makarova et al, 2002) | WB |
| anti-SART3 | Rabbit | D. Staněk (Staněk et al, 2003) | Immunofluorescence |

| | | | |
|---|---|---|---|
| anti-SF3a60 | Rabbit | Abcam | WB |
| anti-SF3b49 | Mouse | Abcam | WB |
| anti-SmB/B' (Y12) | Mouse | Facility of IMG ASCR | WB |
| anti-U2B" | Mouse | Progen | WB |

**Table 4.2. List of primary antibodies used for Western blot (WB) and immunofluorescent staining.**

### 4.1.3 siRNAs

| **siRNA** | **Sequence of sense strand (5' to 3')** | **Producer** |
|---|---|---|
| coilin<br>(Silencer Select Pre-designed siRNA) | GAAGGACUAUAGUCUGUUAtt | Ambion |
| SART3<br>(Silencer Select siRNA) | ACUGCUACGUGGAGUUUAAtt | Ambion |
| SF3a60<br>(Silencer Pre-designed siRNA) | GCUCAAUGCCAUUUCAGGAtt | Ambion |
| SmB/B'<br>(Silencer Select Pre-designed siRNA) | UGGUCUCAAUGACAGUAGAtt | Ambion |
| SmG<br>(Silencer Select Pre-designed siRNA) | GUAAUACGAGGAAAUAGUAtt | Ambion |
| NC1<br>(Silencer Negative Control 1 siRNA) | | Ambion |
| NC5<br>(Silencer Negative Control 5 siRNA) | | Ambion |

**Table 4.3. List of siRNAs.**

### 4.1.4 Instruments

| |
|---|
| Olympus IX70 microscope with DeltaVision system (Applied Precision) |
| Scan^R microscopic system (Olympus) |
| LightCycler 480 System (Roche Applied Science) |
| MJ Mini Personal Thermal Cycler (BioRad) |
| Imager LAS-3000 (Fujifilm) |
| Imager Universal Hood II (BioRad) |
| Spectrophotometer NanoDrop 2000 (Thermo Scientific) |
| Flow box (Clean Air Techniek B.V.) |

| |
|---|
| $CO_2$ incubator (SANYO) |
| Shaking $CO_2$ incubator New Brunswick (Eppendorf) |
| Water bath (Julabo) |
| 37 °C incubator Raven2 (LTE Scientific) |
| Centrifuge 5417R (Eppendorf) |
| Centrifuge Biofuge pico (Heraeus) |
| Vacuum centrifuge Speed Vac (Savant) |
| Thermomixer Comfort (Eppendorf) |
| Heat-stir CB162 (P-Lab) |
| Orbital Shaker OS-10 (BIOSAN) |
| Mini Rocker MR-1 (BIOSAN) |
| Sample Mixer (Invitrogen) |
| 37 °C shaker ORBI-SAFE (SANYO) |
| Vortex Genie2 (Scientific Industries) |
| Sonicator U50 Control (IKA Labortechnik) |
| Power supply PowerPac Universal (BioRad) |
| Horizontal electrophoresis |
| Vertical electrophoresis Mini-Protean Tetra Cell (BioRad) |
| Vertical electrophoresis for RNA gels |

**Table 4.4. List of instruments.**

## 4.2 Methods

### 4.2.1 Cell culture

HeLa cells were cultured in high-glucose (4500 mg/l) DMEM medium (Sigma-Aldrich) supplemented with 10% fetal bovine serum, penicillin and streptomycin. Cells were grown in 37 °C and 5% $CO_2$ atmosphere.

HEK293T cells were cultured in FreeStyle F17 Expression medium (Life Technologies) in the shaking incubator in 37 °C and 5% $CO_2$ atmosphere.

### 4.2.2 Transfection of plasmid DNA and siRNA to HeLa cells

HeLa cells used for immunoprecipitation were transfected by X-tremeGENE HP DNA transfection reagent (Roche) used in ratio 2 µl of the reagent to 1 µg of plasmid DNA. Mixture of transfection reagent, DNA and Opti-MEM reduced serum medium (Invitrogen) was incubated at room temperature for 15 min and then added to cell medium. Cells were used for experiments 24 h after the transfection.

| | **Opti-MEM** | **X-tremeGENE HP** | **DNA** |
|---|---|---|---|
| 10 cm Petri dish (10 ml medium) | 250 µl | 6 µl | 3 µg |

**Table 4.5. Composition of X-tremeGENE HP DNA transfection mixture.**

HeLa cells used for microscopic observation were transfected with Lipofectamine LTX reagent (Invitrogen) used in ratio 3 µl of the reagent to 2 µg of plasmid DNA. Transfection reagent, DNA and serum free DMEM medium were mixed, incubated at room temperature for 20 min and added to cell medium. Cells were used for experiments 24 h after the transfection.

| | **DMEM** | **Lipofectamine LTX** | **DNA** |
|---|---|---|---|
| 6-well plate (2 ml medium) | 500 µl | 3 µl | 2 µg |
| 12-well plate (1 ml medium) | 250 µl | 1,5 µl | 1 µg |

**Table 4.6. Composition of Lipofectamine LTX DNA transfection mixture.**

siRNA was transfected into HeLa cells using Oligofectamine reagent (Invitrogen). Transfection reagent and siRNA were mixed separately with Opti-MEM medium and incubated at room temperature for 5 min. Then both mixtures were merged and added to cell medium after 20 min incubation. Final siRNA concentration in the medium varied from 20nM to 50nM, depending on particular siRNA. Cells were used for experiments 48 h after the transfection.

| | Opti-MEM (mix with siRNA) | siRNA | Opti-MEM (mix with Oligofectamine) | Oligofectamine |
|---|---|---|---|---|
| 10 cm Petri dish (4 ml medium) | 140 µl | to required final concentration | 40 µl | same as siRNA |
| 6-well plate (2 ml medium) | 70 µl | to required final concentration | 20 µl | same as siRNA |
| 12-well plate (1 ml medium) | 35 µl | to required final concentration | 10 µl | same as siRNA |

**Table 4.7. Composition of Oligofectamine siRNA transfection mixture.**

### 4.2.3 Transfection of plasmid DNA to HEK293T cells

HEK293T cells cultured in FreeStyle F17 Expression medium were transfected using linear polyethylenimine (Polysciences). 1 µg DNA used per 0.5 ml of cell suspension in F17 medium was diluted in PBS to 25 µl and mixed with linear polyethylenimine. The reaction mixture was incubated at room temperature for 10 min and added to cells. After 4 h equal amount of EX-CELL 293 serum-free medium (Sigma-Aldrich) was added. Cells were used for experiments 72 h after the transfection.

| | DNA | PBS | Linear polyethylenimine (1 mg/ml) |
|---|---|---|---|
| 0.5 ml cell suspension (1 ml after EX-CELL 293 medium addition) | 1 µg | to 25 µl | 3 µl |

**Table 4.8. Composition of linear polyethylenimine DNA transfection mixture for HEK293T cells.**

### 4.2.4 Immunoprecipitation

HeLa cells in 10 cm Petri dish were washed 3x with PBS buffer, harvested into 1 ml PBS and transferred to microcentrifuge tube. Cells were then centrifuged at 1 000 g and 4 °C for 5 min, PBS was removed and cells resuspended in 500 µl NET2 buffer + 1 µl Recombinant RNasin (Promega) + 2 µl Protease Inhibitor Cocktail Set III (Calbiochem). The samples were sonicated on ice bath by 3x 30 pulses (0.5 s pulses at 60% amplitude) and then centrifuged at 20 000 g and 4 °C for 10 min. Pellets were

discarded and 2x 15 µl of the supernatant were saved as 'input' samples. The rest of the supernatant was pre-cleaned for subsequent immunoprecipitation by incubating on a rotator with 10 µl Protein G Sepharose beads (GE Healthcare) at 4 °C for 1 h.

For immunoprecipitation step, 30 µl Protein G Sepharose beads were washed 3x with NET2 buffer and incubated on a rotator with 450 µl NET2 + 0.4 µl goat anti-GFP antibody (15 mg/ml, D. Drechsel, MPI-CBG Dresden) at 4 °C for at least 2 h. Beads were then washed 3x with PBS and incubated on a rotator with pre-cleaned cell lysates at 4 °C over night. Next day, lysates were discarded, beads washed 3x with NET2 buffer and divided into 2 halves for protein and RNA isolation fractions, respectively.

| |
|---|
| 150mM NaCl |
| 50mM Tris-HCl pH 7.4 |
| 0.05% IGEPAL CA-630 (Sigma-Aldrich) |

**Table 4.9. Composition of NET2 buffer.**

### 4.2.5 Immunoprecipitation with RNase treatment

In the immunoprecipitation experiment with RNase treatment, cells were harvested and centrifuged as described above and resuspended in 1 ml NET2 buffer + 4 µl Protease Inhibitor Cocktail Set III and divided into 2 halves. 1 µl Recombinant RNasin was added into first half, cells in both tubes were sonicated and centrifuged. Supernatants were then transferred to new tubes, second half was supplemented with 5 µl RNase A and both were incubated at 37 °C for 10 min. Subsequent pre-cleaning of cell lysates and immunoprecipitation were performed as described above.

### 4.2.6 Immunoprecipitation with recombinant N-SART3

Cell lysates from HeLa cells grown in 10 cm Petri dish were prepared as described above. Then 5 µg of purified recombinant N-SART3 protein were added and the mixture was incubated on a rotator at 4 °C for 3 h to allow N-SART3 incorporation into snRNP particles. Pre-cleaning of the lysates was done by incubation with 10 µl Protein G Sepharose beads over night.

To decrease a background, 30 µl Protein G Sepharose beads were blocked by 50 µl BSA (10 mg/ml) + 25 µl Sheared salmon sperm DNA (10 mg/ml, Ambion) in

425 µl NET2 buffer at 4 °C for 1 h. Blocked beads were then incubated with 1 µl mouse anti-FLAG M2 (1 mg/ml) or 4 µl rabbit anti-FLAG antibody (0.8 mg/ml, both produced by Sigma-Aldrich) in 450 µl NET2 on the rotator at 4 °C over night. These steps as well as the subsequent immunoprecipitation step were done in siliconized microcentrifuge tubes (Sigma-Aldrich) to further eliminate the background.

Next day, 2x 15 µl of pre-cleaned lysates were saved as input and the rest was poured on the beads 3x washed with PBS. Immunoprecipitation was performed at 4 °C for 4 h. Beads were afterwards 3x washed with NET2 and divided into 2 halves.

### 4.2.7 Protein isolation after immunoprecipitation

After immunoprecipitation (IP), NET2 was removed from the beads which were then mixed with 30 µl 2x concentrated sample buffer. Input samples taken from cells lysates were mixed with 15 µl 2x concentrated sample buffer. All mixtures were incubated at 95 °C for 5 min and then stored at -20 °C. These samples were used for SDS-PAGE, 10 µl of both IP and input were loaded.

| |
|---|
| 20% glycerol |
| 4% SDS |
| 2% 2-mercaptoethanol |
| 250mM Tris-HCl pH 6.8 |
| 0.02% bromphenol blue |

**Table 4.10. Composition of 2x sample buffer.**

### 4.2.8 RNA isolation after immunoprecipitation

After IP, beads were resuspended in 50 µl NET2 buffer, input samples were mixed with 35 µl NET2. Both IP and input samples were mixed with 50 µl phenol-chloroform 5:1 pH 4.3-4.7, thoroughly vortexed for 1 min and centrifuged at 20 000 g and 4 °C for 10 min. After centrifugation, 45 µl from the upper aqueous phase was transferred to new microcentrifuge tube, while the lower organic phase was discarded. 5 µl 3M sodium acetate pH 5.2, 3 µl glycogen and 125 µl 100% ethanol were added and the mixture was incubated at -20 °C for at least 30 min. Precipitated RNA was centrifuged at 20 000 g and 4 °C for 10 min, pellet washed with 500 µl 70% ethanol and

centrifuged again at same conditions. Ethanol was then removed, pellets in opened tubes were dried by air and resuspended in 10 µl nuclease-free water. Isolated RNA was stored at -80 °C.

### 4.2.9 qRT-PCR

For reverse transcription, 2 µl of RNA isolated from IP and input samples were used in the 20 µl reaction mixture. 1 µl of HeLa total RNA was used in the control reaction.

| | |
|---|---|
| RNA (or HeLa total RNA) | 2 µl (or 1 µl) |
| Random hexamers | 1 µl |
| 10mM dNTPs | 1 µl |
| $dH_2O$ | 10.5 µl (or 11.5 µl) |

**Table 4.11a. Composition of reverse transcription reaction mixture.**

Reaction mixture was incubated at 65 °C for 5 min. Then following reagents were added:

| | |
|---|---|
| 5x First Strand buffer (Invitrogen) | 4 µl |
| 0.1M DTT (Invitrogen) | 1 µl |
| SuperScript III Reverse Transcriptase (Invitrogen) | 0.5 µl |

**Table 4.11b. Composition of reverse transcription reaction mixture.**

All reagents were mixed by pipetting up and down and incubated at 25 °C for 5 min and then at 50 °C for 40 min. The reaction was inactivated by heating at 70 °C for 15 min.

Complementary DNA (cDNA) was diluted 1:10 for qPCR reactions with U2 and U6 snRNA primers and 1:100 for reactions with U4 and U5 snRNA primers.

| | |
|---|---|
| diluted cDNA | 2 µl |
| SYBR Green I Master (Roche) | 2.5 µl |
| 10µM primer F | 0.25 µl |
| 10µM primer R | 0.25 µl |

**Table 4.12. Composition of qPCR reaction mixture.**

All samples were loaded on the plate in doublets. Following qPCR program was used for cDNA amplification:

| | | | |
|---|---|---|---|
| Initial denaturation: | 95 °C | 7 min | 1 cycle |
| Quantification - Denaturation: | 95 °C | 20 s | 40 cycles |
| Annealing: | 61 °C | 20 s | |
| Elongation: | 72 °C | 35 s | |
| Melting curves analysis: | 95 °C | 15 s | 1 cycle |
| | 55 °C | 1 min 1 s | |
| | 37 °C | 1 min 1 s | |

**Table 4.13. Program for quantitative PCR.**

Average Ct value was calculated for each doublet and Ct of IP samples were then normalized to input values:

$N = 2^{(Ct_{input} - Ct_{IP})}$.

### 4.2.10 Silver stained RNA gel

RNA isolated after immunoprecipitation was also analysed by separation on denaturing polyacrylamide gel. 5 µl of IP or input RNA samples were mixed with 10 µl 0.5x concentrated sample buffer and evaporated to the final volume of 5 µl. 4M urea from 0.5x sample buffer got concentrated to 8M during this step. Samples were then incubated at 65 °C for 15 min.

| |
|---|
| 4M urea |
| 10mM Tris-HCl pH 8.0 |
| 0.1% xylene blue |

**Table 4.14. Composition of 0.5x sample buffer.**

For our RNA gel apparatus 40 ml gel mixture was needed. First, urea, 10x TBE and acrylamide were mixed and heated on the stirrer at 75 – 90 °C until urea got dissolved. Then the gel was cooled to room temperature, mixed with APS and TEMED and water was added to final volume. The gel was immediately poured into the apparatus. After polymerization, the gel was warmed by the pre-run at 500 V for at least

15 min. 2.5 µl of input RNA and 5 µl of IP RNA were loaded and separated at 500 V for 2 h 40 min.

| | |
|---|---|
| urea | 19.2 g |
| 30% acrylamide/bis-acrylamide 19:1 | 13.3 ml |
| 10x TBE | 4 ml |
| 10% APS | 400 µl |
| TEMED | 25 µl |
| $dH_2O$ | to 40 ml |

**Table 4.15. Composition of 7M urea/10% polyacrylamide RNA gel.**

| |
|---|
| 0.89M Tris base |
| 0.89M boric acid |
| 20mM EDTA |

**Table 4.16. Composition of 10x TBE buffer.**

When RNA was sufficiently separated, RNA gel was fixed by 40% methanol + 10% acetic acid for 30 min. Then it was treated by 3.4mM $K_2Cr_2O_7$ + 3.2mM $HNO_3$ for 10 min, 3x briefly washed with water and silver stained for 30 min (12mM $AgNO_3$). After staining, the gel was washed with water again and the image was developed by 280mM $Na_2CO_3$ + 0.02% formaldehyde. The reaction was stopped by washing in 1% acetic acid for at least 10 min.

### 4.2.11 SDS PAGE

SDS PAGE gels were poured into BioRad gel chamber, 10% separating gel was used in all experiments. Samples were loaded together with 4 µl PageRuler Plus pre-stained protein ladder (Thermo Scientific). The gel was ran at 75 V for 30 min and then at 110 V for 1 h – 1 h 20 min until sufficient separation.

| | |
|---|---|
| 40% acrylamide/N,N'-methylenebisacrylamide 37.5:1 | 1.36 ml |
| 1.5M Tris-HCl pH 8.8 | 1.3 ml |
| 10% SDS | 50 µl |
| $dH_2O$ | 2.24 ml |

| | |
|---|---|
| 10% APS | 50 µl |
| TEMED | 2 µl |

**Table 4.17. Composition of 10% separating gel (5 ml).**

| | |
|---|---|
| 40% acrylamide/N,N'-methylenebisacrylamide 37.5:1 | 272 µl |
| 1M Tris-HCl pH 6.8 | 260 µl |
| 10% SDS | 20 µl |
| $dH_2O$ | 1.426 ml |
| 10% APS | 20 µl |
| TEMED | 2 µl |

**Table 4.18. Composition of stacking gel (2 ml).**

| |
|---|
| 25mM Tris-HCl |
| 192mM glycine |
| 0.1% SDS |

**Table 4.19. Composition of SDS running buffer.**

### 4.2.12 Coomassie blue staining

The gel was microwaved 3x in distilled water for 1 min and then in Coomassie blue staining solution for 1 min 30 s. It was incubated on the shaker with staining solution for another 5 min, microwaved in water for 1 min 30 s and decoloured by shaking in water over night.

### 4.2.13 Western blotting

After electrophoresis, the gel was soaked in transfer buffer. Protran 0.45 µm nitrocellulose membrane (GE Healthcare), filter papers and sponges were treated in the same way. Blotting sandwich was made by the gel and the membrane in the middle with 3 layers of filter paper and 1 sponge on each side. Western blotting was performed in BioRad wet blotting apparatus at a constant current of 360 mA for 1 h. The membrane was then washed in PBST (PBS supplemented with 0.05% Tween) and blocked by incubation on the shaker in 10% low fat milk/PBST for 1 h. Primary antibody was diluted to required concentration in 1% low fat milk/PBST and the membrane was

stained on shaker at room temperature for 1 h. Afterwards, it was 3x washed in 1% low fat milk/PBST, stained by appropriate secondary antibody in 1% low fat milk/PBST for 1 h and washed again. For all experiments, AffiniPure goat secondary antibodies conjugated with horse radish peroxidase (Jackson ImmunoResearch) were used. The membrane was incubated with SuperSignal West chemiluminescent substrate (Thermo Scientific) for 2 min and the image developed in LAS-3000 imager.

| |
|---|
| 25mM Tris-HCl |
| 192mM glycine |
| 20% methanol |

**Table 4.20. Composition of transfer buffer.**

### 4.2.14 PCR

Phusion High-Fidelity DNA Polymerase (New England BioLabs) was used for DNA amplification. For each primer set 8 PCR reactions per 20 µl were prepared and ran at different annealing temperatures to find optimal conditions for DNA amplification.

| | |
|---|---|
| Phusion High-Fidelity DNA Polymerase | 0.2 µl |
| 5x Phusion HF Buffer | 4 µl |
| 10mM dNTPs | 0.4 µl |
| 10µM primer F | 1 µl |
| 10µM primer R | 1 µl |
| template DNA | 2 ng |
| $dH_2O$ | to 20 µl |

**Table 4.21. Composition of PCR reaction mixture.**

| | | | |
|---|---|---|---|
| Initial denaturation: | 98 °C | 30 s | 1 cycle |
| Denaturation: | 98 °C | 10 s | 20 cycles |
| Annealing: | 55-68 °C gradient | 15 s | |
| Elongation: | 72 °C | 1 min | |
| Final elongation: | 72 °C | 5 min | 1 cycle |

**Table 4.22. Program for PCR.**

### 4.2.15 Gateway cloning

The Gateway Technology (Invitrogen) was advantageously used for N-SART3 cloning. PCR product flanked by attB recombination sites was purified by DpnI restriction and centrifugation in 30% PEG 8000/30mM $MgCl_2$ as described in manufacturer's protocol. This treatment helped to reduce background in subsequent recombination by removal of template plasmid DNA and primers. The BP recombination between the PCR product and pDONR221 vector (Invitrogen) was performed accordingly to manufacturer's recommendations as well as the LR recombination between the entry clone (result of the BP recombination) and the destination vector pDEST_MM322 (kindly provided by C. Bařinka, IBT AS CR).

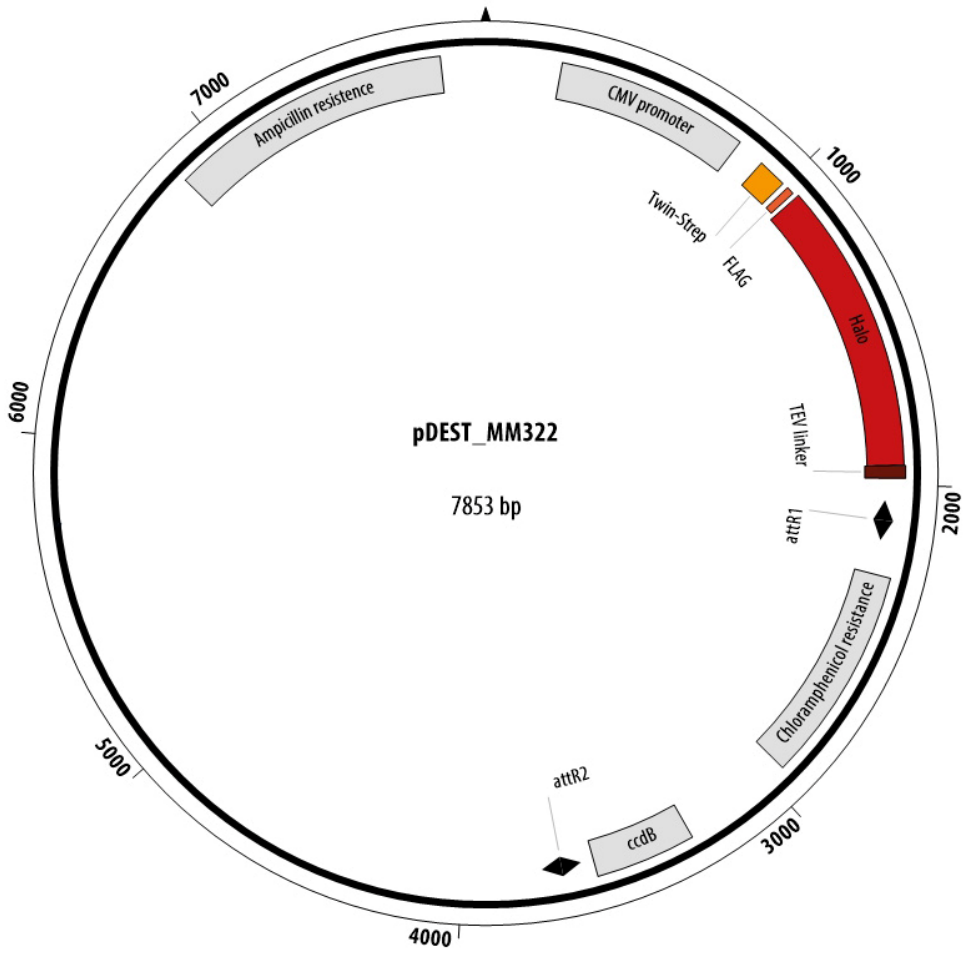


**Figure 4.1. Map of the Gateway destination vector pDEST_MM322.**

### 4.2.16 Transformation of competent bacteria and plasmid DNA preparation

Competent strain of DH5α bacteria was thawed on ice and 30 µl were used for one transformation reaction. It was mixed with 1 µl of recombination reaction product and incubated on ice for 20 min. Heat shock was performed by incubation in water bath at 42 °C for 30 – 40 s. Afterwards, 350 µl SOC medium was added and the mixture was incubated on the thermoshaker at 37 °C for 1 h. Bacteria were grown on agar plates supplemented with appropriate antibiotics at 37 °C over night.

| | |
|---|---|
| tryptone | 2 g |
| yeast extract | 0.5 g |
| 1M NaCl | 1 ml |
| 1M $MgSO_4$ | 1 ml |
| 1M $MgCl_2$ | 1 ml |
| 2M glucose | 1 ml |
| $dH_2O$ | to 100 ml |

**Table 4.23. Composition of SOC medium (100 ml).**

Bacterial colonies were inoculated to 2 ml LB medium (1% tryptone + 0.5% yeast extract + 1% NaCl) supplemented with appropriate antibiotics and grown on the shaker at 37 °C over night but not longer than 16 h. Plasmid DNA preparation was performed using QIAprep Spin Miniprep kit (Qiagen) as described in manufacturer's protocol.

### 4.2.17 Restriction digestion

Recombination products of the Gateway cloning were restriction analyzed using ApaI restriction enzyme (New England BioLabs). Reaction mixture containing approximately 1 µg DNA, ApaI enzyme and 10x NEBuffer 4 (New England BioLabs) was incubated at 25 °C for 1 h. Digestion products were detected on 1% agarose gels.

| | |
|---|---|
| DNA | 1 µg |
| ApaI restriction enzyme | 0.5 µl |
| 10x NEBuffer 4 | 1 µl |
| $dH_2O$ | to 10 µl |

**Table 4.24. Composition of digestion reaction mixture.**

### 4.2.18 Purification of recombinant N-SART3

For recombinant N-SART3 purification, 19 ml of HEK293T cells transfected with pDEST_MM322_N-SART3 plasmid DNA were used. Cells were centrifuged at 500 g and 4 °C for 5 min and resuspended in 1.5 ml lysing buffer A mixed with 1/10 cOmplete protease inhibitor cocktail tablet (Roche) and 0.25 µl benzonase. Cells were sonicated 3x by 10 s pulses at 3 V, 30 µl 10% IGEPAL CA-630 were added and the mixture was incubated on ice for 30 min. Afterwards, it was centrifuged at 4 500 g and 4 °C for 15 min and the pellet was discarded, this step was then repeated at 20 000 g and 4 °C for 30 min and 5 µl of the supernatant was saved as 'supernatant' fraction.

The purification was performed via N-terminal Twin-Strep tag under gravity flow. The column was filled with 3 ml Strep-Tactin Superflow 50% suspension (IBA) and the resin was 2x washed with buffer A. 1.5 ml cell lysate supernatant was loaded on the column and saved as 'flow through' fraction. The resin was then 2x washed with 7.5 ml buffer A (saved as 'wash 1, 2' fraction). Recombinant N-SART3 bound to the resin was eluted by elution buffer C, 7 fractions of 2 ml were saved as 'elution 1-7'. All the fractions were analyzed on SDS PAGE, 1 µl supernatant and flow through samples and 15 µl wash and elution samples were loaded.

| |
|---|
| 0.1M Tris pH 8 |
| 10mM NaCl |
| 5mM KCl |
| 2mM $MgCl_2$ |
| 10% glycerol |

**Table 4.25. Composition of buffer A.**

| |
|---|
| 50mM Tris pH 8 |
| 150mM NaCl |
| 10mM KCl |
| 10% glycerol |
| 3mM desthiobiotin |

**Table 4.26. Composition of buffer C.**

### 4.2.19 Immunofluorescent staining and FISH

Cells grown on coverslips were washed 3x with PBS buffer, fixed with 4% paraformaldehyde/PIPES for 10 min, washed again 3x with PBS and permeabilized with 0.5% Triton X-100/PBS for 5 min. After another wash with PBS, cells were incubated on drops of 0.1M glycine/0.2M Tris-HCl pH 7.4 for 5 min and washed with 4x concentrated SSC. Before the hybridization step, coverslips were put on 2x SSC/50% formamide for at least 10 min. Then they were transferred to pre-warmed moisture chamber on drops of 7.5 µl formamide + 7.5 µl FISH master mix + 0.5 µl fluorescent probe labeled by Cy3 at its 5' end and incubated at 37 °C for 1 h.

Cells with bound probe were washed with 2x SSC/50% formamide at 37 °C for 20 min, 2x SSC at 37 °C for 20 min, 1x SSC at room temperature for 20 min and finally with 4x SSC at room temperature for 10 min. Cells were then blocked with 5% Normal goat serum (Jackson ImmunoResearch) in 4x SSC for 5 min. Staining with primary antibody diluted in 4x SSC was done at room temperature for 1 h. Coverslips were washed 3x with 4x SSC, stained with appropriate secondary antibody in 4x SSC for 1 h and washed again 3x with 4x SSC. AffiniPure goat secondary antibodies conjugated with DyLight488 or Cy5 (Jackson ImmunoResearch) were used. Coverslips were mounted to microscope slides using DAPI Fluoromount-G medium (SouthernBiotech).

| | |
|---|---|
| paraformaldehyde | 2 g |
| 0.2M PIPES pH 6.9 | 25 ml |
| 1M $MgCl_2$ | 100 µl |
| 0.5M EGTA pH 8 | 125 µl |
| $dH_2O$ | 25 ml |

**Table 4.27. Composition of 4% paraformaldehyde/PIPES.**

| |
|---|
| 3M NaCl |
| 300mM sodium citrate |
| pH 7.0 (adjusted with HCl) |

**Table 4.28. Composition of 20x SSC.**

| | |
|---|---|
| 50% dextran sulphate | 40 µl |
| 5% BSA | 40 µl |
| 20x SSC | 20 µl |

**Table 4.29. Composition of FISH master mix (100 µl).**

| **Target RNA** | **Probe sequence (5' to 3')** |
|---|---|
| U2 snRNA | GAACAGATACTACACTTGATCTTAGCCAAAAGGCCGAGAAGC |

**Table 4.30. Sequence of hybridization probe.**

Cells were imaged by DeltaVision microscopic system (Applied Precision) coupled to Olympus IX70 microscope. The microscope was equipped with oil immersion objective (60x/1.4 NA) and 20 Z-stacks with 200 nm spacing were collected for each image. Image restoration was done by SoftWorx deconvolution system (Applied Precision) using a measured point spread function, as described previously (Novotný et al, 2011).

For high-content microscopy, images were automatically acquired by Scan^R system (Olympus) equipped with oil immersion objective (60x/1.35 NA). In each sample, 225 positions were scanned and 10 Z-stacks with 300 nm spacing were taken at each position. Cell nuclei were automatically detected based on DAPI fluorescence intensity and Cajal bodies were identified by coilin immunostaining using Analysis Scan^R software. For both nuclei and Cajal bodies total intensity and area were measured and the ratio between fluorescence intensity of the probe in CBs and nucleoplasm was calculated according to:

$$R = \frac{\frac{\sum total\ IF_{CBs}}{\sum area_{CBs}}}{\frac{total\ IF_{nucleus} - \sum total\ IF_{CBs}}{area_{nucleus} - \sum area_{CBs}}},$$

as described previously (Novotný et al, 2015).

# 5 Results

## 5.1 N-terminal part of SART3 interacts with U2 snRNP

For mapping of SART3 domain interaction with individual snRNPs, we used three different plasmid constructs. Full length SART3 and its two truncation mutants called N-SART3 and C-SART3 were cloned into pEGFP-N1 vectors, resulting in SART3 proteins GFP-tagged at their C-termini. N-SART3 covered first two thirds of the full length protein containing E domain, TPR motifs and nuclear localization signal (NLS), while C-SART3 comprised of NLS and C-terminal RRM and CT domains (Fig. 5.1). All of these constructs were used before and no influence of the GFP tag on protein functionality has been observed (Staněk et al, 2003).

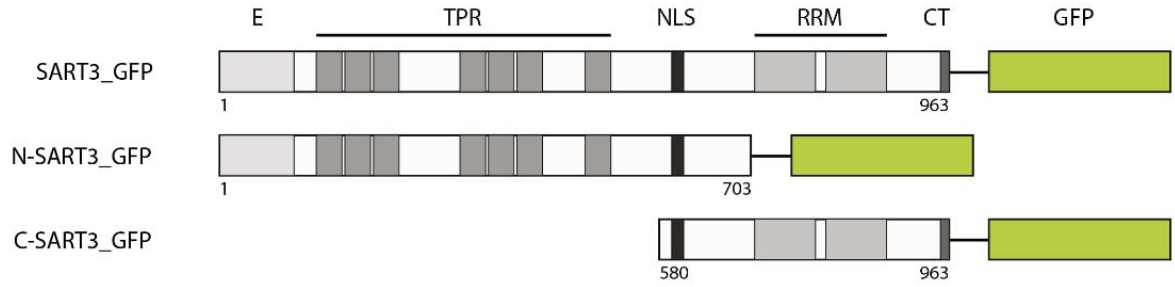


**Figure 5.1. Domain organization of SART3 constructs.**
Full length SART3 and two different truncation mutants were used; all of them were GFP-tagged at their C-termini. N-SART3 (amino acids 1 to 703) contains E domain, TPR motifs and NLS, C-SART3 (amino acids 580 to 963) contains NLS, RRM domain and C-terminal domain.

We transiently transfected SART3 constructs into HeLa cells and immunoprecipitated them from cell lysates via GFP tag. Co-precipitated RNA was isolated and resolved on silver stained denaturating RNA gel (Fig. 5.2a, Novotný et al, 2015) . As expected, U6 snRNA was pulled down mainly by full length (FL) SART3 and C-SART3, confirming the essential role of the C-terminus in U6 snRNP binding. Low but still detectable U6 snRNA signal in N-SART3 sample possibly reflects residual di-snRNP particles bound by TPR domain through hPrp3 protein. Similarly, U4 snRNA can

be pulled down as a part of di-snRNP by C-SART3 interacting with U6 snRNP. However, N-SART3 co-immunoprecipitated higher amounts of U4 snRNA than can be explained by its presence in di-snRNP, suggesting an additional, U6-independent interaction with SART3.

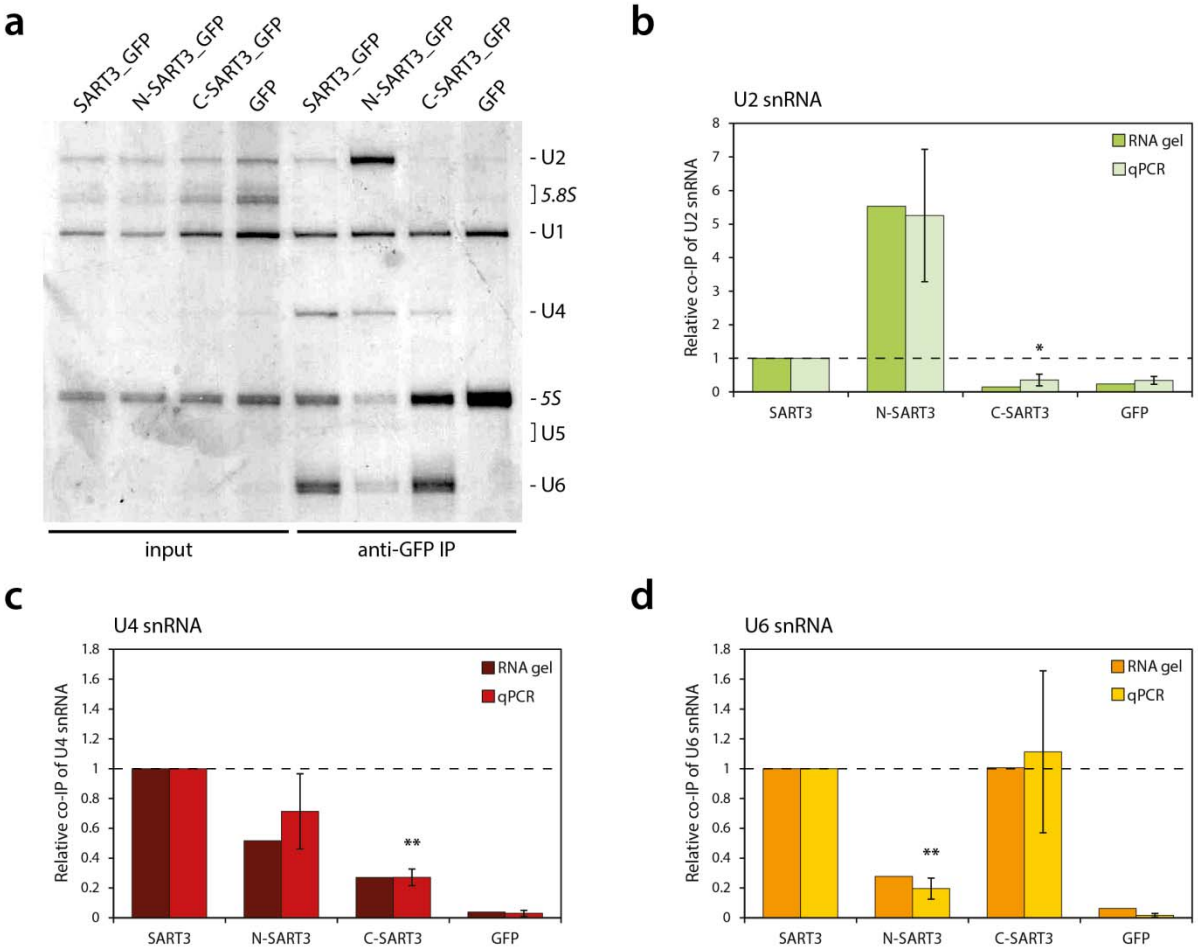


**Figure 5.2. SART3 pulls down U2 snRNA in addition to U4 and U6 snRNAs.**

(a) snRNAs co-precipitated with SART3 constructs were resolved on denaturating polyacrylamide gel and silver stained. The experiment revealed a strong interaction between U2 snRNA and N-SART3. Positions of individual U snRNAs are marked, rRNAs are indicated in italics.

(b) U2 snRNA pull down by N-SART3 was confirmed by quantitative RT-PCR.

(c) U4 snRNA interacts in particular with FL SART3 and partially also with N-SART3.

(d) U6 snRNA interacts with FL SART3 and C-SART3.

(b-d) Quantitative PCR data from three independent experiments were normalized to input values. The average values together with SEM are shown. The significance of N-SART3 and C-SART3 differences in respect to FL SART3 was assayed by t-test; * $p \leq 0.05$ and ** $p \leq 0.01$.

Strikingly, U2 snRNA which is supposed not to interact with SART3 at all was strongly co-purified with N-SART3. Small amount of U2 snRNA co-precipitated also with FL SART3. These data indicate presence of a novel interaction between SART3 N-terminus and U2 snRNP. To support the results from the RNA gel, we further examined co-precipitated RNA by quantitative RT-PCR; three independent biological replicas were done (Fig. 5.2b-d). Quantitative PCR yielded data consistent with the RNA gel.

## 5.2 SART3 associates specifically with immature U2 particles

We decided to focus on the interaction between SART3 and U2 snRNP and analyzed it in more details. As described above in chapter 2.3.2, U2 snRNP undergoes a complicated assembly pathway and it can generally be found in three distinct maturation stages. First, there is the immature 12S particle that contains, in addition to U2 snRNA, only Sm ring and U2A'/U2B" heterodimer. Second, the 15S particle which is still incomplete but contains SF3b proteins. And finally, the mature 17S particle consisting of Sm ring, U2A'/U2B" and SF3b and SF3a complexes (Fig. 2.5). We therefore investigated which U2 snRNP particle SART3 interacts with.

We transfected SART3_GFP, N-SART3_GFP and C-SART3_GFP constructs into HeLa cells, performed IP and examined levels of co-immunoprecipitated U2-specific proteins. In addition to SART3 constructs, U2A'_GFP and empty pEGFP-N1 vector (marked as GFP) were used and served as positive and negative controls, respectively. U2A'_GFP pulled down all the tested U2 proteins, namely U2B", SF3b49 and SF3a60, indicating the proper incorporation of GFP-tagged U2A' into U2 snRNP particle (Fig. 5.3). Co-precipitation between individual U2-specific proteins showed that U2 snRNPs are not disrupted during the IP experiment.

From the three tested U2 proteins, N-SART3 efficiently pulled down only U2B" protein. There was only a very weak interaction with SF3b49 and no co-immunoprecipitation (co-IP) of SF3a60 was observed. These data suggest that N-SART3 associates preferentially with immature 12S particle which lacks SF3a and SF3b complexes. It can possibly interact with 15S particle as well but the interaction seems to

be weak or unstable. Interestingly, C-SART3 weakly co-precipitated SF3b49 even though almost no U2B" was pulled down (Fig. 5.3) indicating that there might be an additional interaction between SF3b49 protein or the whole SF3b complex and C-terminal part of SART3.

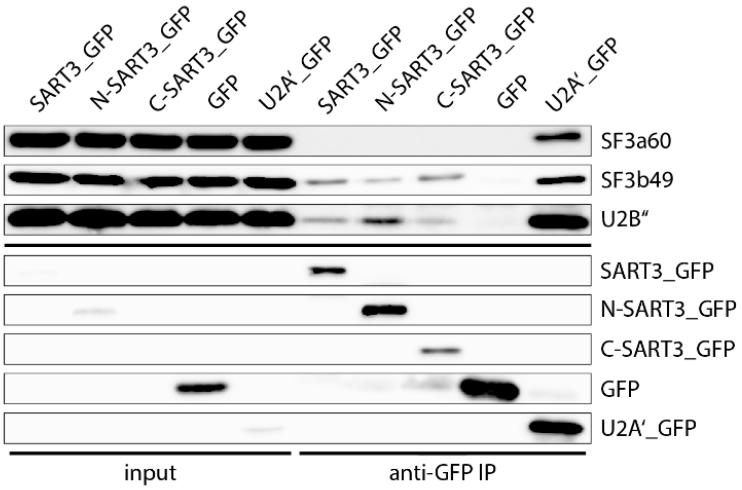


**Figure 5.3. N-SART3 interacts with immature U2 snRNPs.**

Cells were transfected by three different SART3 constructs and GFP and U2A'_GFP which served as negative and positive control, respectively. Anti-GFP immunoprecipitation was then performed and the co-IP of U2-specific proteins was analyzed on Western blot. N-SART3 pulled down mainly U2B" indicating the interaction with the immature 12S U2 particle.

## 5.3 Recombinant N-SART3 pulls down U2 snRNP *in vitro*

To further investigate specificity of U2 snRNP binding by N-SART3, we decided to prepare recombinant N-SART3 protein, purify it and perform an *in vitro* pull down assay. In previous study, Medenbach et al (2004) showed that SART3 mutant containing amino acids 2 to 688 can be efficiently purified. We therefore prepared the similar N-terminal SART3 fragment which lacked the last 15 amino acids compared to N-SART_GFP construct used in IP experiments above. The desired SART3 fragment was cloned using the Gateway Technology (Invitrogen) from SART3_GFP into pDEST_MM322 destination vector (Fig. 4.1) suitable for expression in mammalian cells. The resulting protein was fused at the N-terminus to Twin-Strep, FLAG and Halo tags through a short linker containing TEV protease cleavage site (Fig. 5.4a).

We termed this construct FLAG_N-SART3 and expressed it in human embryonic kidney 293T (HEK293T) cells which are suitable for large-scale protein expression. The recombinant protein was then affinity purified via Twin-Strep tag which composes of two Strep-tag II sequences connected by a short linker and provides more efficient protein purification than a single tag (Schmidt et al, 2013). As shown in Fig. 5.4b, we successfully purified sufficient amount of FLAG_N-SART3 for *in vitro* pull down assay in the first elution fraction.

For the *in vitro* pull down, we incubated purified N-SART3 with cell lysates prepared from untransfected HeLa cells and then immunoprecipitated the protein via FLAG tag. FLAG_N-SART3 pulled down low but detectable amount of di-snRNP-specific hPrp31 protein, indicating that the protein was folded properly and is able to partially interact with U4/U6 snRNP even though it lacks the C-terminus. In addition, N-SART3 also co-precipitated U2B" protein (Fig. 5.4c).

To confirm the specific binding of immature U2 snRNPs, we analyzed co-precipitation of U2-specific SF3a60, SF3b49 and U2B" proteins with recombinant FLAG_N-SART3. Similarly to the *in vivo* experiment, N-SART3 pulled down only U2B" (Fig. 5.4d). In this case, no SF3b49 protein was detected, however it may be caused by a slight decrease of the overall level of U2 snRNP pull down. Together these data show that N-SART3 functionality does not depend on the tag position and that purified N-SART3 binds U2 snRNP in *in vitro* conditions similarly to plasmid-expressed N-SART3 *in vivo*.

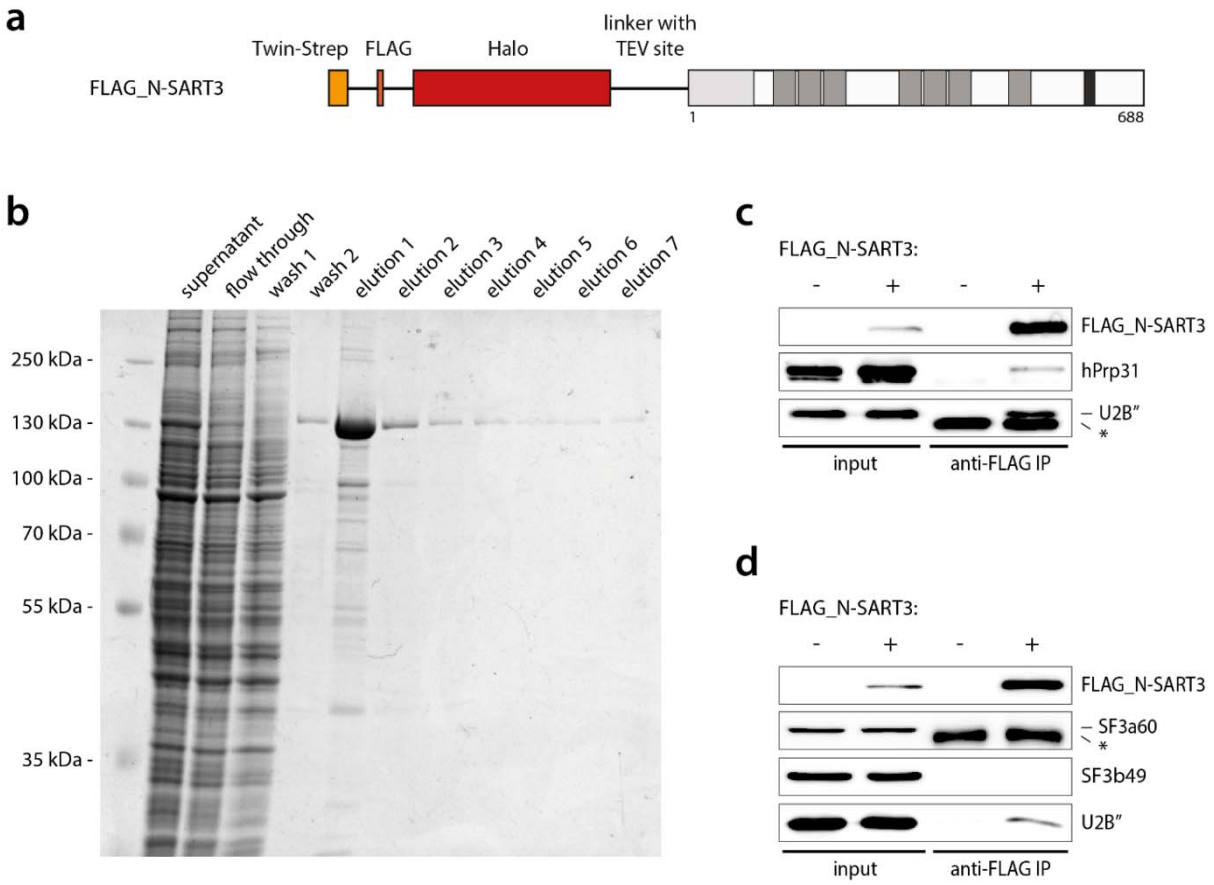


**Figure 5.4. Recombinant SART3 interacts with U2 snRNP.**

(a) Domain organization of FLAG_N-SART3 construct. N-SART3 (amino acids 1 to 688) was fused to Twin-Strep, FLAG and Halo tags through a short linker containing TEV protease cleavage site.

(b) Individual fractions obtained during affinity purification of N-SART3 protein. The purification was performed via the Twin-Strep tag. In the elution 1 fraction, sufficient amount of FLAG_N-SART3 was isolated.

(c) Purified N-SART3 was incubated with cell lysates and then pulled down using mouse anti-FLAG antibody. It co-purified with di-snRNP-specific hPrp31 and U2-specific U2B" proteins. Unspecific detection of IgG is marked by *.

(d) After incubation with cell lysates, recombinant N-SART3 was pulled down using rabbit anti-FLAG antibody. From all analyzed U2-specific proteins, N-SART3 co-precipitated only U2B" protein. Unspecific detection of IgG is marked by *.

## 5.4 SART3 interaction with snRNPs is mediated by Sm proteins

Since both SART3 and U2 snRNP are highly accumulated in nuclear Cajal bodies and both of them can interact with CB scaffolding protein coilin, we hypothesized that the U2-SART3 interaction is indirect and is mediated by coilin. SART3 binding to coilin is mediated by its N-terminal E domain and N-SART3 pulls down coilin distinctly more than full length protein (data not shown, see Novotný et al, 2015) . Amounts of co-precipitated coilin and U2 snRNA thus correlate. To confirm or rule out the involvement of coilin in U2-SART3 interaction, we knocked coilin down by siRNA and analyzed co-IP of U2B" protein which was used as U2 snRNP marker. However, we did not observe any difference in U2B" co-precipitation between coilin-depleted and control cells (Fig. 5.5), indicating that N-SART3 interaction with U2 snRNP occurs independently on coilin.

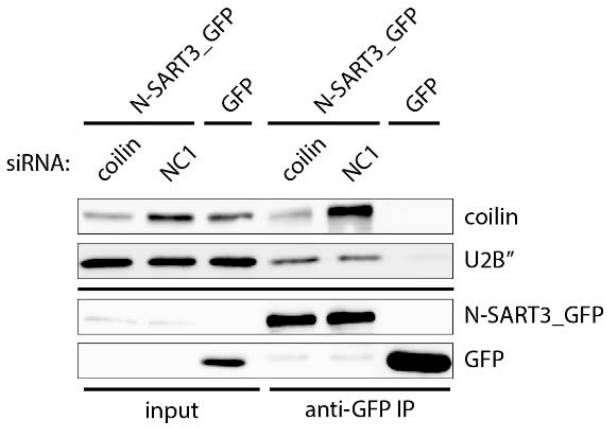


**Figure 5.5. U2-SART3 interaction is coilin independent.**
HeLa cells were treated with negative control or anti-coilin siRNAs and then transfected by N-SART3. U2 snRNP co-IP with N-SART3 was detected via U2B" protein. GFP was used as a negative control.

We then aimed to investigate which U2 snRNP components mediate the U2-SART3 interaction. SART3 associates preferentially with the immature 12S U2 particles, making the list of potential candidates restricted to U2 snRNA, Sm ring proteins and U2A'/U2B" dimer. We therefore treated the cell lysates with RNase A to cause the

snRNPs disassembly and probed for direct protein-protein interactions. After the treatment, we observed complete loss of U2B" co-precipitation with both SART3 and N-SART3. However, it had no considerable effect on SmB/B' pull down by both SART3 constructs (Fig. 5.6), raising an interesting possibility of Sm proteins mediating the U2-SART3 interaction.

Sm ring is assembled around the snRNA in a complicated multi-step process and then is believed to stay stably bound throughout the whole spliceosomal cycle. It is therefore likely that the snRNA degradation can cause disintegration of the Sm ring structure. To test this hypothesis, we expressed, aside from SART3_GFP and N-SART3_GFP, also GFP_SmD1 and examined its interaction with SmB/B' which is located next to it in the Sm ring (Pomeranz Krummel et al, 2009; Fig. 2.1a). Under normal conditions SmD1 co-precipitates SmB/B' as well as U2B", suggesting efficient incorporation of GFP-tagged SmD1 into snRNPs. However, neither U2B" nor SmB/B' are pulled down after the RNase A treatment (Fig. 5.6). From this we conclude that the Sm ring is disintegrated in the absence of U2 snRNA and that SART3 does not require intact ring structure for the interaction with SmB/B'.

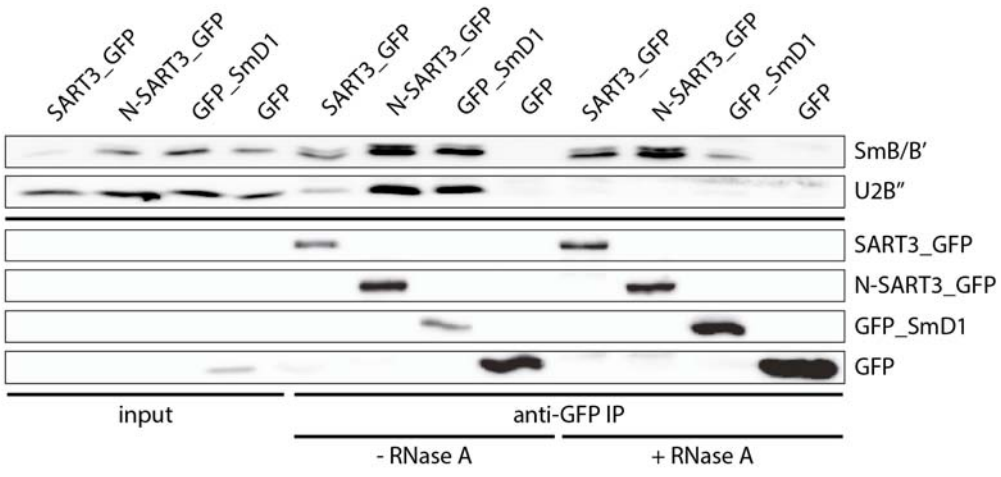


**Figure 5.6. U2 snRNP association with SART3 is mediated by its Sm ring.**

Cell lysates were split into two halves and one was treated with RNase A. Immunoprecipitation was performed afterwards. Due to technical limitations, only inputs from non-treated samples are shown on the Western blot. We however did not observe any difference in inputs before and after the treatment. Contrary to U2B", SmB/B' co-IP with SART3 was preserved even in the absence of snRNA.

To test the possible role of Sm proteins in the U2-SART3 interaction, we knocked down one of these proteins, SmB/B'. We assumed that depletion of an individual Sm ring component prevents the other Sm proteins from forming the ring structure and leads to their degradation, which we confirmed afterwards, as shown below in Fig. 5.8. Although SmB/B' knock down did not affect the total U2B" amount in cells, U2B" co-precipitation with SART3 and N-SART3 was significantly lowered compared to negative control (NC) siRNA transfected cells (Fig. 5.7a). We also examined the influence of SmB/B' depletion on U2B" pull down by purified FLAG_N-SART3 protein and we observed a very similar effect (Fig. 5.7b).

To further confirm these data, we decided to investigate the level of U2 snRNA co-IP. We thus prepared a silver stained RNA gel and observed that the interaction between SART3 and U2 snRNA is indeed SmB/B'-dependent (Fig. 5.7c). Interestingly, the same was true also for all the other U snRNAs. Co-precipitation of U2, U4, U5 and U6 snRNAs in control versus SmB/B' depleted cells was measured by quantitative RT-PCR as well; the average values from three independent biological replicas are shown in Fig. 5.7d-g. It should be noted here that SmB/B' knock down resulted in partial degradation of snRNA molecules, as visible at inputs in Fig. 5.7c. However, all the co-IP values from quantitative PCR were normalized to inputs and snRNA decay thus did not contribute to the differences between knocked down and control samples. Furthermore, snRNA co-precipitation with SART3 and N-SART3 in NC siRNA treated cells is consistent with non-treated cells in Fig. 5.2 for both the RNA gel and qPCR data.

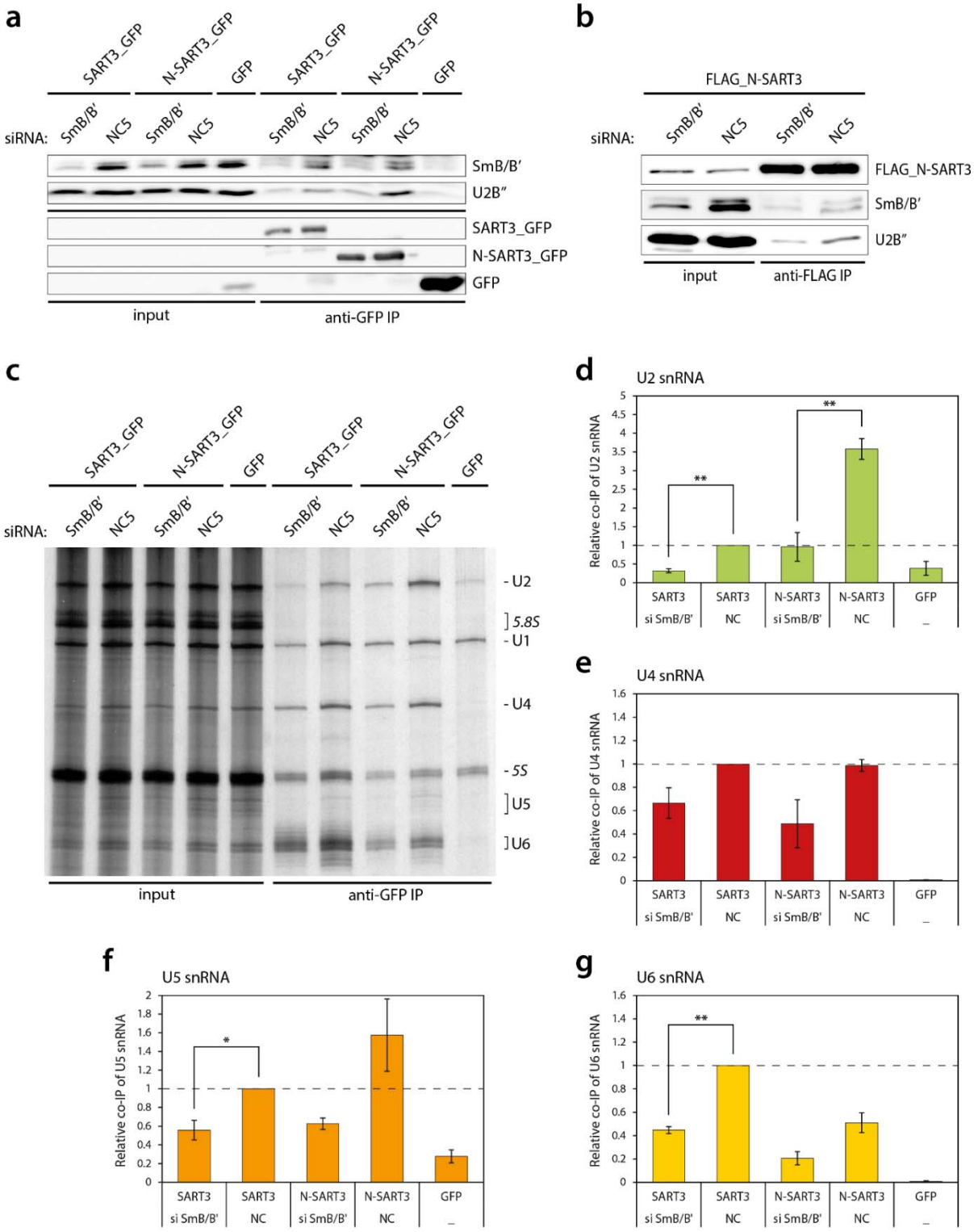


**Figure 5.7. SmB/B' depletion causes decrease of snRNA co-IP with SART3.**

(a) SART3_GFP and N-SART3_GFP were immunoprecipitated from NC and anti-SmB/B' siRNA treated cells. GFP was used as a negative control. Efficiency of SmB/B' knock down was tested by anti-SmB/B' antibody, co-precipitation of U2 snRNP by anti-U2B" antibody. (*Continued on next page*)

(b) FLAG_N-SART3 was incubated with NC or anti-SmB/B' siRNA treated cell lysates and then pulled down using rabbit anti-FLAG antibody.
(c) snRNAs isolated from the same samples as proteins in (a) were resolved on denaturating polyacrylamide gel and silver stained. Positions of individual U snRNAs are marked, rRNAs are indicated in italics.
(d-g) FL SART3 and N-SART3 pulled down U2, U4, U5 and U6 snRNAs in SmB/B'-dependent manner. Quantitative RT-PCR data from three independent experiments were normalized to input values. The average values together with SEM are shown. The significance of differences between SmB/B' knocked down samples and NC siRNA treated samples was assayed by t-test; * $p \leq 0.05$ and ** $p \leq 0.01$.

To further support the results, we repeated this experiment, but this time we knocked down SmG protein instead of SmB/B'. Unfortunately, we did not have anti-SmG antibody to detect siRNA efficiency. We used however anti-SmB/B' antibody and observed similar decrease of SmB/B' as if SmB/B' was targeted directly (Fig. 5.8a). It suggests that anti-SmG siRNA functioned properly and also confirms our assumption that depletion of one Sm protein leads to degradation of the others. As expected, SmG knock down resulted in lowered co-precipitation of U2B" (Fig. 5.8a) and snRNAs (Fig. 5.8b) with SART3, as well as with N-SART3. The decrease of snRNP-SART3 co-IP was comparable to that caused by anti-SmB/B' siRNA, suggesting that neither of the siRNAs had off-target effects. Together, these data show a crucial role of Sm proteins in mediation of U2-SART3 interaction.

Our data strongly indicate that not only U2 but also other snRNPs may use the Sm ring for the interaction with SART3. We thus decided to examine the specificity of U5 snRNP pull down by SART3. We took advantage of immunoprecipitation experiment in combination with RNase A treatment, and analyzed co-precipitation of two U5-specific proteins, hPrp8 and hPrp6. Consistently with the SmB/B'-dependent interaction between U5 snRNA and SART3, both U5-specific proteins were pulled down by SART3 and N-SART3 in normal conditions but not after the treatment (Fig. 5.9). Notice that we used samples from the same experiment as in Fig. 5.6 for the Western blot.

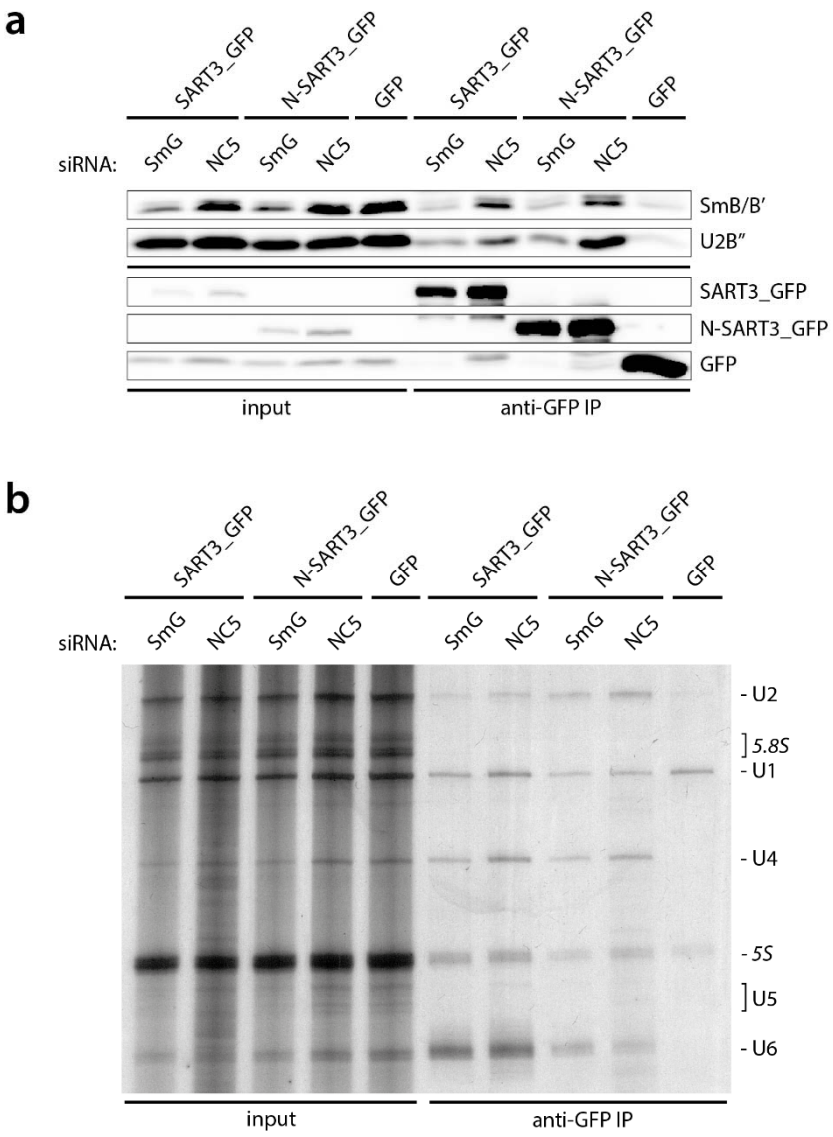


**Figure 5.8. SmG depletion causes degradation of SmB/B' and decrease of snRNA-SART3 co-IP.**

(a) SART3_GFP and N-SART3_GFP were immunoprecipitated from NC and anti-SmG siRNA treated cells. Efficiency of SmG knock down was tested indirectly by anti-SmB/B' antibody, co-precipitation of U2 snRNP by anti-U2B" antibody.

(b) snRNAs isolated from the same samples as proteins in (a) were resolved on denaturating polyacrylamide gel and silver stained. Positions of individual U snRNAs are marked, rRNAs are indicated in italics. The effect of SmG depletion on snRNA co-IP with SART3 was similar to SmB/B' knock down.

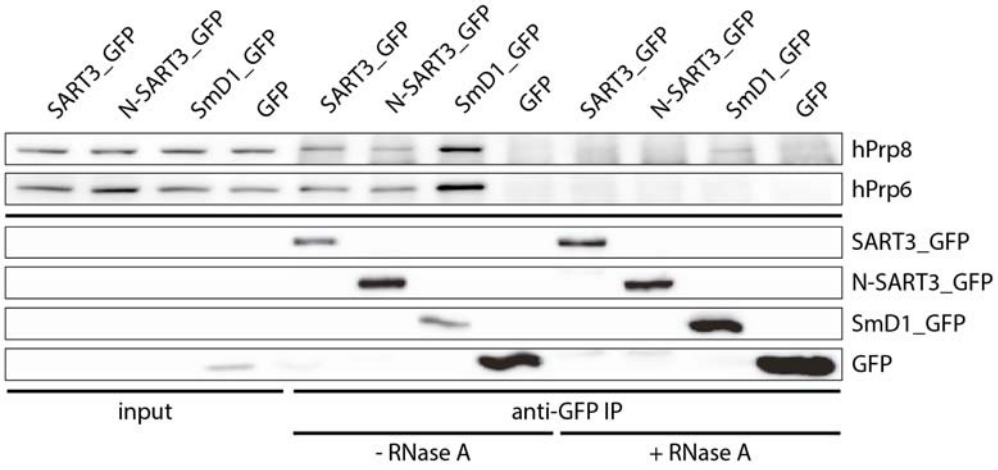


**Figure 5.9. Interaction between SART3 and U5 snRNP proteins is RNA-dependent.**
Cell lysates were treated with RNase A and then the immunoprecipitation was performed. Both U5-specific hPrp8 and hPrp6 proteins were pulled down by SART3, N-SART3 and SmD1 in normal conditions but not after the RNase A treatment.

## 5.5 Searching for the function of the U2-SART3 interaction

The results shown above provided us with a strong evidence for an interaction between SART3 protein and the immature U2 snRNP. However, the functional role of the interaction remained unclear. In an attempt to elucidate a possible function of SART3 in U2 snRNP biogenesis, we decided to investigate whether U2 snRNP localization in the cell depends on SART3.

First, we over-expressed full length SART3_GFP and N-SART3_GFP and compared their nuclear distribution. Both of them exhibited highly similar localization patterns which did not differ from endogenous SART3 (Fig. 5.10, compare with Fig. 5.11a). Both full length SART3 and N-SART3 were uniformly distributed throughout the nucleoplasm, excluded from nucleoli and accumulated in Cajal bodies, which were marked by anti-coilin immunostaining. We detected U2 snRNA using *in situ* hybridization technique with a fluorescent probe. U2 snRNA displayed a characteristic

speckled pattern and accumulated in Cajal bodies, as previously described (Matera & Ward, 1993). Neither SART3_GFP, N-SART3_GFP, nor GFP itself influenced U2 snRNA distribution as compared with untransfected cells (Fig. 5.10).

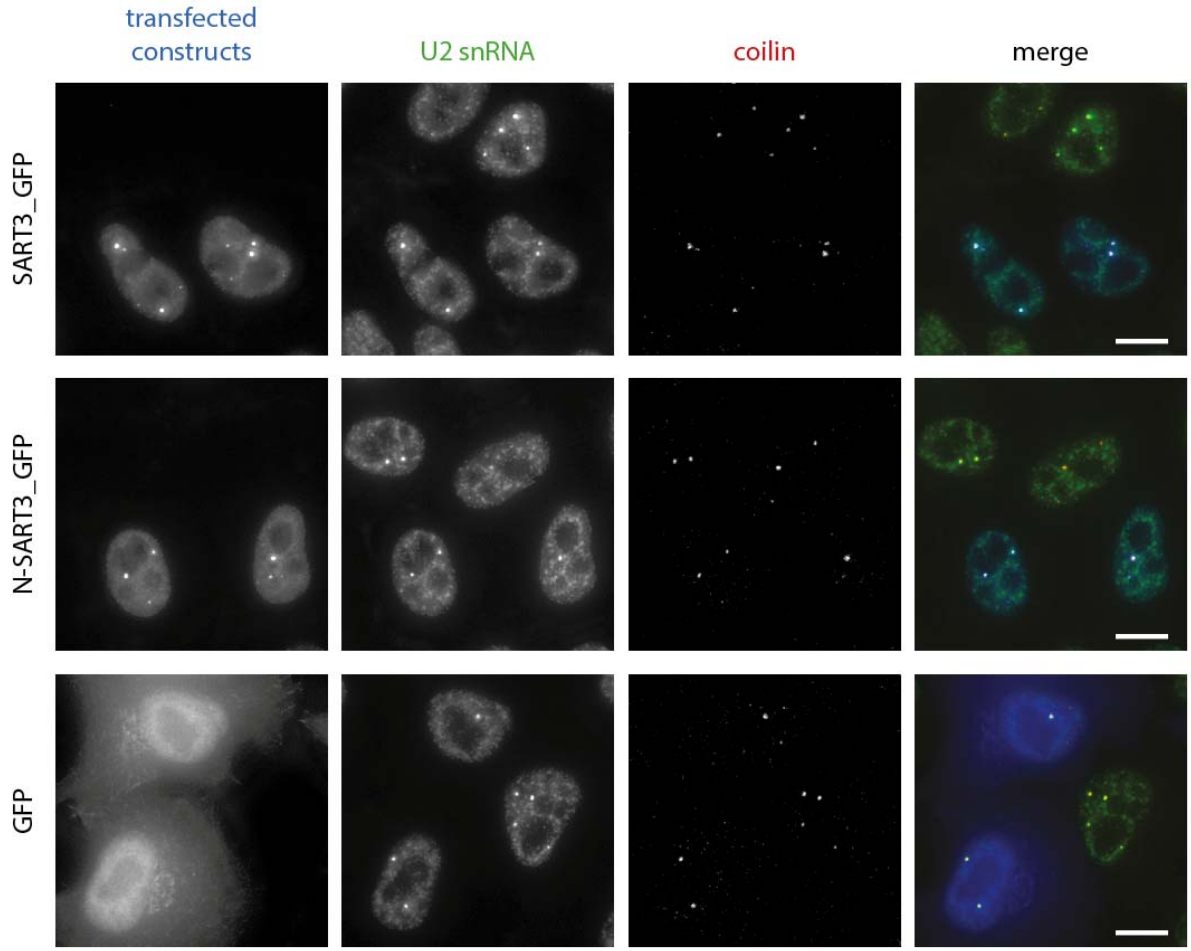


**Figure 5.10. Over-expressed SART3 and N-SART3 do not cause impaired cellular distribution of U2 snRNA.**

SART3 and N-SART3 constructs (in blue) were properly localized in the nucleus and their distribution did not differ from the endogenous SART3 (compare with Fig. 5.11a). U2 snRNA was visualized by *in situ* hybridization (in green) and its distribution remained unchanged in all the transfected as well as untransfected cells. Coilin was detected by indirect immunofluorescence (in red) and served as a marker of Cajal bodies. The merge images of all three color channels are shown. Scale bars represent 10 μm.

We have previously shown that SART3 bridges tri-snRNP components to coilin and is thus important for U4, U5 and U6 snRNPs accumulation in Cajal bodies (Novotný et al, 2015). Interacting specifically with immature U2 snRNP particles, SART3 could possibly play a similar role in the U2 snRNP assembly. To test this hypothesis, we performed a knock down experiment with anti-SART3 and anti-SF3a60 siRNAs and their combination. Depletion of U2-specific SF3a60 protein is known to prevent 17S particle assembly and cause increased 12S U2 snRNP accumulation in CBs (Tanackovic & Krämer, 2005). We observed this effect in our SF3a60-knocked down cells as well. U2 snRNA labeled by FISH probe was strongly reduced in nuclear speckles but enriched in Cajal bodies (Fig. 5.11a), suggesting an existence of a CB-sequestration mechanism for incomplete U2 snRNPs. However, SART3 depletion had, contrary to SF3a60, no visible effect on U2 snRNA localization in both negative control and anti-SF3a60 siRNA treated cells (Fig. 5.11a).

To quantify these results, we performed a high-throughput microscopy analysis using Scan^R automated acquisition system (Olympus). In each sample, two to three thousands of cells were imaged and the intensity of U2 snRNA FISH probe was measured in both the nucleoplasm and CBs. The ratio of fluorescence signal in CBs versus the nucleoplasm is plotted in Fig. 5.11b. This experiment was done in two independent biological replicas and results from both exhibited the same trend (data from the second replica not shown), indicating that SART3 is not crucial for U2 snRNP anchoring in Cajal bodies.

Knowing that SART3 is not essential for the sequestration of U2 particles in CBs, we hypothesized that it might play a role in enhancing U2 snRNP assembly. To test this option, we knocked down SART3 and immunoprecipitated U2 snRNPs from cell lysates through U2A'_GFP. We deduced the efficiency of U2 snRNP assembly from levels of U2-specific proteins co-precipitated with the core U2A' protein. However, we did not reveal any significant difference between SART3-depleted and control cells (Fig. 5.12). That indicates that either SART3 does not influence U2 biogenesis or our approach was not sensitive enough to detect a difference.

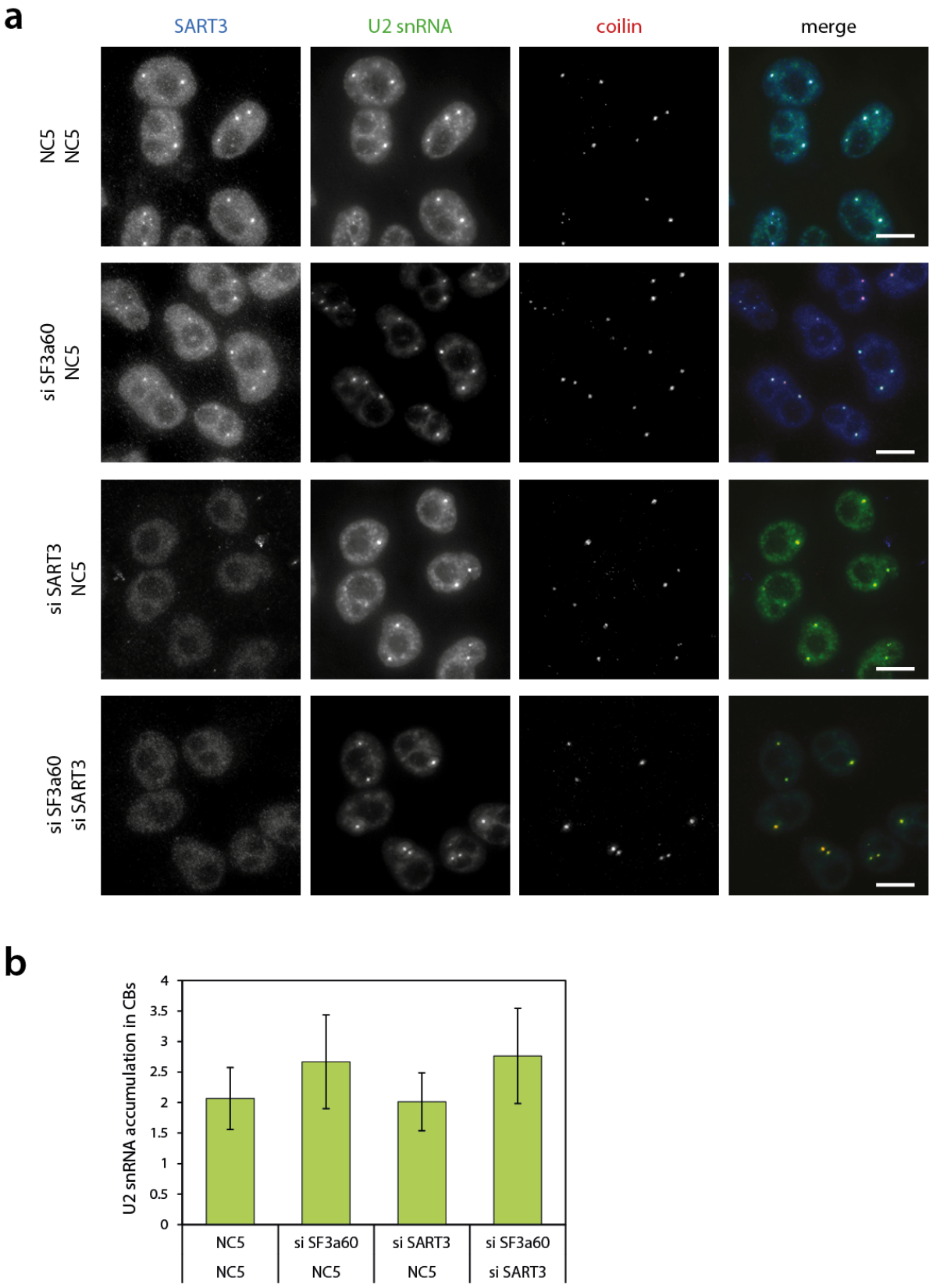


**Figure 5.11. U2 snRNA accumulates in Cajal bodies in SART3-independent manner.**

(*Legend on next page*)

(a) Single knock downs of SF3a60 and SART3 in combination with NC siRNA, as well as double knock down of both were performed. Cells transfected by NC siRNA only served as a positive control. In all samples the same final concentration of siRNA was used. U2 snRNA was detected by *in situ* hybridization (in green), SART3 and coilin were visualized by indirect immunostaining (in blue and red, respectively). The merge images of all three color channels are shown. Scale bars represent 10 µm.

(b) U2 snRNA accumulation in Cajal bodies was measured using high-throughput microscopy (Scan^R acquisition system). Two to three thousands of cells were analyzed in each sample. The nucleus and Cajal bodies were automatically detected using DAPI and coilin staining, respectively, and the U2 snRNA probe signal was measured in both compartments. Average values from one experiment are shown together with standard deviation.

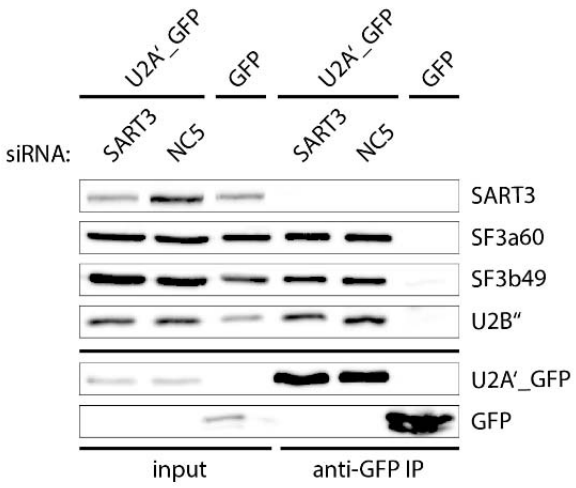


**Figure 5.12. SART3 depletion does not influence U2 snRNP assembly.**

U2 snRNPs were immunoprecipitated via U2A'_GFP from SART3-depleted and control cells. No significant difference in co-IP of three U2-specific proteins was detected after the anti-SART3 siRNA treatment.

# 6 Discussion

SART3 is an essential di-snRNP assembly factor. It interacts with U6 snRNP and di-snRNP-specific hPrp3 protein and helps U6 to associate with U4 snRNP in nuclear Cajal bodies. Here, we focused on the function of individual SART3 domains with the aim to elucidate a mode of SART3 interactions with spliceosomal snRNPs. For this purpose, we used three different SART3 constructs in our study: full length protein (amino acids 1 to 963), N-terminal part (1 to 703) and C-terminal part (580 to 963); both mutants partially overlapped in the middle region that contains nuclear localization signal. Performing a series of immunoprecipitation experiments, we observed a strong interaction between N-SART3 and U2 snRNP and provided evidence that the N-terminal half of SART3 binds preferentially immature 12S U2 particles. We then showed that U2 snRNP interacts with both ectopically expressed and recombinant N-SART3, providing an additional confirmation of the results.

We have several reasons to propose that the U2-SART3 interaction is mediated by the Sm proteins. First, SmB/B' co-precipitates with SART3 even after depletion of snRNAs. Second, the U2-SART3 interaction depends on SmB/B' and SmG proteins. And third, N-terminal part of SART3 interacts, aside from U2, also with U4 and U5 snRNPs and the only components shared by all the three particles are Sm proteins. Moreover, U4 and U5 snRNPs associate with SART3 in SmB/B'-dependent manner.

We showed that Sm-class snRNPs interact specifically with the N-terminal part of SART3 protein which is composed of two different regions. At the very N-terminus there is the E domain that consists of the first 100 amino acids. It is characterized by a high rate of glutamic acid, but no secondary structure motif has been predicted in this region. The E domain is followed by the long stretch of half-a-TPR motifs. Tetratricopeptide repeat (TPR) is an evolutionarily conserved motif that can mediate a variety of protein-protein interactions. TPR-containing proteins are often involved in cell cycle regulation, protein transport, transcriptional control or chaperone complexes (Blatch & Lässle, 1999).

A single TPR motif consists of 34 amino acids, eight of which are more conserved than the others and generate a characteristic consensus sequence. On the secondary structure level, TPR motif packs into two anti-parallel α-helices connected by

a short linker (Blatch & Lässle, 1999; D'Andrea & Regan, 2003). Half-a-TPR (HAT) repeat is a variant of TPR, but it is much less ubiquitous compared to TPR and has been found exclusively in proteins involved in RNA processing. Although lacking two amino acids from the TPR consensus (Preker & Keller, 1998), HAT secondary structure strongly resembles the TPR repeat (Bai et al, 2007; Champion et al, 2009). Both TPR and HAT motifs are often arranged in tandem repeats, and their number in different proteins usually varies between 3 and 16 (D'Andrea & Regan, 2003; Preker & Keller, 1998). Anti-parallel helices of individual TPR or HAT motifs in these tandems closely neighbor the helices of preceding and following motifs and together form a right-handed superhelical structure which provides an ideal protein-protein interaction platform (Bai et al, 2007; D'Andrea & Regan, 2003).

Interestingly, clusters of three consecutive TPR repeats are most common and their surface is sufficient for the binding of a target protein. Some proteins with two or more three-TPR clusters are thus able to accommodate two different protein targets simultaneously as, for example, in the case of Hsp70/Hsp90 organizing protein (Hop) which binds both Hsp70 and Hsp90 and facilitates their assembly into a multiprotein complex. Sometimes, a single TPR repeat is present in addition to the clusters and enhances stability of protein interactions (L. Regan, personal communication; D'Andrea & Regan, 2003) .

This organization of HAT motifs is present also in SART3 protein, which contains two three-HAT clusters and one individual repeat. So far, di-snRNP-specific hPrp3 is the only protein identified to interact with the TPR domain of SART3. However, the first HAT cluster is entirely sufficient for hPrp3 binding (Medenbach et al, 2004), raising a possibility that there might be another target protein able to bind to the second HAT cluster of SART3 TPR domain. We therefore considered an option that Sm proteins interact with one half of SART3 TPR domain while the other half of the domain binds hPrp3.

Unfortunately, there is no TPR-binding consensus sequence known, which would allow easy identification of TPR-target proteins. Nevertheless, if both hPrp3 and Sm proteins interacted with the TPR domain of SART3, they could conceivably share a similar SART3-binding region. To support this idea, we decided to compare sequences of all seven Sm proteins with the short C-terminal region of hPrp3 (amino acids 416 to 550) which has been shown to associate with SART3 (Medenbach et al, 2004). We used NCBI

protein-protein BLAST tool, and strikingly, we revealed 28 % identity and 59 % similarity between amino acids 438 to 469 of hPrp3 and 41 to 70 of SmB (Fig. 6.1).

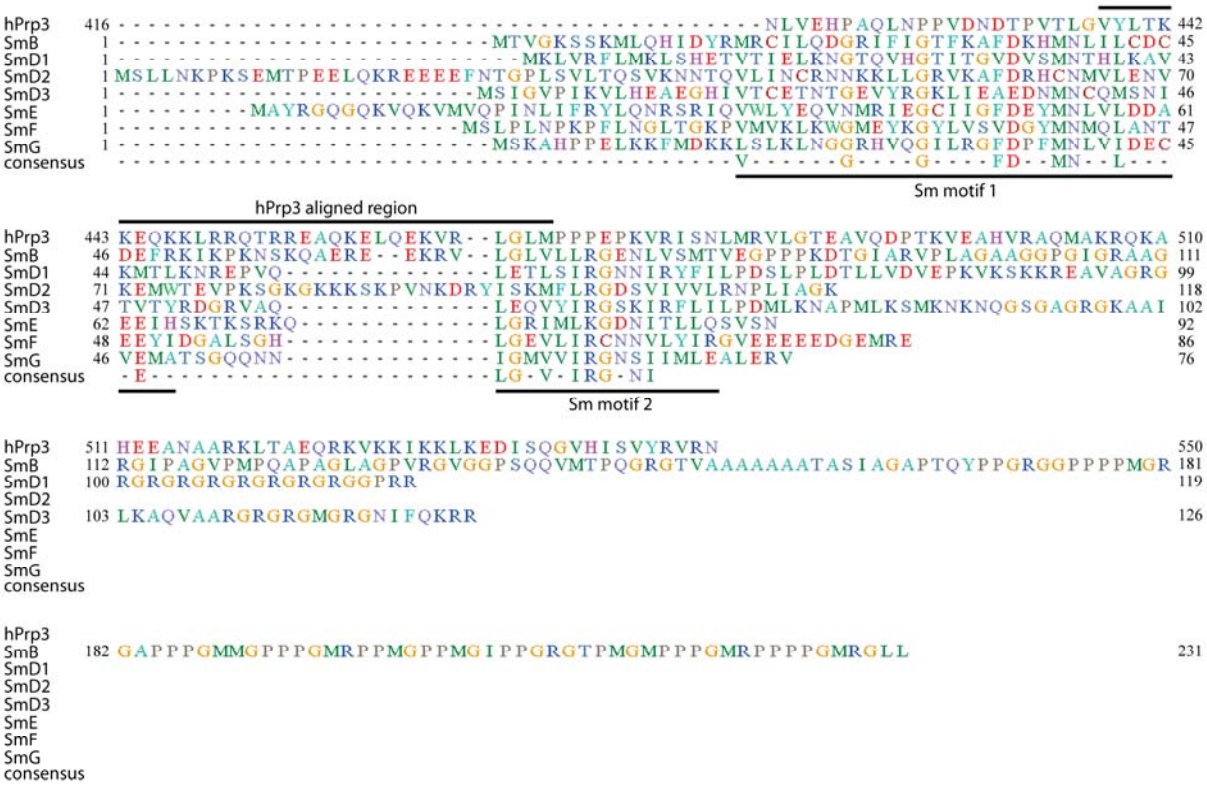


**Figure 6.1. Sequence alignment of hPrp3 (amino acids 416 to 550) and Sm proteins.**
Sm proteins were aligned according to Hermann et al, 1995. Sm consensus sequence is shown in the last row and both conserved Sm motifs are indicated by solid lines below the sequences. The region 416-550 of hPrp3 was aligned to all Sm proteins using NCBI protein-protein BLAST tool. From all Sm proteins, only SmB exhibited a significant similarity to hPrp3; the successfully aligned region is indicated by a solid line above the sequences. Individual amino acids are distinguished by different colors; amino acids that exhibit similar properties are in the same color. Gaps are marked by dash. Complete sequences of Sm proteins are shown.

Although the aligned region partially overlaps the conserved Sm motifs, most of it lays in the variable linker between both motifs. Because individual Sm proteins significantly differ in the linker sequence, no other Sm protein except for SmB has been detected by the BLAST tool to share a homology region with the hPrp3 fragment. However, we noticed that the hPrp3 region 466 to 469 (LGLM sequence) is identical to

the first four amino acids of the Sm motif 2 and that these amino acids are more or less conserved in all the Sm proteins. This suggests that there might be a sequence conservation between hPrp3 and the Sm proteins. An experimental evidence is however needed to confirm whether SmB or other Sm proteins are capable to bind the TPR domain of SART3.

It is not clear from our data which of the seven Sm proteins is/are responsible for the interaction with SART3. Some of our findings indicate that SmB might be involved in the interaction; it exhibits a sequence similarity with another SART3-binding partner, hPrp3, and it stayed bound to SART3 after the RNase A treatment even though the Sm ring was destabilized. We do not know, however, whether the Sm ring disintegrated to individual Sm proteins or bigger subcomplexes, such as B/D3, D1/D2 and E/F/G which serves as building blocks during the Sm ring assembly. On the other hand, SmB/B' did not co-precipitate with SART3 after the SmG depletion which could have two possible causes. Either we did not detect SmB/B' due to its overall decrease or the Sm-SART3 interaction is mediated by another Sm protein.

Our results imply that the interaction between SART3 and Sm proteins is not restricted only to U2 snRNP but occurs also in case of U4 and U5 snRNPs. This might mean that the role of SART3 in snRNP biogenesis and recycling is much more extensive than previously thought. The suggestion that SART3 may assist the recycling of U2 and U5 snRNPs is consistent with a study done by Trede et al, who reported a zebrafish mutant lacking a C-terminal part of SART3, i.e. RRM and CT domains. Mutant embryos suffered multiple organ-specific defects and died within 7 to 8 days postfertilization. Interestingly, a lot of genes encoding snRNP-specific proteins were up-regulated in mutant embryos, suggesting that the cells attempted to compensate a defect in snRNP recycling. Among the identified up-regulated genes, there were Sm and LSm proteins, di-snRNP and tri-snRNP-specific proteins and surprisingly also half of all U5-specific proteins and U2-specific SF3a and SF3b proteins. In contrast, expression levels of all U1-specific proteins remained unchanged (Trede et al, 2007).

Although we have provided several lines of evidence showing that SART3 interacts with U2 snRNP, we were not successful in determining the biological role of the interaction. To locate the interaction within the cell nucleus, we took advantage of Förster resonance energy transfer (FRET) approach. FRET allows detection of close interactions (up to 10 nm distance) between two fluorescently tagged proteins in

individual cellular compartments. Since we wanted to analyze U2-SART3 interaction specifically, we decided to measure the energy transfer between SART3 and U2B" as the U2 snRNP marker. However, the distance between both proteins was probably too high, so we did not detect any positive signal (data not shown). We then used fluorescence *in situ* hybridization and immunoprecipitation to test possible functions of SART3 in the U2 snRNP assembly. We hypothesized that SART3 may either target unassembled U2 particles to Cajal bodies or assist the U2 snRNP assembly. However, we did not confirm either of these options. The biological relevance of the U2-SART3 interaction thus remains to be elucidated.

In case of U4 and U5 snRNPs, the function of the Sm-SART3 interaction seems to be much clearer. We have recently shown that U4, U5 and U6 snRNPs accumulate in Cajal bodies in SART3-dependent manner, indicating that SART3 is important for CB targeting of these snRNP particles (Novotný et al, 2015). It is known that SART3 interacts with CB-scaffolding protein coilin via the N-terminal E domain and functions as a bridging protein between coilin and snRNPs (Novotný et al, 2015). It was also established that SART3 directly binds U6 snRNP via the C-terminal RRM and CT domains (Bell et al, 2002; Rader & Guthrie, 2002), but it is currently unknown how the interaction with U4 and U5 snRNPs is mediated. Here, we propose that SART3 interacts with Sm proteins of immature U4 and U5 snRNPs and targets thus these particles to Cajal bodies.

Interestingly, both immature 12S U2 snRNP and U4 snRNP are quite small particles that contain only two snRNP-specific proteins in addition to the Sm proteins. Moreover, the snRNP-specific proteins are always positioned on the other end of the snRNA molecule than the Sm ring. This organization provides a possible explanation how SART3 distinguish between immature and fully assembled particles. In the incomplete snRNP, Sm proteins might be exposed on the surface enough to be accessible for SART3, and then, during the assembly process in CBs, the ring might be covered by other proteins. The spatial organization of the mature particle thus would not allow SART3 binding and this would result in snRNP release from the Cajal body.

The Sm-SART3 interaction may work in a similar way also in case of U5 snRNP. The exact positions of U5-specific proteins within the snRNP are not known, but even though U5 is a huge particle which contains eight different snRNP-specific proteins, it is possible that the Sm ring remains accessible during the whole assembly process. This is

supported by the fact that SART3 pulls down U5 snRNA together with hPrp8 and hPrp6 proteins; while hPrp8 is the first or one of the first U5-specific proteins loaded on the snRNA (Novotný et al, 2015), hPrp6 is incorporated into U5 snRNP during later steps of the assembly and could reflect the complete U5 particle (Liu et al, 2006; Novotný et al, 2015). The immature U5 snRNP thus may be anchored by SART3 in the Cajal body during the assembly until it interacts with di-snRNP and forms the tri-snRNP particle, the conformation of which does not support SART3 binding. U4/U6·U5 snRNP is then released from the CB.

Having implicated SART3 in U4, U5 and U6 snRNPs targeting to CBs and knowing that SART3 binds specifically immature U2 snRNPs, we speculated that SART3 could play a crucial role also in CB targeting of U2 snRNP. However, we found out that incomplete U2 particles accumulate in CBs in SART3-independent manner. Contrary to tri-snRNP components, U2 snRNP thus must be sequestered in the CB by a different factor, which may or may not collaborate with SART3.

Despite the lack of information about the mechanism of U2 snRNP targeting to CBs, we propose that the interaction between Sm proteins and SART3 is essential for the targeting of U4 and U5 snRNPs. In Cajal bodies, U4/U6 snRNP is assembled with the assistance of SART3. We suggest that aside from the targeting function, SART3 interaction with the Sm ring of U4 snRNP is important also for the di-snRNP assembly itself. Since SART3 uses the C-terminal part of the molecule for the U6 snRNP binding, it could conceivably interact, at the same time, with U4 snRNP using the N-terminal part. SART3 would thus function as a chaperone, actively arranging both U4 and U6 snRNP to appropriate positions. Moreover, if our hypothesis is true and Sm proteins really interact with one of the three-HAT clusters of the TPR domain, the other one could simultaneously bind di-snRNP-specific hPrp3 protein and further enhance the di-snRNP assembly. Alternatively, SART3 might employ only one of both HAT clusters and hPrp3 would then have to replace Sm proteins in the binding site of SART3.

Taken together, we propose that SART3 interacts via the N-terminal TPR domain with spliceosomal Sm-class snRNPs and that the interaction is mediated by Sm proteins. Namely, we suggest that the interaction occurs with U2, U4 and U5 snRNPs; unfortunately, we do not have any data about U1 snRNP. We further suggest that the Sm-SART3 interaction represents a molecular mechanism how SART3 targets immature U4 and U5 particles to Cajal bodies. However, our results do not support this hypothesis in

case of U2 snRNP, indicating that U2 particles are targeted to CBs differently. Further experiments will be thus needed to confirm our speculations and to reveal the biological relevance of the U2-SART3 interaction.

# 7 Conclusions

Here, we report a novel interaction between di-snRNP assembly factor SART3 and Sm proteins. Sm proteins are bound around the snRNA in a heptameric ring structure and form the core of four out of five major spliceosomal snRNPs. SART3 is generally considered to be a U6-specific and di-snRNP-specific protein, we however revealed that it associates with immature U2 snRNP particles as well, and identified Sm proteins as a mediator of the interaction. Our data further imply that SART3 binds in the same manner also U4 and U5 snRNPs, suggesting existence of a common snRNP-SART3 binding mechanism. The Sm-SART3 interaction may provide a new insight into a molecular mechanism how SART3 targets immature U4 and U5 snRNPs to Cajal bodies, and we propose that the interaction plays a crucial role in the snRNP assembly process.

We have shown here that the region of SART3 responsible for the interaction with Sm proteins lays within the N-terminal part of the protein which is composed of glutamic-acid-rich domain and a long stretch of TPR motifs. In our further work, we intend to localize this region in more detail. In parallel, we plan to identify the particular components of the Sm ring which contribute to the interaction. Finally, we want to take advantage of *in vitro* systems and test the ability of SART3 to bind directly these Sm components as well as the whole spliceosomal snRNPs.

# References


Achsel T, Brahms H, Kastner B, Bachi A, Wilm M, Lührmann 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

Almeida F, Saffrich R, Ansorge W, Carmo-Fonseca M (1998) Microinjection of anti-coilin antibodies affects the structure of coiled bodies. *J Cell Biol* **142:** 899-912

Bai Y, Auperin TC, Chou CY, Chang GG, Manley JL, Tong L (2007) Crystal structure of murine CstF-77: dimeric association and implications for polyadenylation of mRNA precursors. *Mol Cell* **25:** 863-875

Behzadnia N, Golas MM, Hartmuth K, Sander B, Kastner B, Deckert J, Dube P, Will CL, Urlaub H, Stark H, Lührmann R (2007) Composition and three-dimensional EM structure of double affinity-purified, human prespliceosomal A complexes. *EMBO J* **26:** 1737-1748

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

Bessonov S, Anokhina M, Will CL, Urlaub H, Lührmann R (2008) Isolation of an active step I spliceosome and composition of its RNP core. *Nature* **452:** 846-850

Black DL, Pinto AL (1989) U5 small nuclear ribonucleoprotein: RNA structure analysis and ATP-dependent interaction with U4/U6. *Mol Biol Cell* **9:** 3350-3359

Blatch GL, Lässle M (1999) The tetratricopeptide repeat: a structural motif mediating protein-protein interactions. *Bioessays* **21:** 932-939

Boelens W, Scherly D, Beijer RP, Jansen EJ, Dathan NA, Mattaj IW, van Venrooij WJ (1991) A weak interaction between the U2A' protein and U2 snRNA helps to stabilize their complex with the U2B" protein. *Nucleic Acids Res* **19:** 455-460

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

Bouveret E, Rigaut G, Shevchenko A, Wilm M, Séraphin B (2000) A Sm-like protein complex that participates in mRNA degradation. *EMBO J* **19:** 1661-1671

Brahms H, Meheus L, de Brabandere V, Fischer U, Lührmann R (2001) Symmetrical dimethylation of arginine residues in spliceosomal Sm protein B/B' and the Sm-like protein LSm4, and their interaction with the SMN protein. *RNA* **7:** 1531-1542

Branlant C, Krol A, Ebel J-P, Lazar E, Haendlerl B, Jacobl M (1982) U2 RNA shares a structural domain with U1, U4, and U5 RNAs. *EMBO J* **1:** 1259-1265

Bringmann P, Appel B, Rinke J, Reuter R, Theissen H, Lührmann R (1984) Evidence for the existence of snRNAs U4 and U6 in a single ribonucleoprotein complex and for their association by intermolecular base pairing. *EMBO J* **3:** 1357-1363

Burge CB, Padgett RA, Sharp PA (1998) Evolutionary fates and origins of U12-type introns. *Mol Cell* **2:** 773-785

Carmo-Fonseca M, Ferreira J, Lamond AI (1993) Assembly of snRNP-containing coiled bodies is regulated in interphase and mitosis - evidence that the coiled body is a kinetic nuclear structure. *J Cell Biol* **120:** 841-852

D'Andrea LD, Regan L (2003) TPR proteins: the versatile helix. *Trends Biochem Sci* **28:** 655-662

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

Darzacq X, Jády BE, Verheggen C, Kiss AM, Bertrand E, Kiss T (2002) Cajal body-specific small nuclear RNAs: a novel class of 2'-O-methylation and pseudouridylation guide RNAs. *EMBO J* **21:** 2746-2756

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

Dundr M, Hebert MD, Karpova TS, Staněk D, Xu H, Shpargel KB, Meier UT, Neugebauer KM, Matera AG, Misteli T (2004) In vivo kinetics of Cajal body components. *J Cell Biol* **164:** 831-842

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

Ferreira JA, Carmo-Fonseca M, Lamond AI (1994) Differential interaction of splicing snRNPs with coiled bodies and interchromatin granules during mitosis and assembly of daughter cell nuclei. *J Cell Biol* **126:** 11-23

Fornerod M, Ohno M, Yoshida M, Mattaj IW (1997) CRM1 is an export receptor for leucine-rich nuclear export signals. *Cell* **90:** 1051-1060

Fourmann JB, Schmitzová J, Christian H, Urlaub H, Ficner R, Boon KL, Fabrizio P, Lührmann R (2013) Dissection of the factor requirements for spliceosome disassembly and the elucidation of its dissociation products using a purified splicing system. *Genes Dev* **27:** 413-428

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

Fury MG, Zieve GW (1996) U6 snRNA maturation and stability. *Exp Cell Res* **228:** 160-163

Gall JG (2003) The centennial of the Cajal body. *Nat Rev Mol Cell Biol* **4:** 975-980

Ganot P, Jády BE, Bortolin M-L, Darzacq X, Kiss T (1999) Nucleolar factors direct the 2′-O-ribose methylation and pseudouridylation of U6 spliceosomal RNA. *Mol Cell Biol* **19:** 6906-6917

Girard C, Neel H, Bertrand E, Bordonné R (2006) Depletion of SMN by RNA interference in HeLa cells induces defects in Cajal body formation. *Nucleic Acids Res* **34:** 2925-2932

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

Golas MM, Sander B, Will CL, Lührmann R, Stark H (2003) Molecular architecture of the multiprotein splicing factor SF3b. *Science* **300:** 980-984

Golas MM, Sander B, Will CL, Lührmann R, Stark H (2005) Major conformational change in the complex SF3b upon integration into the spliceosomal U11/U12 di-snRNP as revealed by electron cryomicroscopy. *Mol Cell* **17:** 869-883

Grimm C, Chari A, Pelz JP, Kuper J, Kisker C, Diederichs K, Stark H, Schindelin H, Fischer U (2013) Structural basis of assembly chaperone-mediated snRNP formation. *Mol Cell* **49:** 692-703

Grimmler M, Otter S, Peter C, Muller F, Chari A, Fischer U (2005) Unrip, a factor implicated in cap-independent translation, associates with the cytosolic SMN complex and influences its intracellular localization. *Hum Mol Genet* **14:** 3099-3111

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

Gu J, Shumyatsky G, Makan N, Reddy R (1997) Formation of 2',3'-cyclic phosphates at the 3' end of human U6 small nuclear RNA in vitro. Identification of 2',3'-cyclic phosphates at the 3' ends of human signal recognition particle and mitochondrial RNA processing RNAs. *J Biol Chem* **272:** 21989-21993

Hallais M, Pontvianne F, Andersen PR, Clerici M, Lener D, Benbahouche Nel H, Gostan T, Vandermoere F, Robert MC, Cusack S, Verheggen C, Jensen TH, Bertrand E (2013) CBC-ARS2 stimulates 3'-end maturation of multiple RNA families and favors cap-proximal processing. *Nat Struct Mol Biol* **20:** 1358-1366

Hao le T, Fuller HR, Lam le T, Le TT, Burghes AH, Morris GE (2007) Absence of gemin5 from SMN complexes in nuclear Cajal bodies. *BMC Cell Biol* **8**

Harada K, Yamada A, Yang D, Itoh K, Shichijo S (2001) Binding of a SART3 tumor-rejection antigen to a pre-mRNA splicing factor RNPS1: a possible regulation of splicing by a complex formation. *Int J Cancer* **93:** 623-628

Hebert MD, Shpargel KB, Ospina JK, Tucker KE, Matera AG (2002) Coilin methylation regulates nuclear body formation. *Dev Cell* **3:** 329-337

Hermann H, Fabrizio P, Raker VA, Foulaki K, Hornig H, Brahms H, Lührmann R (1995) snRNP Sm proteins share two evolutionarily conserved sequence motifs which are involved in Sm protein-protein interactions. *EMBO J* **14:** 2076-2088

Huang CJ, Ferfoglia F, Raleff F, Kramer A (2011) Interaction domains and nuclear targeting signals in subunits of the U2 small nuclear ribonucleoprotein particle-associated splicing factor SF3a. *J Biol Chem* **286:** 13106-13114

Huber J, Dickmanns A, Lührmann R (2002) The importin-beta binding domain of snurportin1 is responsible for the Ran- and energy-independent nuclear import of spliceosomal U snRNPs in vitro. *J Cell Biol* **156:** 467-479

Champion EA, Kundrat L, Regan L, Baserga SJ (2009) A structural model for the HAT domain of Utp6 incorporating bioinformatics and genetics. *Protein Eng Des Sel* **22:** 431-439

Chu JL, Elkon KB (1991) The small nuclear ribonucleoproteins, SmB and B', are products of a single gene. *Gene* **97:** 311-312

Izaurralde E, Lewis J, Gamberi C, Jarmolowski A, McGuigan C, Mattaj IW (1995) A cap-binding protein complex mediating U snRNA export. *Nature* **376:** 709-712

Jády BE, Darzacq X, Tucker KE, Matera AG, Bertrand E, Kiss T (2003) Modification of Sm small nuclear RNAs occurs in the nucleoplasmic Cajal body following import from the cytoplasm. *EMBO J* **22:** 1878-1888

Kaiser TE, Intine RV, Dundr M (2008) De novo formation of a subnuclear body. *Science* **322:** 1713-1717

Kambach C, Mattaj IW (1994) Nuclear transport of the U2 snRNP-specific U2B'' protein is mediated by both direct and indirect signalling mechanisms. *J Cell Sci* **107:** 1807-1816

Karaduman R, Dube P, Stark H, Fabrizio P, Kastner B, Lührmann R (2008) Structure of yeast U6 snRNPs: arrangement of Prp24p and the LSm complex as revealed by electron microscopy. *RNA* **14:** 2528-2537

Karijolich J, Yu Y-T (2010) Spliceosomal snRNA modifications and their function. *RNA Biol* **7:** 192-204

Kitao S, Segref A, Kast J, Wilm M, Mattaj IW, Ohno M (2008) A compartmentalized phosphorylation/dephosphorylation system that regulates U snRNA export from the nucleus. *Mol Cell Biol* **28:** 487-497

Klingauf M, Staněk D, Neugebauer KM (2006) Enhancement of U4/U6 small nuclear ribonucleoprotein particle association in Cajal bodies predicted by mathematical modeling. *Mol Biol Cell* **17:** 4972-4981

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

Laggerbauer B, Liu S, Makarov EM, Vornlocher H-P, Makarova OV, Ingelfinger D, Achsel T, Lührmann R (2005) The human U5 snRNP 52K protein (CD2BP2) interacts with U5-102K (hPrp6), a U4/U6.U5 tri-snRNP bridging protein, but dissociates upon tri-snRNP formation. *RNA* **11:** 598-608

Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, Devon K, Dewar K, Doyle M, FitzHugh W, et al (2001) Initial sequencing and analysis of the human genome. *Nature* **409:** 860-921

Lemm I, Girard C, Kuhn AN, Watkins NJ, Schneider M, Bordonné R, Lührmann R (2006) Ongoing U snRNP biogenesis is required for the integrity of Cajal bodies. *Mol Biol Cell* **17:** 3221-3231

Leung AK, Nagai K, Li J (2011) Structure of the spliceosomal U4 snRNP core domain and its implication for snRNP biogenesis. *Nature* **473:** 536-539

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

Li X, Zhang W, Xu T, Ramsey J, Zhang L, Hill R, Hansen KC, Hesselberth JR, Zhao R (2013) Comprehensive in vivo RNA-binding site analyses reveal a role of Prp8 in spliceosomal assembly. *Nucleic Acids Res* **41:** 3805-3818

Licht K, Medenbach J, Lührmann R, Kambach C, Bindereif A (2008) 3'-cyclic phosphorylation of U6 snRNA leads to recruitment of recycling factor p110 through LSm proteins. *RNA* **14:** 1532-1538

Lin PC, Xu RM (2012) Structure and assembly of the SF3a splicing factor complex of U2 snRNP. *EMBO J* **31:** 1579-1590

Liu Q, Dreyfuss G (1996) A novel nuclear structure containing the survival of motor neurons protein. *EMBO J* **15:** 3555-3565

Liu S, Li P, Dybkov O, Nottrott S, Hartmuth K, Lührmann R, Carlomagno T, Wahl MC (2007) Binding of the human Prp31 Nop domain to a composite RNA-protein platform in U4 snRNP. *Science* **316:** 115-120

Liu S, Rauhut R, Vornlocher HP, Lührmann R (2006) The network of protein-protein interactions within the human U4/U6.U5 tri-snRNP. *RNA* **12:** 1418-1430

Liu Y, Li J, Kim BO, Pace BS, He JJ (2002) HIV-1 Tat protein-mediated transactivation of the HIV-1 long terminal repeat promoter is potentiated by a novel nuclear Tat-interacting protein of 110 kDa, Tip110. *J Biol Chem* **277:** 23854-23863

Liu Y, Timani K, Ou X, Broxmeyer HE, He JJ (2013) C-MYC controlled TIP110 protein expression regulates OCT4 mRNA splicing in human embryonic stem cells. *Stem Cells Dev* **22:** 689-694

Long L, Thelen JP, Furgason M, Haj-Yahya M, Brik A, Cheng D, Peng J, Yao T (2014) The U4/U6 recycling factor SART3 has histone chaperone activity and associates with USP15 to regulate H2B deubiquitination. *J Biol Chem* **289:** 8916-8930

Machyna M, Kehr S, Straube K, Kappei D, Buchholz F, Butter F, Ule J, Hertel J, Stadler PF, Neugebauer KM (2014) The coilin interactome identifies hundreds of small noncoding RNAs that traffic through Cajal bodies. *Mol Cell* **56:** 389-399

Makarov EM, Makarova OV, Urlaub H, Gentzel M, Will CL, Wilm M, Lührmann R (2002) Small nuclear ribonucleoprotein remodeling during catalytic activation of the spliceosome. *Science* **298:** 2205-2208

Makarov EM, Owen N, Bottrill A, Makarova OV (2012) Functional mammalian spliceosomal complex E contains SMN complex proteins in addition to U1 and U2 snRNPs. *Nucleic Acids Res* **40:** 2639-2652

Makarova OV, Makarov EM, Liu S, Vornlocher H-P, Lührmann 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

Massenet S, Pellizzoni L, Paushkin S, Mattaj IW, Dreyfuss G (2002) The SMN complex is associated with snRNPs throughout their cytoplasmic assembly pathway. *Mol Cell Biol* **22:** 6533-6541

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

Matera AG, Ward DC (1993) Nucleoplasmic organization of small nuclear ribonucleoproteins in cultured human cells. *J Cell Biol* **121:** 715-727

Medenbach J, Schreiner S, Liu S, Lührmann R, Bindereif A (2004) Human U4/U6 snRNP recycling factor p110: mutational analysis reveals the function of the tetratricopeptide repeat domain in recycling. *Mol Cell Biol* **24:** 7392-7401

Meister G, Bühler D, Laggerbauer B, Zobawa M, Lottspeich F, Fischer U (2000) Characterization of a nuclear 20S complex containing the survival of motor neurons (SMN) protein and a specific subset of spliceosomal Sm proteins. *Hum Mol Genet* **9:** 1977-1986

Meister G, Eggert C, Bühler D, Brahms H, Kambach C, Fischer U (2001) Methylation of Sm proteins by a complex containing PRMT5 and the putative U snRNP assembly factor pICln. *Curr Biol* **11:** 1990-1994

Montemayor EJ, Curran EC, Liao HH, Andrews KL, Treba CN, Butcher SE, Brow DA (2014) Core structure of the U6 small nuclear ribonucleoprotein at 1.7-A resolution. *Nat Struct Mol Biol* **21:** 544-551

Mouaikel J, Narayanan U, Verheggen C, Matera AG, Bertrand E, Tazi J, Bordonne R (2003) Interaction between the small-nuclear-RNA cap hypermethylase and the spinal muscular atrophy protein, survival of motor neuron. *EMBO Rep* **4:** 616-622

Narayanan U, Achsel T, Luhrmann R, Matera AG (2004) Coupled in vitro import of U snRNPs and SMN, the spinal muscular atrophy protein. *Mol Cell* **16:** 223-234

Narayanan U, Ospina JK, Frey MR, Hebert MD, Matera AG (2002) SMN, the spinal muscular atrophy protein, forms a pre-import snRNP complex with snurportin1 and importin beta. *Hum Mol Genet* **11:** 1785-1795

Nesic D, Tanackovic G, Krämer A (2004) A role for Cajal bodies in the final steps of U2 snRNP biogenesis. *J Cell Sci* **117:** 4423-4433

Nottrott S, Hartmuth K, Fabrizio P, Urlaub H, Vidovic I, Ficner R, Lührmann R (1999) Functional interaction of a novel 15.5kD [U4/U6.U5] tri-snRNP protein with the 5' stem-loop of U4 snRNA. *EMBO J* **18:** 6119-6133

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

Novotný I, Blažíková M, Staněk D, Herman P, Malinsky J (2011) In vivo kinetics of U4/U6.U5 tri-snRNP formation in Cajal bodies. *Mol Biol Cell* **22:** 513-523

Novotný I, Malinová A, Stejskalová E, Matějů D, Klimešová K, Roithová A, Švéda M, Knejzlík Z, Staněk D (2015) SART3-Dependent Accumulation of Incomplete Spliceosomal snRNPs in Cajal Bodies. *Cell Rep* **10:** 1-12

Novotný I, Podolská K, Blažíková M, Valášek LS, Svoboda P, Staněk D (2012) Nuclear LSm8 affects number of cytoplasmic processing bodies via controlling cellular distribution of Like-Sm proteins. *Mol Biol Cell* **23:** 3776-3785

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

Ospina JK, Gonsalvez GB, Bednenko J, Darzynkiewicz E, Gerace L, Matera AG (2005) Cross-talk between snurportin1 subdomains. *Mol Biol Cell* **16:** 4660-4671

Otter S, Grimmler M, Neuenkirchen N, Chari A, Sickmann A, Fischer U (2007) A comprehensive interaction map of the human survival of motor neuron (SMN) complex. *J Biol Chem* **282:** 5825-5833

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

Patel SB, Bellini M (2008) The assembly of a spliceosomal small nuclear ribonucleoprotein particle. *Nucleic Acids Res* **36:** 6482-6493

Pellizzoni L, Kataoka N, Charroux B, Dreyfuss G (1998) A novel function for SMN, the spinal muscular atrophy disease gene product, in pre-mRNA splicing. *Cell* **95:** 615-624

Pellizzoni L, Yong J, Dreyfuss G (2002) Essential role for the SMN complex in the specificity of snRNP assembly. *Science* **298:** 1775-1779

Pomeranz Krummel DA, Oubridge C, Leung AK, Li J, Nagai K (2009) Crystal structure of human spliceosomal U1 snRNP at 5.5 A resolution. *Nature* **458:** 475-480

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

Price SR, Evans PR, Nagai K (1998) Crystal structure of the spliceosomal U2B"-U2A' protein complex bound to a fragment of U2 small nuclear RNA. *Nature* **394:** 645-650

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

Raker VA, Plessel G, Lührmann 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

Raška I, Andrade LEC, Ochs RL, Chan EKL, Chang CM, Roos G, Tan EM (1991) Immunological and ultrastructural studies of the nuclear coiled body with autoimmune antibodies. *Exp Cell Res* **195:** 27-37

Reddy R, Ro-Choi TS, Henning D, Busch H (1974) Primary sequence of U-1 nuclear ribonucleic acid of Novikoff hepatoma ascites cells. *J Biol Chem* **249:** 6486-6494

Salgado-Garrido J, Bragado-Nilsson E, Kandels-Lewis S, Séraphin B (1999) Sm and Sm-like proteins assemble in two related complexes of deep evolutionary origin. *EMBO J* **18:** 3451-3462

Sander B, Golas MM, Makarov EM, Brahms H, Kastner B, Lührmann R, Stark H (2006) Organization of core spliceosomal components U5 snRNA loop I and U4/U6 di-snRNP within U4/U6.U5 tri-snRNP as revealed by electron cryomicroscopy. *Mol Cell* **24:** 267-278

Segref A, Mattaj IW, Ohno M (2001) The evolutionarily conserved region of the U snRNA export mediator PHAX is a novel RNA-binding domain that is essential for U snRNA export. *RNA* **7:** 351-360

Selenko P, Sprangers R, Stier G, Bühler D, Fischer U, Sattler M (2001) SMN tudor domain structure and its interaction with the Sm proteins. *Nat Struct Biol* **8:** 27-31

Shimba S, Reddy R (1994) Purification of human U6 small nuclear RNA capping enzyme. Evidence for a common capping enzyme for gamma-monomethyl-capped small RNAs. *J Biol Chem* **269:** 12419-12423

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

Schmidt C, Gronborg M, Deckert J, Bessonov S, Conrad T, Lührmann R, Urlaub H (2014) Mass spectrometry-based relative quantification of proteins in precatalytic and catalytically active spliceosomes by metabolic labeling (SILAC), chemical labeling (iTRAQ), and label-free spectral count. *RNA* **20:** 406-420

Schmidt TG, Batz L, Bonet L, Carl U, Holzapfel G, Kiem K, Matulewicz K, Niermeier D, Schuchardt I, Stanar K (2013) Development of the Twin-Strep-tag(R) and its application for purification of recombinant proteins from cell culture supernatants. *Protein Expr Purif* **92:** 54-61

Schneider C, Will CL, Makarova OV, Makarov EM, Lührmann R (2002) Human U4/U6.U5 and U4atac/U6atac.U5 tri-snRNPs exhibit similar protein compositions. *Mol Cell Biol* **22:** 3219-3229

Schwer B (2008) A conformational rearrangement in the spliceosome sets the stage for Prp22-dependent mRNA release. *Mol Cell* **30:** 743-754

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

Sleeman JE, Lamond AI (1999) Newly assembled snRNPs associate with coiled bodies before speckles, suggesting a nuclear snRNP maturation pathway. *Curr Biol* **9:** 1065-1074

Staněk D, Neugebauer KM (2004) Detection of snRNP assembly intermediates in Cajal bodies by fluorescence resonance energy transfer. *J Cell Biol* **166:** 1015-1025

Staněk D, Přidalová-Hnilicová J, Novotný I, Huranová M, Blažíková M, Wen X, Sapra AK, Neugebauer KM (2008) Spliceosomal small nuclear ribonucleoprotein particles repeatedly cycle through Cajal bodies. *Mol Biol Cell* **19:** 2534-2543

Staněk D, Rader SD, Klingauf M, Neugebauer KM (2003) Targeting of U4/U6 small nuclear RNP assembly factor SART3/p110 to Cajal bodies. *J Cell Biol* **160:** 505-516

Stejskalová E, Staněk D (2014) The splicing factor U1-70K interacts with the SMN complex and is required for nuclear gem integrity. *J Cell Sci* **127:** 3909-3915

Tanackovic G, Krämer A (2005) Human splicing factor SF3a, but not SF1, is essential for pre-mRNA splicing in vivo. *Mol Biol Cell* **16:** 1366-1377

Trede NS, Medenbach J, Damianov A, Hung LH, Weber GJ, Paw BH, Zhou Y, Hersey C, Zapata A, Keefe M, Barut BA, Stuart AB, Katz T, Amemiya CT, Zon LI, Bindereif A (2007) Network of coregulated spliceosome components revealed by zebrafish mutant in recycling factor p110. *Proc Natl Acad Sci USA* **104:** 6608-6613

Trippe R, Guschina E, Hossbach M, Urlaub H, Lührmann R, Benecke BJ (2006) Identification, cloning, and functional analysis of the human U6 snRNA-specific terminal uridylyl transferase. *RNA* **12:** 1494-1504

Turunen JJ, Niemela EH, Verma B, Frilander MJ (2013) The significant other: splicing by the minor spliceosome. *Wiley Interdiscip Rev RNA* **4:** 61-76

Verdone L, Galardi S, Page D, Beggs JD (2004) Lsm proteins promote regeneration of pre-mRNA splicing activity. *Curr Biol* **14:** 1487-1491

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

Weber G, Trowitzsch S, Kastner B, Lührmann R, Wahl MC (2010) Functional organization of the Sm core in the crystal structure of human U1 snRNP. *EMBO J* **29:** 4172-4184

Will CL, Lührmann R (2011) Spliceosome structure and function. *Cold Spring Harb Perspect Biol* **3**

Will CL, Schneider C, Hossbach M, Urlaub H, Rauhut R, Elbashir S, Tuschl T, Lührmann R (2004) The human 18S U11/U12 snRNP contains a set of novel proteins not found in the U2-dependent spliceosome. *RNA* **10:** 929-941

Will CL, Schneider C, MacMillan AM, Katopodis NF, Neubauer G, Wilm M, Lührmann R, Query CC (2001) A novel U2 and U11/U12 snRNP protein that associates with the pre-mRNA branch site. *EMBO J* **20:** 4536-4546

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

Xu H, Pillai RS, Azzouz TN, Shpargel KB, Kambach C, Hebert MD, Schumperli D, Matera AG (2005) The C-terminal domain of coilin interacts with Sm proteins and U snRNPs. *Chromosoma* **114:** 155-166

Yang D, Nakao M, Shichijo S, Sasatomi T, Takasu H, Matsumoto H, Mori K, Hayashi A, Yamana H, Shirouzu K, 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

Yong J, Kasim M, Bachorik JL, Wan L, Dreyfuss G (2010) Gemin5 delivers snRNA precursors to the SMN complex for snRNP biogenesis. *Mol Cell* **38:** 551-562

Yoshimoto R, Kataoka N, Okawa K, Ohno M (2009) Isolation and characterization of post-splicing lariat-intron complexes. *Nucleic Acids Res* **37:** 891-902

Younis I, Dittmar K, Wang W, Foley SW, Berg MG, Hu KY, Wei Z, Wan L, Dreyfuss G (2013) Minor introns are embedded molecular switches regulated by highly unstable U6atac snRNA. *Elife* **2**

Yu YT, Shu MD, Steitz JA (1998) Modifications of U2 snRNA are required for snRNP assembly and pre-mRNA splicing. *EMBO J* **17:** 5783-5795

Zhou L, Hang J, Zhou Y, Wan R, Lu G, Yin P, Yan C, Shi Y (2014) Crystal structures of the Lsm complex bound to the 3' end sequence of U6 small nuclear RNA. *Nature* **506:** 116-120