

---

# Oligomerization of Biomacromolecules – Example of RNA Binding Sm/LSm Proteins

---

Bozidarka L. Zaric

Additional information is available at the end of the chapter

<http://dx.doi.org/10.5772/57592>

---

## 1. Introduction

### 1.1. A multitude of RNAs

The central dogma of molecular biology stated that genetic information only flows in one direction, from DNA to proteins via an intermediate called messenger ribonucleic acid (mRNA) [1,2,3]. Originally, ribonucleic acid (RNA) was thought to have roles in information transfer and structure maintenance. Today, we know that RNA performs a remarkable range of functions in the living cell, (control of gene expression, chromosome –end maintenance, housekeeping activities, sorting of proteins in the cell and defines metazoan development) [3]. Although, proteins have enzymatic activities mostly, in the early 1980s has been shown that RNA molecules can catalyze a chemical reaction and RNAs with catalytic activity are called **ribozymes**. The discovery of ribozymes led to the hypothesis that RNA could have been the original molecule of life on earth about four billion years ago; a biopolymer with the ability to self-replicate and that could both store information and catalyze chemical reactions. RNA would have been self-sufficient as the original molecule of life [4]. Discovery of the unexpectedly wide variety of functions carried out by RNA was accompanied by the identification of a multitude of further types of small, non-coding RNAs (small nuclear RNA, small nucleolar RNAs, small interfering RNAs, micro RNAs) highlighting the versatility of RNA as a biochemical tool for the cell [2].

**rRNAs** represent structural and catalytic elements of the ribosome. In the nucleolus of eukaryotic cells, more than 100 tandemly repeated units of rRNA genes are transcribed into long precursor transcripts [7,8]. Following transcription, pre-rRNA is subsequently cleaved to form mature rRNAs and with approximately 80 proteins to form the large and small ribosomal subunits prior to their export to the cytoplasm. **SnoRNAs (Small nucleolar**

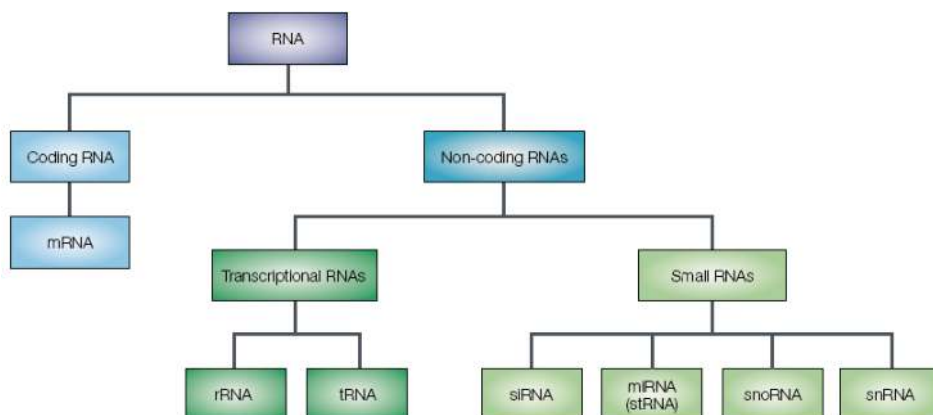


Figure 1. The RNA family (Reprinted from reference 2)

**RNAs**) participate in both the modification and cleavage events that occur during ribosome biogenesis [5]. A second group of small RNAs are the **tRNAs**, which are essential in translation [8,10]. **Micro RNAs** are non-coding RNAs of 22-24 nucleotides in length. They down regulate gene expression by attaching themselves to mRNA, thereby preventing them from being translated into protein. Another type of non-coding RNA is the **small interfering RNA**. These small RNAs mediate life time of RNA by interacting with mRNAs and labeling it for destruction [2, 4]. **PiRNA** is the large class of small non-coding RNAs which acts with piwi proteins. These piRNA have been linked to both epigenetic and posttranslational gene silencing of retrotransposons and other genetic elements in germline cells [11,12]. Telomerase is complex of proteins and RNA, and is responsible for maintaining of natural end of chromosomes. Telomerase acts as reverse transcriptase because its mechanism of action is copying RNA template into DNA [10].

Small nuclear RNAs are components of the macromolecular machinery (spliceosome) that has a role in the maturation of mRNA. They are termed **U snRNAs** (stands for uridyl rich small nuclear RNA). U1, U2, U4, U5, U7, U11, and U12 are synthesized in the nucleus by RNA polymerase II. After that they are transported to the cytoplasm where association with the U snRNP proteins (proteins which associate with uridine rich small nuclear RNAs, generating UsnRNP, uridyl rich small nuclear ribonuclear particles) occurs, followed by re-import into the nucleus [1, 13]. U6 and U8 snRNA belong as well to class of small nuclear RNA but their synthesis, biogenesis and function differs from mentioned UsnRNAs. U3 snRNA shares common denomination but as well this small RNA is found in the nucleolus, and has role in pre-rRNA processing and has C/D box motif, which technically make it a member of the C/D class of snoRNAs [14].

## 1.2. Structure of RNA and association with proteins

DNA and RNA have similar covalent structures, the only difference being the change from a 2'-deoxyribose sugar to a ribose sugar and from a methyl group in thymine to a hydrogen in

uracil. RNA has much wider biological activities and adopts a wider range of structures. DNA double helices preferentially assume the B-form structure in solution and RNA double helices are found in the A-form. The **RNA A-form double helix** has a narrow and deep major groove, which prevents proteins to recognize RNA in a manner analogous to the way they recognize DNA. An RNA molecule can locally adopt several types of secondary structure (bulges, hairpins, internal loops) [15, 16]. Eukaryotic mRNAs are almost always associated with RNA-binding proteins. RNA-binding proteins generally have a modular structure and contain RNA-binding domains of 70–150 amino acids that mediate RNA recognition. Three major classes of eukaryotic RNA-binding protein domains are known: the RNA-recognition motif (RRM), the double stranded RNA binding domain (dsRBD) and the K-homology (KH) domain [17].

### 1.3. All mRNA processing steps are coupled

Eukaryotic gene expression is a complex, stepwise process that begins with transcription (synthesis of pre-mRNA) [18]. Mature mRNAs are produced in the cell nucleus from primary transcripts of coding genes (pre-mRNAs) by a series of processing events which include capping, splicing, and 3' end polyadenylation. Mature mRNAs are transported to the cytoplasm. All modification steps are coupled and influence each other. RNA polymerase II is a key molecular coordinator of these processing events, and phosphorylation of it has regulatory role [19, 20, 21,].

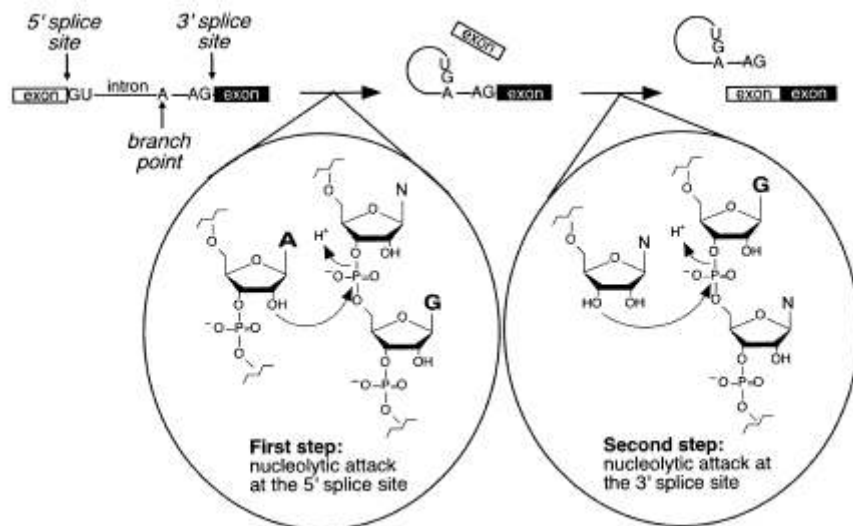
### 1.4. Removal of introns and the splicing reaction

In 1977, a number of research groups discovered that the genes of higher organisms are often made up of a sequence of coding (called exons) and non-coding base sequences (introns). During transcription, all parts of the gene are copied to form a strand of pre-mRNA. Introns are removed and the exons stitched together so that the now continuous exons can be translated to produce a protein. This splicing of the pre-mRNA is a multistage process, carried out by complex macromolecular machinery known as the spliceosome, which is among the most complex macromolecular machineries in the cell [22].

Splicing of precursors to mRNAs occurs in two steps, both involving a single transesterification reaction [23]. Assembly and function of the spliceosome requires approximately 300 polypeptides and five snRNAs, not considering gene-specific RNA-binding factors [23]. There are two distinct types of spliceosome in most cells. The major class U2-type spliceosome is universal in eukaryotes, whereas the minor class or U12-type spliceosome is not present in some organisms. The evolutionary relation between these two spliceosomes is uncertain.

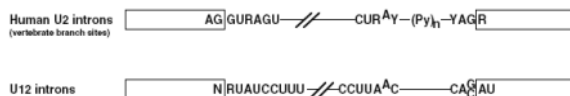
### 1.5. Types of introns

The pre-mRNA contains conserved elements at its intron/ exon boundaries that determine the proper sites for the splicing reaction (Figure 3.). The 5' splice site contains a conserved consensus sequence, which is AG/GURAGU (R=purine, / denotes the exon/ intron boundary). The branch site lies between 100 and 18 bases upstream of the 3' splice site and has the consensus: CURAY for vertebrates (A branching nucleotide, Y is pyrimidine). In higher



**Figure 2.** The two chemical steps of splicing (Reprinted from reference 23)

eukaryotes, a polypyrimidine tract variable in length is often located between the branch site and the 3' splice site. The 3' splice site has the consensus:YAG/R for mammals. This class of introns is spliced by U2 spliceosome. The U12 type introns have different consensus sequences and are spliced by the U12 spliceosome [24]. The number of known U12 introns is still very small. U12 type introns are present in many vertebrates, nematodes, insect, and plant species.



**Figure 3.** Splice site consensus sequences. Comparison of splice site consensus sequences for human U2 dependent and U12 dependent introns. The most conserved regions, 5' splice site (5' SS), branch point (BP), 3' splice site (3' SS), are shown with their consensus sequences (R=purine, Y=pyrimidine). The polypyrimidine tract often present in metazoan U2 dependent introns is indicated as (Py). (Reprinted from reference 25)

### 1.6. Small nuclear ribonucleoproteins, snRNPs

Small ribonucleoproteins (RNPs) are tight complexes of one or more proteins with a short RNA molecule (usually 60-300 nucleotides). RNPs inhabit nuclear and cytoplasmical compartments of the eukaryotic cell [25]. Those that reside in the nucleus, the small nuclear ribonucleoproteins (snRNPs) can themselves be divided into two families. There are **snRNPs** of the nucleoplasm, whose function lies in preparing messenger RNA for export into the cytoplasm. A different set of snRNPs, called **snoRNPs**, reside in the nucleolus [26]. The **nucleolus** is a

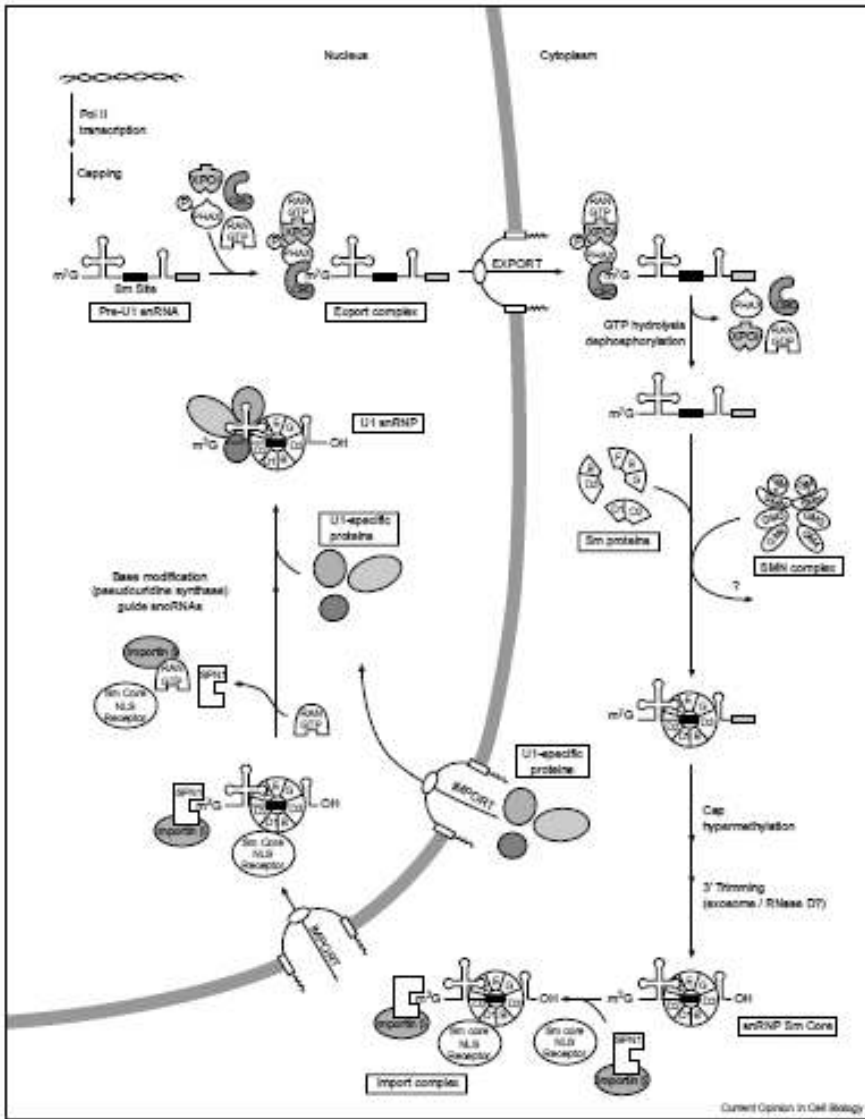
structure in the nucleus composed of proteins and nucleic acids. Its function is to transcribe ribosomal RNA (rRNA) and combine it with proteins to form ribosomes. There are about 200 distinct kinds of snRNPs (they differ according to the RNA or protein components) with abundances between  $10^4$  (for snoRNPs) to over  $10^6$  copies per cell (for snRNPs in major spliceosomes). They generally play a role in gene expression. One exception is the telomerase snRNP, essential for genome maintenance, which is present only in a few copies per cell. Spliceosomes are formed around the pre-mRNA substrate by the successive assembly of five small nuclear ribonucleoprotein-particles (snRNPs): U1, U2, U4, U5, and U6. These particles are composed each of a small nuclear RNA (snRNA), seven Sm core proteins common to all snRNPs (except for the U6 snRNP, which contains a related set of seven proteins, the Sm-like proteins) and several snRNP-specific proteins. The snRNPs play a central role in the process of splicing. They are responsible for the recognition of splice sites, definition of exon/ intron boundaries. These interactions are partially mediated through base pairing and are dynamic so that the spliceosome complex changes during the process of splicing.

### 1.7. U snRNP biogenesis

Subsequent to transcription by RNA polymerase II and capping, pre-U1 snRNA assembles with several factors including cap-binding proteins (CBP), a phosphorylated adaptor for RNA export (PHAX), Crm1, and Ran-GTP, which all together mediate export of U1 snRNA to the cytoplasm. After export, Sm proteins interact with the U snRNAs to form the snRNP Sm core. This step is facilitated by the SMN complex (survival of motor neurons complex). The SMN complex is composed from SMN protein and the other proteins called Gemins (Gemins 2-7). Nuclear re-import is mediated by snurportin-1 (SPN1), which binds to the snRNAs m3G cap structure. After import, these factors dissociate. The U1 specific proteins are imported independently into the nucleus, where assembly into mature U1 snRNP occurs [27]. This is a pathway shared with U1, U2, U4 and U5 snRNPs.

### 1.8. Spliceosome assembly

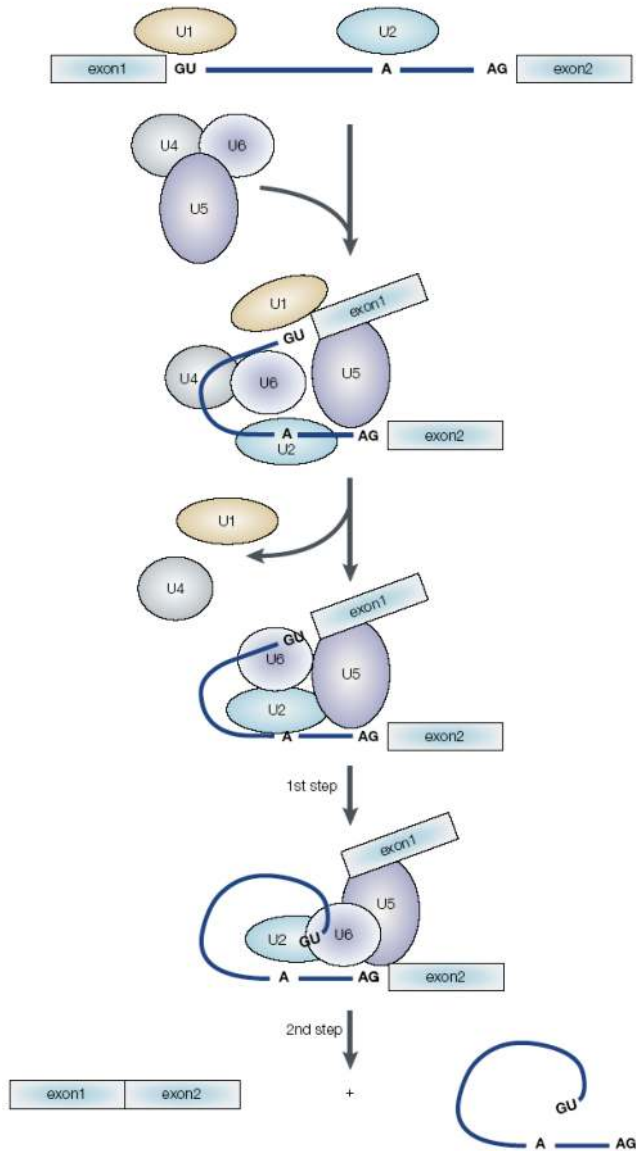
Assembly of a spliceosome for excision of an intron requires recognition of sequences at the 5' splice site as well as the branch site and nearby 3' splice site. U1 snRNA binds to the 5' end of the intron using sequence complementarities. There are reports which show that the U1 snRNA recognizes the 5' splice site in a preassembled penta-snRNP complex [28]. U2 snRNP complex associates with the branch region. Early snRNP/pre-mRNA complexes are preferentially committed to splicing as compared to free RNA and thus are called **commitment complexes** (CCs). The process of U2 snRNP association is ATP-dependent and four proteins are critical for recognition [29]. Subsequent to the binding of U2 snRNP complex, a tri-snRNP complex containing U4/U6 snRNP and U5 snRNP associates in an ATP-dependent manner to form complex A2-1. It is likely that the U1 snRNA/pre-mRNA duplex dissociates at this stage. The 5' splice site sequence is probably paired on the intron side to U6 snRNA and on the exon side to U5 snRNA. The transition between complex A2-1 and A1 requires destabilization of at least U4/U6 di-snRNA. As only three snRNAs U2, U6 and U5 are associated with the spliceosome at the moment of catalysis, and as U5 snRNA pairing with exon sequences is not essential, the



**Figure 4.** The U snRNP biogenesis pathway (Reprinted from reference 27)

catalytic site is either created by U6 snRNA by U2 snRNA or both. The action of certain proteins is required for the transition to the second step in splicing. The catalytic site for the second step is created by either U6 snRNA, U2 snRNA, or associated proteins. The reannealing of released U4 and U6 snRNP and association with U5 forms the U4/U6 U5 tri-snRNP complex which is then ready to reassemble on another commitment complex.

The classical view of spliceosome assembly has been challenged by Stevens et al [30]. This group isolated from yeast a penta-snRNP complex which when supplied with soluble components, does splice pre-mRNA.



**Figure 5.** A simplified view of the splicing process (Reprinted from reference 5)

## 1.9. mRNA stabilization, degradation

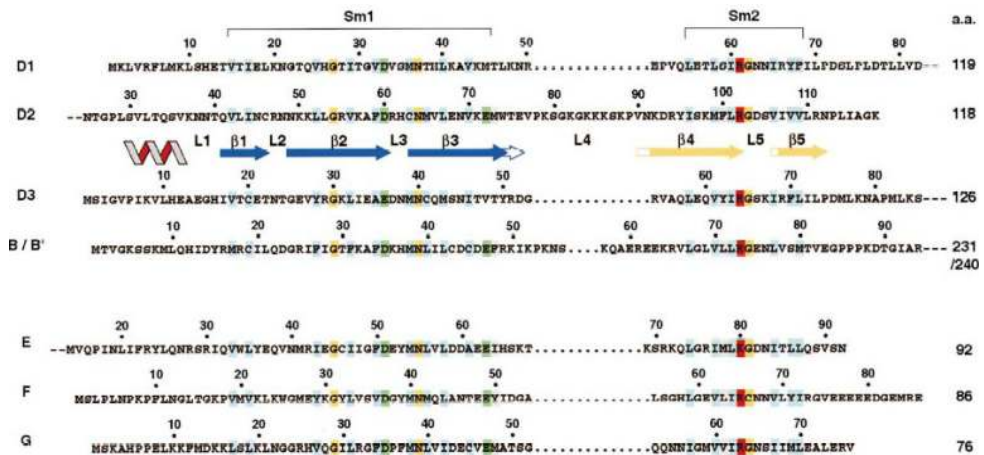
Regulation of mRNA decay rates is an important control point in determining the abundance of cellular transcripts. Some mRNA has half-lives that are 100 times shorter than cellular generation times and some mRNA have half-lives spanning several cell cycles [21]. The poly (A) tail is important in stabilization of mRNA. It interacts with the poly (A) binding protein (PABP), which makes direct contact with a specific region of the translation-initiation factor (eIF4E). Translation initiation factor (eIF4) interacts with the cap binding proteins. In this way, a ternary (PABP-translation initiation,-cap binding protein, poly (A) tail) complex is formed which circularizes mRNA *in vitro*, promoting translation and stabilization of mRNAs [21]. Several sequence elements can regulate the mRNA turnover rate, either by its promotion (destabilizer elements) or by its inhibition (stabilizer elements). Important elements are A+U-rich elements (ARE), located in the 3' untranslated regions (UTR) of mRNAs [31]. At least four different ways of mRNA degradation have been reported in eukaryotic cells [32]. In most cases, degradation of the transcript begins with the shortening of the poly (A) tail at the 3' end of mRNA. After shortening of the poly (A) tail follows the removal of the 5' cap structure (decapping), thereby exposing the transcript to digestion by a 5' to 3' exonuclease. Family of LSm proteins is involved in degradation of mRNAs, in the decapping step. Transcripts can be degraded in the 3'-5' direction after deadenylation. This process is catalyzed by the exosome [33]. One mRNA degradation pathway is the **nonsense mediated decay (NMD)**, which provides strongest evidence for a link between translation and turnover [34].

## 2. Family of Sm-LSm proteins

### 2.1. Sm proteins, assembly of U1, U2, U4, U5 snRNPs

The Sm proteins were first discovered as antigens targeted by so-called **Anti-Sm antibodies** in a patient with a form of Systemic lupus erythematosus (SLE), a debilitating autoimmune disease. They were named Sm proteins in honor of Stephanie Smith, a patient who suffered from SLE. Other proteins with very similar structures were subsequently discovered and named LSm proteins. The common proteins for U1, U2, U4 and U5 snRNPs are named Sm proteins due to their recognition by anti-Sm autoantibodies (isolated from the serum of patients with autoimmune diseases [35, 36]). Eight proteins: B', B, D1-D3, E, F and G have been characterized in human cells. All of the Sm core proteins are encoded by separate genes [37], with the exception of B and B'. The B and B' that result from an alternative splicing of gene 6628 located on chromosome 20, locus 20p13, only differs in 11 amino acids at the C-terminus [38]. In neural tissues, SmN replaces SmB and SmB' [39]. Two sequence motifs, named Sm1 and Sm2, are found in all known Sm proteins, what is reason that they are called Sm proteins [40]. The N terminal Sm1 motif is composed of 32 amino acids. The Sm2 motif, located closer to the C terminus, is shorter spanning only 14 amino acids [35]. Sm motif 1 and Sm motif 2 are separated by a linker of variable length. The alignment of the sequences of human Sm proteins reveals a striking conservation of the two motifs. Majority of the Sm proteins have amino or carboxy-terminal extensions.

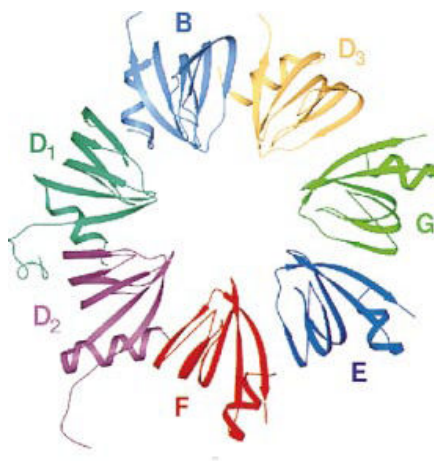




**Figure 6.** Primary and secondary structure of Sm proteins (Reprinted from reference 41). Amino acid sequence alignment of the human Sm (D1, D2, D3, B/B', E, F, and G) proteins with secondary structure elements. Wavy line, helix; arrows,  $\beta$  strands. The  $\beta$  strands within the Sm1 and Sm2 motifs are colored blue and yellow, respectively. The  $\beta$  strands and interconnecting loops are numbered consecutively from the N terminus. The conserved Sm1 and Sm2 motifs are indicated and the conserved residues within these motifs are highlighted in blue (hydrophobic), grey (hydrophobic, less well conserved), orange (neutral polar), red (basic) and green (acidic). (Reprinted from reference 41)

Solved structures of this protein family members (pdb codes: 1d3b,1b34,1hk9, 1h64,1i8f,1i4k, 1kq1,3bw1,1th7) show that the fold is highly conserved. It is defined by an N-terminal helix, followed by a five-stranded anti-parallel  $\beta$  sheet. Strands  $\beta 1$ ,  $\beta 2$ , and  $\beta 3$  are part of the Sm1 motif, whereas the Sm2 motif forms strands  $\beta 4$  and  $\beta 5$ . The five stranded  $\beta$  sheet is strongly bent in the middle and the conserved hydrophobic residues form a hydrophobic core [41].

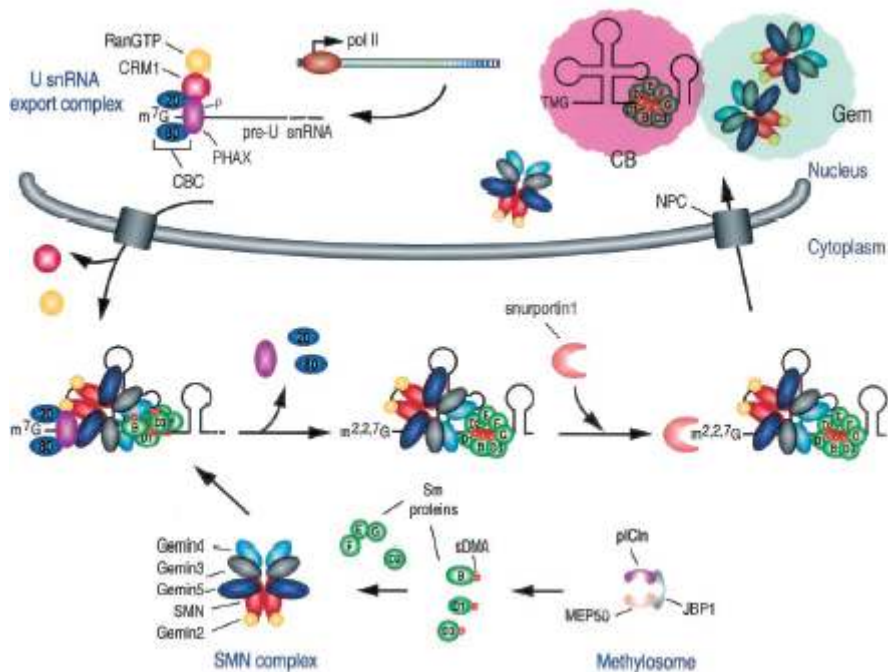
The Sm proteins bind to the Sm site of U snRNAs [42]. The Sm site consensus sequence (PuAU4-6Gpu) has a central, uridine rich tract and flanking purines. *In vitro* studies reveal that the single-stranded U rich region and the 5' adenosine of the Sm site play critical roles in Sm protein assembly. The uridine bases and the 2' hydroxyl groups collectively provide binding determinants [43]. In the absence of U snRNA, the seven Sm proteins form three stable subcomplexes (D3B, or D3B', D1D2, and EFG). These sub complexes then form a heptameric ring around the snRNA Sm site, and as such the complex is termed the **Sm core**. SnRNP core assembly is an ordered pathway that involves formation of a sub-core particle followed by formation of the full Sm core, which promotes cap hypermethylation and pre-snRNP import [44]. The Sm fold is necessary and sufficient for the formation of specific inter-subunit interactions. Biochemical results indicate that there is one copy of each Sm protein in the snRNP core domain and therefore support the heptameric ring model of the snRNP core domain [45]. None of the single Sm proteins has a known RNA recognition motif, so another type of interactions with RNA must be involved. Crosslinking studies indicate that Sm motif 1 is responsible for interactions with RNA, and Sm motif 2 for protein-protein interactions [43].



**Figure 7. Proposed Higher-Order Assembly of the Human Core snRNP Proteins.** The seven core Sm proteins (B/B', D1, D2, D3, E, F, and G) are arranged within the seven-membered ring based on the crystal structures of the D1D2(1b34) and D3B (d3b) complexes and pairwise interactions deduced from biochemical and genetic experiments. (Reprinted from reference 41)

Basic residues of human and yeast SmB, SmD1 and SmD3 are reported to be responsible for import of the Sm core particle [45]. *In vitro*, the snRNP core domain can be assembled from purified components [46]. Assembly of the spliceosomal class of snRNPs *in vivo* is an active process that is mediated by several factors, including the product of the SMN gene (survival of motor neurons gene). Mutations of SMN gene are responsible for spinal muscular atrophy disease (SMA). Spinal muscular atrophy is an autosomal recessive disorder correlated with loss of motor neurons, as a result of a mutation on the SMN gene [47]. The SMN protein is ubiquitously expressed in all tissues of metazoan organisms reflecting the fact that it provides a fundamental activity required by all cells. The SMN protein is predominantly cytosolic but it is also found in the nucleus, namely in a few spherical nuclear domains that overlap with the so-called Cajal bodies (where snRNPs and snoRNPs are localized). These spherical domains have been called Gemini of Cajal bodies (Gems). Proteins associated with the SMN protein are called Gemins. The SMN complex interacts *in vitro* with Sm and LSm proteins which contain symmetrically methylated RG (arginine –glycine) repeats [48]. Symmetrically methylated RG repeats of SmD1, SmD3 and LSm4 are generated by action of the so-called methylosome [48,49]. The SMN complex binds to the human hypermethylase which suggests that SMN may have a role in formation of the snRNA m<sup>3</sup>G cap structure. It has been proposed that after binding of SMN to the Sm core proteins, SMN promotes engagement of TGS1 to the m<sup>7</sup>G-capped snRNP particle. According to that model, SMN dissociates from the C terminal part of the B/B' Sm proteins followed by association of TGS1 and Sm core. This step allows formation of the m<sup>3</sup>G-cap. The association of Snurportin 1 with the m<sup>3</sup>G-cap can promote release of TGS1 and generate import-competent snRNP [50]. According to these data, the SMN complex interacts with protein components of U snRNPs, but there are reports [51] on sequence-specific interactions between U1 snRNA and the SMN complex. The nuclear

localization signal of U snRNP is composed of the U snRNAs 2,2,7 tri-methylguanosine cap and Sm core domain. Snurportin 1 binds to the m<sup>3</sup>G cap but not to the Sm core. Snurportin1 has an N-terminal importin beta binding domain and a carboxy terminal m<sup>3</sup>G-cap binding domain [52]. The Importin beta binding domain allows for snRNP's cargo to be imported in a Ran independent fashion. After import of snRNPs into the nucleus, Snurportin1 dissociates from its cargo and is exported back into the cytoplasm using Crm1, a receptor for leucine-rich nuclear export signals [53]. The SMN complex not only mediates snRNP core assembly but is an integral complex component during the entire snRNP core biogenesis in the cytoplasm. It is not excluded that SMN is actually the long-sought Sm core nuclear localization signal receptor [54].



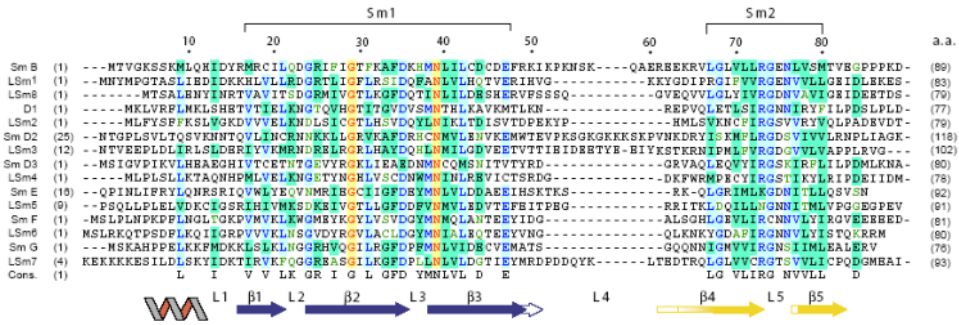
**Figure 8.** Schematic model of the role of SMN in snRNP core biogenesis in the cytoplasm. (Reprinted from reference 54)

## 2.2. LSm proteins

Sm and Sm-like proteins are found in all kingdoms of life: *eukarya*, *archaea* and *bacteria*. These proteins were found even in *Archaea*. Because *Archaea* have been proposed to be related to the ancestor of the eukaryotic nuclear genome, this fact suggests that an LSm protein gene was present in the last common ancestor. *Archaeobacteria* harbour between one and two genes which encode for Sm motif containing proteins. The *in vivo* functions of

archaeal Sm proteins remain unknown (in contrast to the eukaryotic and bacterial homologs, and fact that high resolution structure from archaeal systems is known (pdb code 1ljo) [55]. LSm proteins have been identified in plants as well [56]. Eukaryotic genomes have more than 20 Sm/LSm genes each, corresponding to the LSm and Sm proteins which are components of Sm and LSm complexes. Database searches in the yeast genome, revealed 16 Sm motif containing proteins. Some Sm-like proteins were found to interact weakly with some Sm proteins, most probably via non-specific Sm domain interactions [57], but some of the LSm proteins interact with Sm proteins as part of U7 snRNP. In yeast there are nine LSm proteins, in humans more than eight. Each of the human LSm proteins has one orthologue in yeast. Yeast LSm2p-LSm7p share sequence identity with human LSm2-LSm7 ranging from 41-62%. LSm9p appears to be present only in yeast. Yeast LSm8p aligns best with human LSm8 (26% identity). In addition, LSm proteins are highly conserved throughout all eukaryotic kingdoms, as the homologues in insect, nematode and plant database share between 50 and 75 % identity with their human counterparts. Each of the LSm proteins in humans can clearly be best aligned with one of the canonical Sm proteins. However, their sequence identities are not high enough to allow the conclusion that LSm proteins undergo the same protein-protein interactions like Sm proteins [58].

## Structural Alignment of Human Sm/LSm Proteins



**Figure 9.** Structural Alignment of Human Sm/LSm proteins

Similar to canonical Sm proteins, the LSm proteins are recognized by antibodies from patients suffering from systemic lupus erythematosus (SLE) [59]. Sm/LSm proteins always appear as homomeric (in the case of prokaryotes) or heteromeric (in eukaryotes) ringlike multimers. These ring-shaped complexes, generally containing either six or seven subunits, are the functional LSm protein unit. All canonical Sm proteins are essential for vegetative growth of yeast. LSm proteins have variable effects after depletion in yeast. In mice embryos, LSm4-null zygotes survived to the blastocyst stage, but died shortly after [60].

### **2.3. Role of LSm 2-8 oligomers in U6snRNP assembly**

The LSm2-8 complex was isolated from HeLa cells nuclear extract in an RNA free form. Electron micrographs revealed a doughnut-shaped heterooligomer, similar to the Sm core snRNPs [58]. LSm proteins have a high affinity for single-stranded oligo-U, but they do not recognize the canonical Sm binding site. In yeast and humans, LSm2-8 forms a heteroheptameric ring around the 3' end of U6 snRNA, consisting of a U rich tract. The Sm core RNP is extremely salt stable; however, LSm-U6 snRNA dissociates at salt concentrations higher than 0.5M, or in the presence of competitor RNA, suggesting that the LSm-U6 complex is less stable [58]. U6 snRNA has no conserved Sm site and does not associate with Sm proteins. Its biogenesis pathway differs in many respects from the U1, U2, U4 and U5snRNP pathways; it is transcribed by RNA polymerase III and capped by  $\gamma$ -monomethyltriphosphate. The 3' end of pre-U6 snRNA is elongated during maturation and subsequently trimmed leaving in most organisms a 2'-3' cyclic phosphate. The enzymes involved in this process are specific for U6 snRNA, and U6 snRNA does not leave the nucleus [61]. Mature U6 snRNA shows nucleoplasmic localization [62]. Experimental evidence suggests that U6 snRNA is present in the cytoplasmic compartment of mouse fibroblast cells [63]. This result suggests that the LSm2-8 complex may act as a nuclear localization signal, but the cytoplasmic localization of the U6 snRNP is highly questionable. The actual function of the LSm 2-8 complex associated with U6 snRNA appears to be connected to U6 snRNP assembly and function. Mutants with decreased levels of LSm2-8 show splicing defects correlating with a reduced level of U6 snRNA. How the LSm2-8 complex affects U6 snRNP remains unclear. One possibility is that LSm proteins facilitate conformational rearrangements during the splicing cycle, U4/U6 annealing and formation of U4/U6/U5 tri-snRNP [64].

### **2.4. Role of LSm proteins in protecting mRNA 3' end termini from degradation**

LSm proteins have additional roles apart from splicing. Yeast strains which lack LSm1-7p fail to grow at higher temperatures, and accumulate mRNA shortened at the 3' end by 20-30 nucleotides. The simplest model proposes that LSm1-7 complex binds to the mRNA and sterically inhibits endo and exo-nucleases. Nuclear LSm2-8 binds to the U6 snRNA 3' end, suggesting, that LSm2-8 protects the 3' end of U6 snRNA from degradation [65].

### **2.5. Role of LSm oligomer proteins in U8 snoRNP organization**

U8 snoRNP is required for processing of 5.8S and 28S rRNAs, which together with the 5S rRNA build up the large ribosomal subunit. In *Xenopus* extract, LSm2, 3, 4, 6, 7, and 8 are bound as hetero hexamer to U8 snoRNA on the conserved third stem-loop sequences [66].

### **2.6. LSm oligomers as part of U7 snRNP**

Maturation of the non-polyadenylated histone mRNAs 3' ends occurs by endonucleolytic cleavage mediated by U7 snRNP [67]. U7 snRNA contains a non-canonical Sm site. Purified U7 snRNP lacks D1 and D2 proteins but has LSm10 (14kDa) and LSm 11 (50kDa) instead [68].

## 2.7. LSm protein oligomers in mRNA degradation

Yeast two hybrid assays reveal multiple interactions between the eight LSm proteins, suggesting the existence of more than one LSm protein complex. Each human LSm protein is capable of interacting with multiple other LSm proteins and splicing factors, like prp24, prp4, and SmD1 [69]. Coprecipitation experiments demonstrated that LSm1p (LSmXp, is the nomenclature for yeast LSm proteins) together with LSm2p-LSm7p forms a new seven-subunit complex [70, 71]. The LSm complex LSm1-7 plays a role in mRNA degradation [72], and LSm2-8 has a role in the stabilization of U6 snRNP. These two protein complexes thus have very different functions. LSm1p mutants accumulate full length capped transcripts, but mutations on LSm1p do not stabilize mRNA containing premature stop codons, suggesting that the LSm1-7 complex is not involved in NMD [71]. The function of the LSm1-7 complex is most likely to interact with the mRNA substrate and accelerate decapping. Decapping is mediated by a decapping enzyme that is consisting of Dcp1a, Dcp1b, and the catalytic subunit Dcp2. The LSm1-7 proteins are localized in discrete cytoplasmic foci. The foci contain key decapping factors required for 5'-3' mRNA degradation. Coexpression of LSm proteins increases the number of foci. The cytoplasmic foci contain LSm1-7 proteins [73]. LSm1 and LSm8 are closely related to each other, and to the SmB protein. The 33 C terminal amino acids of LSm1 are necessary but not sufficient for proper cellular localization of hLSm1 [73]. Finally it has been demonstrated [73] that the foci are actual degradation centers, where mRNA degradation occurs. This suggests that the cytoplasm of cells is more organized than previously thought. Bacterial Hfq protein (pdb 1hk9) is able to chaperone RNA-RNA interactions similarly like LSm proteins ability to chaperone RNA/protein interactions and protect the 3' end of a transcript from exonucleolytic decay while encouraging degradation through other pathways [74].

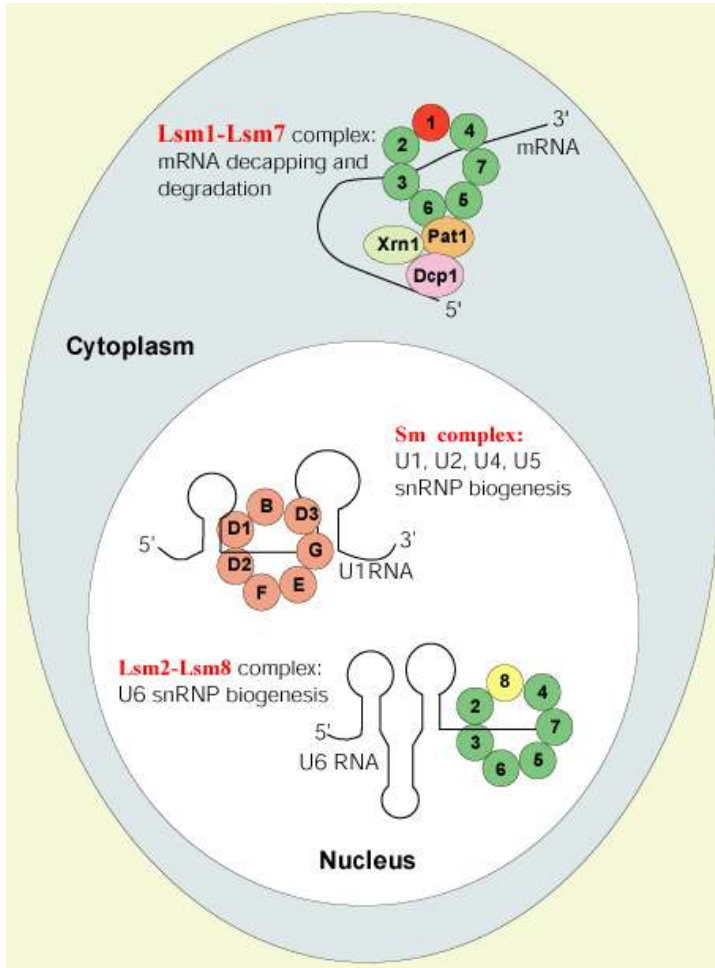
## 2.8. LSm proteins in the processing of pre-tRNAs

It has been reported that depletion of LSm proteins in yeast leads to strong accumulation of unspliced tRNA species. The absence of LSm proteins most probably alters the pattern of processing intermediate [75].

## 2.9. LSm 2-7 complex associated with snR5

An LSm2-7 hexameric complex is found to be associated with snR5 in *Saccharomyces cerevisiae*. This RNA is a member of the class of snoRNAs that function in pseudouridylation of rRNA. The SnR5 associated LSm complex may be a hexamer, but it is not excluded that one LSm protein in this complex is present in more than one copy, or an as yet unidentified yeast protein associates with LSm2-7, thereby closing the heptameric ring [76]. LSm2-8 interacts with external 3' sequences on U6 snRNA. The LSm1-7 complex interacts with 3' UTR mRNA but there could well be secondary structure elements between the LSm1-7 binding site and the mRNA 3' end. U8snoRNA and snR5 bind LSm proteins via internal RNA sequences, suggesting that LSm rings can assemble onto the RNA. LSm proteins have a role in the biogenesis and function of at least a subset of nucleolar RNAs. One possibility is that LSm proteins assist base

pairing of snoRNAs with their rRNA targets, to conduct pseudouridylation and ribose modifications.



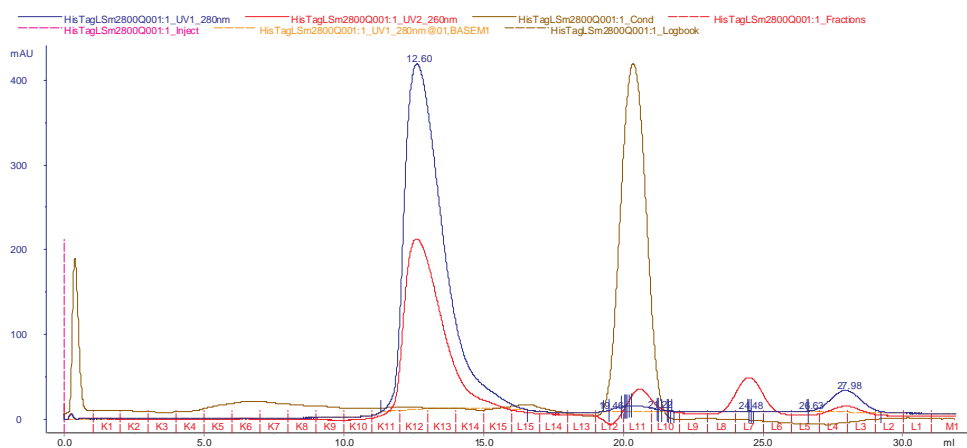
**Figure 10.** Three different heptameric complexes contain Sm or LSm proteins, reprinted from [71]

A particularly interesting example of forming higher order complexes-oligomers is the Sm/LSm protein family (whose various complexes are described above), whose members are engaged in a variety of RNA processing events, forming complexes which differ sometimes only by one out of seven subunits. Another important aspect of the Sm/LSm protein family is that these proteins never occur in isolation; for proper functioning they require complex formation. Hence, the way to better understand Sm/LSm protein function is to study Sm/LSm complexes. It is difficult to determine the connection between the oligomeric state of a given

protein and its function *in vivo*. Reconstitution *in vitro* of two human LSm complexes with seven subunits each, LSm1-7 and LSm 2-8, has been described [77, 78, 79]. The LSm2-8 complex binds to the 3'-end of U6 snRNA in the cell nucleus. The closely related cytoplasmic LSm1-7 complex binds to the 3'UTR of mRNAs destined for degradation. Remarkably, LSm1-7 differs from LSm2-8 only by the exchange of one single subunit, LSm1 for LSm8.

Sequence comparisons of the yeast LSm protein family indicate that each canonical Sm protein has a corresponding LSm protein with the exception of SmB, which aligns almost equally well with LSm1 and LSm8. Based on sequence comparisons co expression vectors encoding the homologs of SmD1D2, LSm23, of SmD3B, LSm48, and of SmEFG, LSm567 were constructed and proteins were expressed in bacteria [77]. LSm4 and LSm1 were singly over expressed for the reconstitution of LSm1-7 [77].

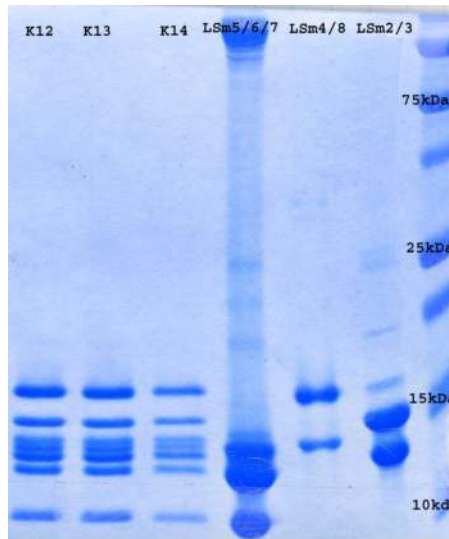
Two heteroheptameric complexes LSm1-7 and LSm2-8 were reconstituted from two heterodimers and one heterotrimer in case of LSm2-8 (LSm2-3, LSm4-8, LSm5-6-7) and one heterotrimer, one heterodimer and two proteins singly expressed (LSm2-3, LSm5-6-7, LSm1, LSm8). Reconstitution of heteroheptamers was achieved by mixing of equimolar amounts of each appropriate protein at 37°C adding 4 M urea in order to disrupt higher order structures, because those proteins have tendency to oligomerize. After incubation, mixture of pure recombinant proteins was dialyzed against native buffer. Mixture was applied on to size exclusion chromatography, followed by the anion exchange chromatography. Last step in purification of homogenous heteroheptamers was size exclusion chromatography (peak profile shown on figure 11., and respective fractions were analyzed on polyacrylamide gel (shown of figure 12).



**Figure 11.** Second size exclusion chromatography step

Negative stain electron micrographs show that reconstituted LSm2-8 has a ring-like architecture with a diameter of about 8 nm. The overall dimensions are similar to those previously observed for the native LSm2-8 complex isolated from HeLa cell nuclear extract (8 nm) and

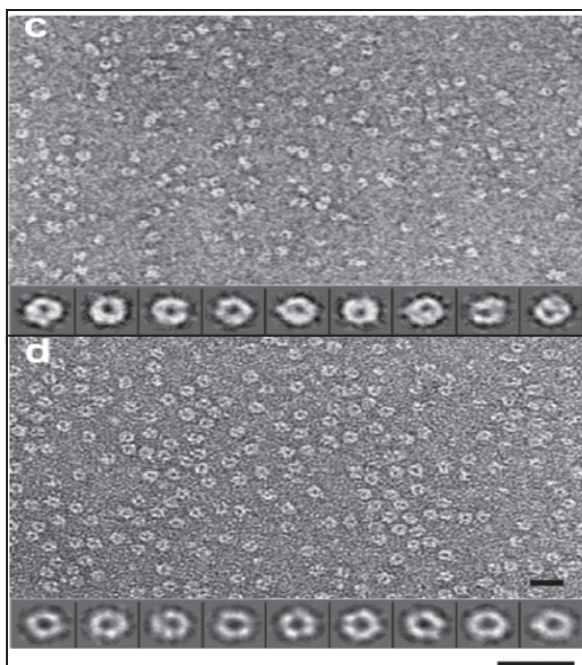




**Figure 12.** SDS PAGE gel

core snRNP domain [58]. The central cavity observed for the recombinant LSm2-8 complex is larger than in the native LSm2-8 complexes (3 vs. 2 nm, respectively). The LSm1-7 rings appear to be slightly smaller, measuring ~ 7 nm across with a pore diameter of less than 1.5 nm. Thus, recombinant LSm1-7 and LSm2-8 complexes are similar to one another and to the native Sm/LSm complexes at this level. In all LSm co-crystal structures solved with RNA oligonucleotides, the RNA molecules mainly wrap around the rim of the pore.

One of the methods which can be used for the identification and characterization of the RNA binding proteins is the electrophoretic mobility shift assay (EMSA). The basis of this method is the change in the electrophoretic mobility of a nucleic acid molecule upon binding to a protein or another molecule. Initially a labeled RNA, which contains the binding sequence, is incubated with a sample containing the RNA binding proteins and the mixture is then analyzed on a non-denaturing gel. The unbound RNA will have a characteristic electrophoretic mobility. Functionality of reconstituted LSm2-8 and LSm1-7 complexes has been demonstrated using this assay *in vitro* [77]. That oligomer complexes are functional *in vivo* has been shown [77], by injecting fluorescently labeled complexes into cytoplasm of living cells. They localized in expected cellular compartment, namely LSm 2-8 took nuclear localization and LSm1-7 complex remained in the cytoplasm. The structure-function relationships within the Sm/LSm protein family reflect three major interconnected features which illustrate why it is so important to solve the structures of Sm/LSm hetero-oligomeric complexes: First, Sm/LSm protein function is in general strictly dependent on complex formation. This holds for RNA binding, Sm/LSm-protein containing RNP biogenesis, interaction with non-Sm protein effector proteins, and RNA processing activity. The required interaction interfaces are apparently always three dimensional structural sites generated from several Sm/LSm subunits. High resolution



**Figure 13.** Electron micrographs of reconstituted complex LSm1-7 (c) and LSm2-8 (d) (reprinted from reference 77).

structural information is clearly required to explain the molecular basis for this phenomenon. Second, exchange of only one or two subunits from one to another heterooligomeric (mostly heptameric) Sm/LSm complex changes its whole biology (see above). How such subtle structural changes can have these very large functional effects can only be addressed by solving the crystal structures of the respective complexes. Lastly, the ability of individual Sm/LSm proteins to assemble with different homologous binding partners to form architecturally very similar, yet functionally diverse complexes argues for a very fine balance between flexibility and specificity for the respective Sm-Sm interactions. Clearly, in order to understand the “molecular recognition code” governing the specificity balance mentioned above, more structural information on such interactions is indispensable. Recently crystal structure of *Saccharomyces cerevisiae* LSm2-8 complex bound to U6 snRNA had been determined (pdb code 4M7D) [80].

## Acknowledgements

This work was supported by the grant No. 172001 from the Ministry of Science and Education, Republic of Serbia.

## Author details

Bozidarka L. Zaric

ICTM – Centre of Chemistry, University of Belgrade, Belgrade, Serbia

## References

- [1] Mark G. Caprara and Timothy W. Nilsen (2000) RNA: Versatility in form and function, *Nature Structural Biology*, (7): 831-835
- [2] Steven Buckingham (2003) The Major World of microRNAs, Horizon symposia, 1-4, <http://www.nature.com/horizon/rna/background/understanding.html>
- [3] Molecular Biology of the Cell (2001) Bruce Alberts, Dennis Bray, Julian Lewis, Martin Raff, Keith Roberts, James D. Watson, Chapter 6, Third edition, Garland Publishing, Inc. New York and London.
- [4] Pietzsch J. (2003) Understanding the RNAissance, Horizon Symposia, 1-4, <http://www.nature.com/horizon/rna/background/understanding.html>
- [5] Hartman E. and Hartmann R. (2003) The enigma of ribonuclease P evolution, *Trends in Genetics*, 19(10): 561-568.
- [6] Altman S., Ribonuclease P, Chapter 14 in *The RNA World* (Second edition), Cold Spring Harbor, New York.
- [7] Tollervey D. and Kiss T. (1997) Function and synthesis of small nucleolar RNAs. *Current Opinion Cell Biology*, (3): 337-342.
- [8] Gerbi S. A. and Lange T. S. (2002) All Small Nuclear RNAs (snRNAs) of U4/U6U5 tri-snRNP Localize to Nucleoli, Identification of the Nucleolar Localization Element of U6 snRNA., *Molecular Biology of the Cell*, (13): 3123-3137.
- [9] Trotta C. R., and Abelson J., (2000) tRNA splicing: An RNA World Add-on or an Ancient Reaction? *RNA World*, second edition, Cold Spring Harbor Laboratory Press.
- [10] Wong JM, Collins K. Telomere maintenance and disease (2003) *Lancet*;362(9388): 983-8.
- [11] Bortvin A. PIWI-interacting RNAs (piRNAs)-a mouse testis perspective. *Biochemistry (Mosc)*. (2013);78(6):592-602.
- [12] Burroughs AM, Ando Y, Aravind L. (2013) New perspectives on the diversification of the RNA interference system: insights from comparative genomics and small RNA sequencing. (2013) *Wiley Interdiscip Rev RNA* Dec 5. doi: 10.1002/wrna.1210.
- [13] Steitz J. A. (1998) ``Snurps``, *Scientific American*, 36-41.

- [14] Pendrak ML, Roberts DD. (2011) Ribosomal RNA processing in candida albicans RNA. (2011) 17(12):2235-48.
- [15] Varani G. and Pardi A. (1994) Structure of RNA, Chapter 1, in *Protein -RNA interactions*, Nagai and Mattaj, IRL press, Oxford
- [16] Varani G. and Kiyoshi N. (1998) RNA Recognition by RNP proteins during RNA processing, *Annu. Rev. Biophys. Biomol. Struct.* 27: 402-45.
- [17] Pérez-Cañadillas J. M. and Varani G. (2001) Recent advances in RNA-protein recognition, *Current Opinion Structural Biology*, 11, 53-58.
- [18] Maniatis T and Reed R. (2002) An extensive network of coupling among gene expression machines, *Nature*, 416, 499-504
- [19] Proudfoot NJ, Furger A, Dye MJ.,(2002) Integrating mRNA processing with transcription, *Cell*, 108 (4), 501-512.
- [20] Kornbliht AR, de la Mata M, Fededa JP, Munoz MJ, Nogues G. (2004) Multiple links between transcription and splicing, *RNA*, 10, 1489-1498.
- [21] Cougot N, van Dijk E, Babajko S, Seraphin B. (2004) Cap-Tabolism, *Trends in Biochemical Sciences*, 29(8): 436-44.
- [22] Jurica Melissa and Melissa J. Moore: Capturing splicing complexes to study structure and mechanism (2002) *Methods*, 28: 336-345.
- [23] Nilsen Timothy: The spliceosome: The most complex macromolecular machine in the cell? (2003) *Bioassays*, 25: 1147-1149
- [24] Will CL, Schneider C, Hossbach M, Urlaub H, Urlaub R, Elbashir S, Tuschl T, Luhrmann R (2004) The human 18S U11/U12 snRNP contains a set of novel proteins not found in the U2-dependent spliceosome. *RNA*, 10(6):929-41.
- [25] Yi-Tao Yu, Scharl E, Smith C. M., and A. Steiz J. A. The growing world of small nuclear ribonucleoproteins, *RNA WORLD*, second edition, Cold Spring Harbor Laboratory Press
- [26] Balakin AG, Smith L, Fournier MJ. (1996) The RNA world in the nucleolus: Two major families of small RNAs defined by different box elements with related function, *Cell*, 86: 823-834.
- [27] Will C. L and Lührmann R. (2001) Spliceosomal U snRNP biogenesis, structure and function, *Current Opinion in Cell Biology*, (13): 290-301.
- [28] Malca H, Shomron N, Ast G (2003) The U1 snRNP Base Pairs with the 5' Splice Site within a Penta-snRNP complex, *Molecular and Cellular Biology*, 23(10): 3442-3455.
- [29] Pascolo E. and Seraphin B. (1997) The Branchpoint Residue Is Recognized during Commitment Complex Formation before Being Bulged out of the U2snRNA-pre-mRNA Duplex, *Molecular and Cellular Biology*, 17(7): 3469-3476.

- [30] Stevens SW, Ryan DE, Ge HY, Moore RE, Young MK, Lee TD, Abelson J (2002) Composition and functional characterization of the yeast spliceosomal penta-snRNP, *Molecular Cell*, 9: 31-44.
- [31] Chen Ching-Yi, Roberto Gherzi, Shao-En Ong, Edvard L. Chan, Reinout Raijmakers, Ger J. M. Pruijijn, Georg Stoecklin, Christof Moroni, Matthias Mann and Mitchael Karin (2001) AU binding proteins recruits the exosome to degrade ARE-containing mRNAs. *Cell*, 107(4) :451-64.
- [32] Tucker M. and Parker R. (2000) Mechanisms and Control of mRNA decapping in *Saccharomyces cerevisiae*, *Annual Rev. Bioch.*, 69: 571-95.
- [33] Lejeune Fabrice, Xiaojie Li and Lynne E. Maquat (2003) Nonsense-Mediated mRNA Decay in Mammalian Cells Involved Decapping, Deadenylation and Exonucleolytic Activities, *Molecular Cell*, 12: 675-687.
- [34] Miles F. Wilkinson and Ann-Bin Shyu (2002) RNA Surveillance by nuclear scanning? *Nature Cell Biology*, 4: 144-147.
- [35] Luhrmann, R., Kastner, B. and Bach, M. (1990) Structure of spliceosomal snRNPs and their role in pre-mRNA splicing, *Biochim. Biophys. Acta*, 1087: 265-292.
- [36] Lerner, E. A., Lerner, M. R., Hardin, J. A., Janeway, C. A and Steitz, J. A. (1981) Monoclonal antibodies to nucleic acid containing cellular constituents: probes for molecular biology and autoimmune disease. *PNAS*, 78: 2737-2741.
- [37] Hermann H, et al., (1995) snRNP Sm proteins share two evolutionary conserved sequence motifs which are involved in Sm protein-protein interaction, *EMBO Journal*, 14(9): 2076-2088.
- [38] van Dam A, Winkel I, Zijlstra-Baalbergen J, Smeenk R, Cuypers HT. (1989) Cloned human snRNP proteins B and B' differ only in their carboxy-terminal part. *The EMBO Journal*. 8(12): 3853-60.
- [39] McAllister, G., Amara, S. G and Lerner, M. R. (1988) Tissue-specific expression and cDNA cloning of small nuclear ribonucleoprotein-associated polypeptide N. *PNAS, USA*, 85(14): 5296-300.
- [40] Salgado-Garrido, J. ; Bragado-Nilsson, E. ; Kandels-Lewis, S. ; Seraphin, B. (1999) Sm and Sm-like proteins assemble in two related complexes of deep evolutionary origin, *The EMBO Journal*, 8(12), 3451-3462.
- [41] Kambach, C. ; Walke, S. ; Young, R. ; Avis, J. M. ; de la, Fortelle E. ; Raker, V. A. ; Luhrmann, R. ; Li, J. ; Nagai, K. (1999) Crystal Structures of Two Sm protein Complexes and their Implications for the Assembly of the Spliceosomal snRNPs, *Cell*, 96:375-387.
- [42] Raker, V. A., Plessel Gabriele and Lurmann Reinhard (1996) The snRNP core assembly pathway: identification of stable core protein heteromeric complexes and an snRNP subcore particle in vitro, *The EMBO Journal*, 15(9): 2256-2269.

- [43] Raker,V. A. ; Hartmuth,K. ; Kastner,B. ; Luhrmann,R. (1999) Spliceosomal U snRNPcore assembly: Sm proteins assemble onto an Sm site RNA Nonanucleotide in a specific and Thermodynamically Stable Manner, *Molecular and Cell Biology*, 19(10): 6554-6565.
- [44] Walke,S. ; Bragado-Nilsson,E. ; Seraphin,B. ; Nagai,K. (2001) Stoichiometry of the Sm proteins in yeast spliceosomal snRNPs supports the heptamer ring model of the core domain, *J. Mol. Biol.*, 308:49-58.
- [45] R. Bordonne (2001) Functional characterization of nuclear localization signals in yeast Sm proteins, *Molecular Cell. Biology*, 20(21):7943-54.
- [46] Stark H, Dube P, Luhrmann R, Kastner B. (2001) Arrangement of RNA and proteins in the spliceosomal U1 small nuclear ribonucleoprotein particle, *Nature*, (409): 539-542.
- [47] SchrankB., GötzR., Gunnensen J. M., Urej. M., Toyka K. V., Smith A., and Sendtner M. (1997) Inactivation of the survival motor neuron gene, a candidate for human spinal muscular atrophy, leads to massive cell death in early mouse embryos, *PNAS*. 94: 9920-9925.
- [48] BrahmS,H. ; Meheus,L. ; de,Brabandere,V; Fischer,U. ; Luhrmann,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.
- [49] Friesen,W. J. ; Wyce,A. ; Paushkin,S. ; Abel,L. ; Rappsilber,J. ; Mann,M. ; Dreyfuss,G. (2002) A Novel WD repeat protein Component of the Methylosome Binds Sm proteins, *Journal of Biol. Chemistry*, 277: 8243-8247.
- [50] GubitZ AK, Feng W, Dreyfuss G. (2004) The SMN complex, *Experimental Cell Research* 296, 51-56.
- [51] Yong J, Pellizzoni L, Dreyfuss G. (2002) Sequence –specific interaction of U1 snRNA with the SMN complex, *The EMBO Journal*, 21(5):1188-1196.
- [52] 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  $\beta$ , *Human Molecular Genetics*, 11(15): 1785-1795.
- [53] Paraskeva E, Izaurralde E, Bischoff FR, Huber J, Kutay U, Hartmann E, Luhrmann R, Gorlich D. (1999) CRM-1 mediated Recycling of Snurportin1 to the Cytoplasm, *The Journal of Cell Biology*, 145(2): 255-264.
- [54] Massenet,S. ; Pellizzoni,L. ; Paushkin,S. ; Mattaj,I. W. ; Dreyfuss,G. et al (2002) The SMN Complex is Associated with snRNPs throughout their Cytoplasmic Assembly Pathway, *Molecular and Cellular Biology*, 22(18):6533-6541.

- [55] Cameron Mura, Peter S. Randolph, Jennifer Patterson and Aaron e. Cozen (2013) Archaeal and eukaryotic homologs of Hfq A structural and evolutionary perspective on Sm function. *RNA Biology* 10:4, 636–651.
- [56] Anna Golisz, Pawel J. Sikorski, Katarzyna Kruszkka and Joanna Kufel (2013) Arabidopsis thaliana LSM proteins function in mRNA splicing and degradation *Nucleic Acid Research* 41:6232-6249.
- [57] Camasses, A. ; Bragado-Nilsson, E. ; Martin, R. ; Seraphin, B. ; Bordonne, R. (1998) Interactions within yeast Sm core complex: from proteins to amino acids., *Mol. Cell. Biol.* 18: 1956-1966.
- [58] Achsel, T. ; Brahm, H. ; Kastner, B. ; Bachi, A. ; Wilm, M. ; Luhrmann, R. (1999) A doughnut-shaped heteromer of human Sm-like proteins binds to the 3' end of U6 snRNA, thereby facilitating U4/U6 duplex formation in vitro, *The EMBO Journal*, 18(20):5789-5802.
- [59] Eystathiou T, Peebles CL, Hamel JC, Vaughn JH, Chan EK. (2002) Autoantibody to the hLSm4 and the Heptameric LSm complex in Anti-Sm Sera, *Arthritis and Rheumatism*, 46(3), 726-734.
- [60] Hirsch, E. ; Oohashi, T. ; Ahmad, M. ; Stamm, S. ; Fassler, R. (2000) Peri-Implantation Lethality in Mice Lacking the Sm Motif-containing protein LSm4, *Molecular and Cellular Biology*, 20 (3): 1055-1062.
- [61] Vankan, P. ; McGuigan, C. ; Mattaj, J. W. (1990) Domains of U4 and U6 snRNAs required for snRNP assembly and splicing complementation, *The EMBO Journal*, 9: 3397-3404.
- [62] Lange TS, Gerbi SA. (2000) Transient Nucleolar Localization of U6 snRNA in *Xenopus* Laevis Oocytes, *Molecular Biology of the Cell*, 11:2419-2428.
- [63] Fury, M. G. and Zieve, G. W. (1996) U6 snRNA maturation and stability, *Experimental Cell Research*, 228: 63-69.
- [64] Mayes, A. E. ; Verdone, L. ; Legrain, P. ; Beggs, J. D. (1999) Characterization of Sm-like proteins in yeast and their association with U6 snRNA, *The EMBO Journal*, 18 (15): 4321-4331.
- [65] Weihai He and Parker R. (2001) The Yeast Cytoplasmic LSmI/pat1p complex protects mRNA 3' termini from partial degradation, *Genetics*, 158:1445-1455.
- [66] Tomasevic, N. ; Peculis, B. A. (2002) *Xenopus* LSm proteins Bind U8 snoRNA via an Internal Evolutionarily Conserved Octamer Sequence, *Molecular and Cellular Biology*, 22(12): 4101-4112.
- [67] Muller B. and Schumperli D. (1997) The U7 snRNP and the hairpin binding protein: Key players in histone mRNA metabolism, *Seminars in cell and developmental biology*, 8: 567-577.

- [68] Pillai RS, Grimm M, Meister G, Will CL, Luhrmann R, Fischer U, Schumperli D. (2003) Unique Sm core structure of U7 snRNPs: assembly by a specialized SMN complex and the role of a new component, LSm11 in histone RNA processing, *Genes and Development* 17: 2321-2333.
- [69] Fromont-Racine M, Mayes AE, Brunet-Simon A, Rain JC, Colley A, Dix I, Decourty L, Joly N, Ricard F, Beggs JD, Legrain P.. (2000) Genome-wide protein interaction screens reveal functional networks involving Sm-like proteins, *Yeast*, 95-110.
- [70] Tharun,S. ; He,W. ; Mayes,A. E. ; Lennertz,P. ; Beggs,J. D. ; Parker,R. (2000) Yeast Sm-like proteins function in mRNA decapping and decay, *Nature*, 404: 515-518.
- [71] Bouveret,E. ; Rigaut,G. ; Shevchenko,A. ; Wilm,M. ; Seraphin,B. I (2000) A Sm-like protein complex that participates in mRNA degradation, *The EMBO Journal*, 19 (7): 1661-1671.
- [72] Panone B. K. and Wolin S. L. (2000) RNA degradation: Sm-like proteins WRING the neck of mRNA, *Current Biology*, 10: 478-482.
- [73] Ingelfinger D, Arndt-Jovin DJ, Luhrmann R, Achsel T. (2002) The human LSm1-7 proteins colocalize with the mRNA-degrading enzymes Dcp1/2 and Xrn1 in distinct cytoplasmic foci, *RNA*, 8: 1489-1501.
- [74] Cougot N, van Dijk E, Babajko S, Seraphin B. (2004) Cytoplasmic foci are sites of mRNA decay in human cells, *The Journal of Cell Biology*, 165(1): 31-40.
- [75] Carol J. Wilusz and Jeffrey Wilusz (2013) Lsm proteins and Hfq Life at the 3' end". *RNA Biology* 10:4, 592-601.
- [76] Kufel J, Allmang C, Verdone L, Beggs JD, Tollervey D. (2002) LSm proteins Are Required for Normal Processing of the Pre-tRNA and Their Efficient Association with La-Homologous Protein Lhp1p, *Molecular and Cellular Biology*, 22(14):5248-56.
- [77] Fernandez CF, Pannone BK, Chen X, Fuchs G, Wolin SL. (2004) An LSm2-7 Complex in *Saccharomyces cerevisiae* Associates with the Small Nucleolar RNA snR5, *Molecular Biology of the Cell*, (15):2842-2852.
- [78] Zaric B., Chami M., Remigy H., Engel A., Ballmer-Hofer K., Winkler FK and Kambach C. (2005) Reconstitution of two Recombinant LSm protein complexes reveals aspects of their architecture, assembly and function, *Journal of Biological Chemistry*, 280: 16066-16075.
- [79] Karen C. M. Moraes, Naomi Bergman, Bozidarka Zaric, Christian Kambach, Carol J. Wilusz & Jeffrey Wilusz, (2007) Lsm proteins bind to and stabilize RNAs containing poly(A) tracts *Nature Structural Molecular Bioogy.* ;14(9):824-31.
- [80] Bozidarka Zaric and Christian Kambach (2008) "Reconstitution of recombinant human LSm complex for biochemical, biophysical and cell biological studies". *Methods in Enzymology* 448:57-74.



- [81] Zhou L, Hang J, Zhou Y, Wan R, Lu G, Yin P, Yan C, Shi Y(2013) Crystal structures of the Lsm complex bound to the 3' end sequence of U6 small nuclear RNA. *Nature* Nov 17. doi: 10.1038/nature12803.

