Spliceosomal UsnRNP biogenesis, structure and function

Cindy L Will* and Reinhard Lührmann†

Significant advances have been made in elucidating the biogenesis pathway and three-dimensional structure of the UsnRNPs, the building blocks of the spliceosome. U2 and U4/U6•U5 tri-snRNPs functionally associate with the pre-mRNA at an earlier stage of spliceosome assembly than previously thought, and additional evidence supporting UsnRNA-mediated catalysis of pre-mRNA splicing has been presented.

Addresses
Max Planck Institute of Biophysical Chemistry, Department of Cellular Biochemistry, Am Fassberg 11, 37077 Göttingen, Germany.
*e-mail: cwill1@gwdg.de
†e-mail: reinhard.luehrmann@mpi-bpc.mpg.de

Current Opinion in Cell Biology 2001, **13**:290–301

0955-0674/01/$ – see front matter
© 2001 Elsevier Science Ltd. All rights reserved.

Abbreviations

CBC	cap-binding complex
NLS	nuclear localization signal
PHAX	phosphorylated adaptor for RNA export
RRMs	RNA recognition motifs
SMN	survival of motor neurons
UsnRNPs	uridine-rich small nuclear ribonucleoproteins

Introduction

Pre-mRNA splicing, the removal of non-coding introns from mRNA precursors, is a pre-requisite for the expression of most eukaryotic genes. This essential reaction proceeds in two steps. 5′ splice site cleavage and ligation of the intron's 5′ end to the so-called branch site occur concomitantly in the first step, and 3′ splice site cleavage (with the resulting excision of the intron) and ligation of the 5′ and 3′ exons take place in the second.

The spliceosome, the complex macromolecular machinery that catalyzes pre-mRNA splicing, is formed from several RNP subunits, termed uridine-rich small nuclear ribonucleoproteins (UsnRNPs), and numerous non-snRNP splicing factors [1]. Each UsnRNP particle consists of a UsnRNA molecule complexed with a set of seven Sm or Sm-like proteins and several particle-specific proteins [2]. The major spliceosomal UsnRNPs U1, U2, U4, U5 and U6 are responsible for splicing the vast majority of pre-mRNA introns (so-called U2-type introns). A group of less abundant snRNPs, U11, U12, U4atac and U6atac, together with U5, are subunits of the so-called minor spliceosome that splices a rare class of pre-mRNA introns, denoted U12-type [3]. The UsnRNPs carry out a number of essential functions during splicing. Via the interactions of their RNA and protein components with the pre-mRNA, they mediate the recognition and subsequent pairing of the 5′ and 3′ splice sites of an intron. In addition, they are responsible for creating the three-dimensional (3D) structure required for the formation of the spliceosomes' two active sites, one for each of the two reaction steps. Components of the UsnRNPs also appear to catalyze the two transesterification reactions leading to excision of the intron and ligation of the 5′ and 3′ exons. Here we describe recent advances in our understanding of snRNP biogenesis, structure and function, focusing primarily on the major spliceosomal UsnRNPs from higher eukaryotes.

Identification of a novel UsnRNA export factor

UsnRNP biogenesis is a complex process, many aspects of which remain poorly understood. Although less is known about the maturation process of the minor U11, U12 and U4atac UsnRNPs, it is assumed that they follow a pathway similar to that described below for the major UsnRNPs (see Figure 1). The UsnRNAs, with the exception of U6 and U6atac (see below), are transcribed by RNA polymerase II as snRNA precursors that contain additional 3′ nucleotides and acquire a monomethylated, m7GpppG (m7G) cap structure. These pre-UsnRNAs must first be exported to the cytoplasm where snRNP assembly is initiated. U1, U2, U4 and U5 snRNA export is dependent on a number of factors, including the snRNA's m7G cap, the m7G cap-binding complex (CBC), the export receptor CRM1/Xpo1 and RanGTP [4–7].

In the past year, an additional factor, PHAX (phosphorylated adaptor for RNA export), which is required specifically for UsnRNA export, has been identified [8••]. PHAX mediates the interaction between complexes containing CBC/UsnRNA and CRM1/RanGTP to form a UsnRNA export complex (Figure 1). Furthermore, it is phosphorylated in the nucleus, an event that is required for the formation of the UsnRNA export complex and, subsequent to export, PHAX is dephosphorylated in the cytoplasm, leading to disassembly of this complex. Thus, together with GTP hydrolysis by Ran, dephosphorylation of PHAX is thought to ensure the directionality of UsnRNA export.

UsnRNP Sm core assembly

Following their export to the cytoplasm, UsnRNA precursors interact in an ordered, stepwise manner with seven Sm proteins, B/B′, D3, D2, D1, E, F and G to form the snRNP Sm core structure (Figure 1). The demonstration of a link between the disease spinal muscular atrophy (SMA) and mutations in a UsnRNP assembly factor has fueled interest in this step of UsnRNP biogenesis in recent years. The Sm proteins form three distinct heteromeric complexes prior to their interaction with the highly conserved Sm site ($PuAU_{4\text{-}6}GPu$ flanked by two stem-loop structures) of the U1, U2, U4 and U5 snRNAs [9]. *In vitro* experiments indicate that the individual complexes do not bind stably on their own; rather the D1–D2 and E–F–G complexes bind concomitantly to the Sm site, forming a

so-called subcore complex, and assembly is completed by the association of B–D3 or B′–D3 (B and B′ differ only in their last 11 amino acids and are encoded by a single gene) [9]. In vertebrate cells, survival of motor neurons (SMN), the SMA disease gene product, facilitates assembly of the U1, U2, U4 and U5 core snRNPs *in vivo* by a currently ill-defined mechanism [10–12].

SMN is present in a protein complex that contains SIP1/Gemin2, Gemin3 and the newly characterized Gemin4 that, like SMN, directly interacts with several of the Sm proteins [10–13]. Upon oligomerization, this complex associates transiently with UsnRNPs in the cytoplasm, dissociating at some point before their import into the nucleus. Recent studies have demonstrated that the carboxy-terminal RG dipeptide repeat domains of SmD1 and SmD3 interact with the tudor domain of SMN [14,15]. NMR studies have now revealed the structure of the tudor domain and also more precisely mapped those residues interacting with the carboxy-terminal RG repeats of D1 [15]. Interestingly, arginines in the RG dipeptides of SmD1 and SmD3 are symmetrically dimethylated, raising the possibility that this unusual modification may modulate the interaction of SMN and thus regulate UsnRNP Sm core assembly [16].

In budding yeast, considerably less is known about UsnRNP biogenesis, and it is not clear whether this process involves a cytoplasmic phase. Brr1, which is distantly related to SIP1/Gemin2, functions in yeast UsnRNP biogenesis [17], but a yeast equivalent of SMN has not been identified. Studies by Xue *et al.* [18•] now reveal that the *Saccharomyces cerevisiae* homologue of the La protein (Lhp1p), an autoantigen known to bind and stabilize nascent Pol III transcripts (including U6 snRNA, see below) [19], also binds to yeast U1, U2, U4 and U5 snRNA precursors and facilitates UsnRNP assembly *in vivo* [18•]. Whether La functionally replaces SMN in yeast and whether it contributes to the assembly of U1, U2, U4 and U5 snRNPs in higher eukaryotes is presently not known.

Subsequent to UsnRNP Sm core assembly, the m7G cap of the snRNA is converted to the 2,2,7-tri-methylated guanosine (m3G) form by a yet to be characterized methyl transferase, and the UsnRNAs undergo 3′ end maturation. Both events are dependent on proper Sm core assembly [20,21]. Precursors of vertebrate snRNAs typically contain less than 20 extra nucleotides, the majority of which are removed prior to re-import into the nucleus by one or more as yet unidentified exoribonucleases ([22] and references therein). In the yeast *S. cerevisiae*, the 3′ ends of pre-UsnRNAs (which often contain considerably longer 3′ extensions) are generated in multiple steps by a processing reaction, and snRNP core assembly is thought to occur concomitantly with this process [21,23]. Although alternative pathways that bypass RNAse III appear to exist, UsnRNA 3′ end maturation in yeast involves cleavage by RNAse III followed by 3′ trimming by the exosome, a complex of multiple 3′ exoribonucleases [24,25], and/or by members of the RNase D family of exonucleases [26•].

UsnRNP import and assembly

The m3G cap and the Sm core domain form a bipartite UsnRNP nuclear localization signal (NLS) required for the subsequent nuclear import of newly assembled core UsnRNPs ([27] and references therein). The UsnRNA's m3G cap is specifically recognized by Snurportin-1, which also interacts with the general import factor Importin β; both are required for UsnRNP import ([28,29]; Figure 1). A distinct import factor appears to interact with the NLS formed by the Sm core, but its identity, as well as the molecular nature of the Sm core NLS, remains to be elucidated [28,29]. Recent studies in *S. cerevisiae* have identified basic amino acids in the carboxy-terminal domains of SmB and SmD1 that exhibit nuclear localiz-ation properties and are collectively essential for yeast viability, suggesting that they may comprise the Sm core NLS [30]. However, an alternative role for these regions in snRNP biogenesis or function could not be excluded, and it remains to be established whether UsnRNPs in yeast transit through the cytoplasm; yeast orthologues of PHAX or Snurportin-1, the two specific mediators of vertebrate UsnRNP export and import, for example, have not been identified, consistent with the idea that UsnRNP biogenesis in yeast takes place exclusively in the nucleus [8••,29]. More definitive experiments are still needed to clarify this issue.

At an undefined step before the association of the particle-specific proteins, the UsnRNAs are internally modified at several positions, primarily by pseudouridylation and 2′-O-methylation (reviewed in [31]). In the case of the human U2 snRNP, post-transcriptional modifications are a prerequisite for the assembly of a functional particle [32]. At present the factors responsible for these modifications are largely unknown, and it is not entirely clear whether U1, U2, U4 and U5 modification takes place in the nucleoplasm or cytoplasm; however, data obtained with *Xenopus* oocytes are most consistent with nuclear modification, at least for the U2 snRNA [32].

Studies by Jády and Kiss [33••] now demonstrate that, similar to the U6 snRNA (see below), pseudouridylation and 2′-O-methylation of specific nucleotides of the human U5 snRNA are generated, at least in part, by a small nucleolar (sno)RNA-guided mechanism [33••]. Two families of snoRNAs, the box C/D or H/ACA snoRNAs, together with their associated proteins, are responsible for post-transcriptional modifications of rRNA; the former direct 2′-O-methylation exclusively and the latter, pseudouridylation (reviewed in [34]). Intriguingly, in the case of the U5 snRNA, the newly identified U85 snoRNP possesses dual activity, directing 2′-O-methylation and pseudouridylation at positions 45 and 46, respectively. Consistent with its dual function, U85 appears to be a hybrid snoRNP, possessing the sequence motifs characteristic of box C/D and H/ACA snoRNAs and containing proteins specific for each

Figure 1

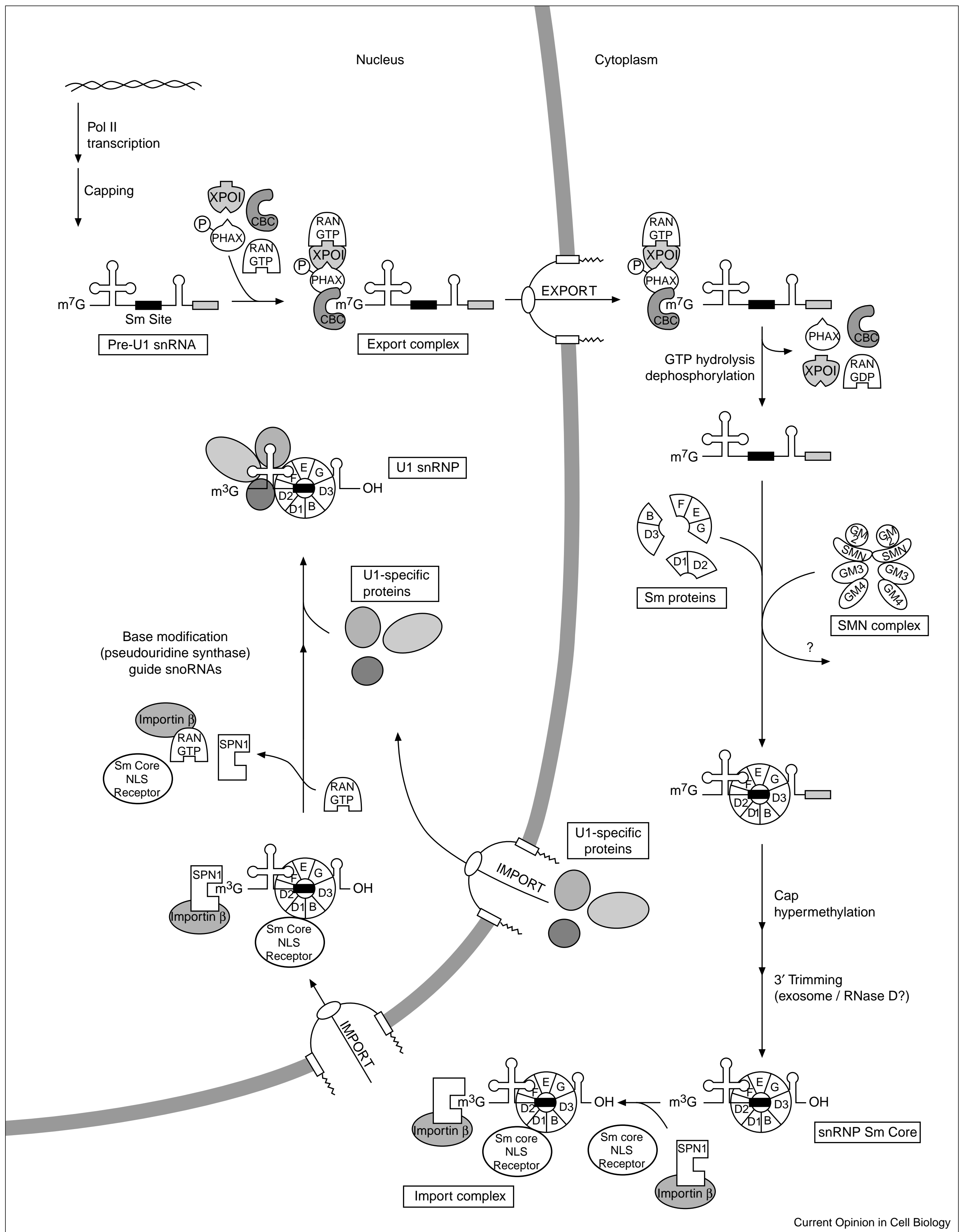


Figure 1 legend

The UsnRNP biogenesis pathway. A cartoon model of the U1 snRNP maturation pathway is shown. Subsequent to transcription and capping, the pre-U1 snRNA becomes complexed with several factors, including the CBC, the newly identified phosphorylated adaptor for RNA export (PHAX), XpoI (also denoted CRM1) and Ran-GTP, which together mediate UsnRNA export to the cytoplasm. Subsequent to export, dephosphorylation of PHAX and GTP hydrolysis of Ran-GTP lead to dissociation of this complex. In the following step the Sm proteins interact as pre-formed heteromic complexes with the UsnRNA's Sm site (indicated by a black box) to form the so-called snRNP Sm core. This step is facilitated by the SMN complex, which consists of the survival of motor neuron protein (SMN), Gemin2 (GM2), Gemin3 (GM3) and Gemin4 (GM4); the SMN complex interacts with the Sm proteins. This complex dissociates from the Sm proteins/UsnRNP core at an unknown step prior to UsnRNP nuclear import. Subsequent to cap hypermethylation and 3′ trimming, the snRNP Sm core is imported into the nucleus. Import is mediated by snurportin-1 (SPN1), which binds to the snRNA's m3G cap, Importin β which forms a complex with SPN1, and an unidentified Sm core NLS receptor. After import, these factors dissociate and the U1-specific proteins, which appear to be imported independently of the U1 snRNA, associate with the U1 snRNP Sm core. The UsnRNA is internally modified, apparently at a step prior to specific protein association.

class of snoRNP [33••]. These results also raise the interesting possibility that modification of some U5 nucleotides may take place in the nucleolus. In yeast, Pus1p, a pseudouridine (Ψ) synthase involved in tRNA modification, catalyzes conversion of uridine 44 to Ψ in U2 snRNA [35]. Thus, at least in the case of Ψ-formation, activities in addition to the snoRNPs, appear to contribute to UsnRNA post-transcriptional modification.

The association of the UsnRNP-specific proteins completes the biogenesis of UsnRNPs. This poorly understood process is thought to occur in the nucleus, although there is currently little evidence to pinpoint precisely at what stage it occurs. Consistent with nuclear assembly, the UsnRNP-specific proteins, U1-A and U2-B′, unlike the Sm proteins, can be imported independently into the nucleus (i.e. without first interacting with an snRNA) ([36] and references therein). Interestingly, recent studies suggest that their nuclear import occurs by a novel, as yet uncharacterized, mechanism [36].

U6 biogenesis

Biogenesis of the U6 snRNP, and presumably also that of the U6atac snRNP, differs in many ways from that of the other spliceosomal UsnRNPs. The U6 snRNA is transcribed by RNA polymerase III and acquires a γ-monomethyl cap structure. Assembly of the U6 snRNP is thought to take place entirely in the nucleus.

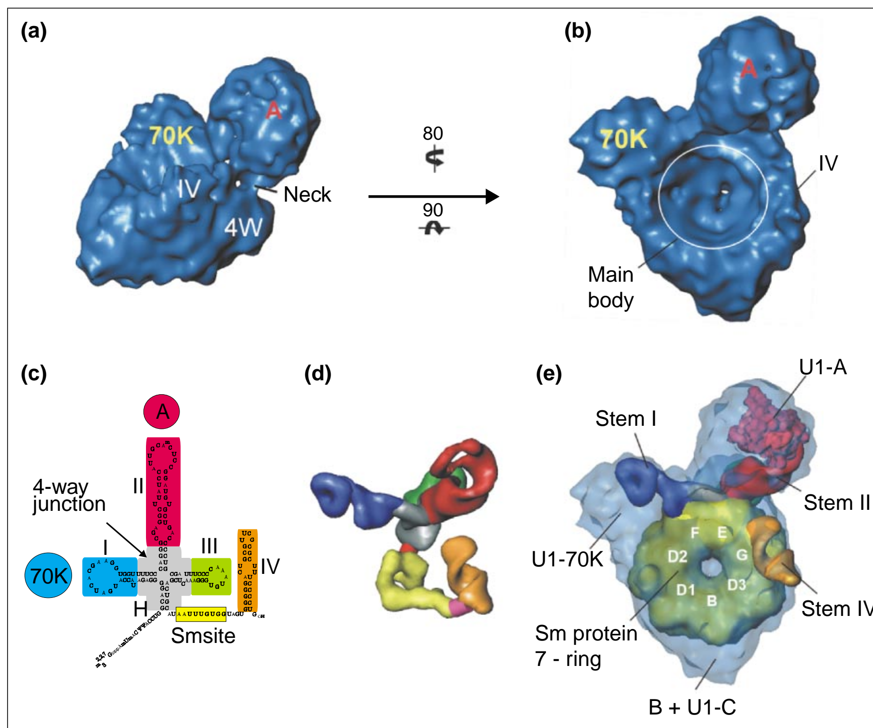
The La protein binds to newly synthesized U6 transcripts at their 3′ end and stabilizes them, thereby contributing to U6 snRNP assembly [18]. The U6 snRNA is also internally modified primarily by pseudouridylation and 2′-O-methylation, a process catalyzed at least in part, if not exclusively, by snoRNPs [37,38]. Factors directing U6 snRNA 2′-O-methylation and pseudouridylation have been localized to the nucleolus in mammalian cells, and recent studies indicate that U6 cycles through the nucleolus to acquire these modifications [38,39]. The U6 snRNA lacks an Sm site and formation of the U6 snRNP involves the association of seven so-called Sm-like proteins (LSm2, LSm3, LSm4, LSm5, LSm6, LSm7 and LSm8) with the U tract at the 3′ end of the U6 snRNA [40,41]. LSm2–8 share the conserved structural motifs characteristic of all Sm protein family members [42,43], but unlike the canonical Sm proteins, all seven LSm proteins can be isolated as a heteromeric complex in the absence of U6 snRNA [40]. However, whether they bind *in vivo* as a pre-formed ring or stepwise in a manner analogous to the Sm proteins is presently not clear. In yeast, the LSm2–LSm8 proteins stabilize the U6 snRNA and are required for assembly of a functional U6 snRNP [18,42,43].

The mechanism whereby the LSm proteins are transported to the nucleus also remains open to question. As U6 does not appear to transit through the cytoplasm, the LSm proteins, in stark contrast to the Sm proteins, are most probably imported as such, rather than as part of an snRNP complex. Interestingly, a subset of the U6-associated LSm proteins, LSm2 to LSm7, together with LSm1, form a complex involved in mRNA degradation that is found predominantly in the cytoplasm [44••,45••]. Thus, substitution of LSm1 for LSm8 has a significant effect not only on the function, but also on the cellular localization of these complexes. In subsequent steps, the U6 snRNP interacts with the U4 snRNP, primarily via U4 and U6 snRNA base pairing, to form the U4/U6 snRNP; LSm proteins facilitate this interaction *in vitro* [40].

The UsnRNP Sm core structure

As the UsnRNPs are spliceosomal subunits, information regarding their structure also provides insight into the probable structure of the spliceosome and its assembly intermediates. Much progress has been made in recent years in uncovering the molecular structure of the UsnRNP Sm core domain, which is formed upon interaction of the Sm proteins with a UsnRNA's Sm site. The hallmark of the Sm protein family, which also includes the LSm proteins, is the presence of the Sm motif, which is composed of two conserved regions, Sm1 and Sm2, that are separated by a linker of variable length [46,47]. A structural model of the mammalian Sm core domain based on electron microscopy (EM) of native snRNP particles, Sm–Sm protein interaction assays and X-ray crystallography of two human Sm protein complexes (B–D3 and D1–D2) was recently proposed [48]. This model depicts the UsnRNP Sm core as a seven-membered ring structure containing one copy of each Sm protein, in the order

Figure 2



The 3D structure of the human U1 snRNP at 10 Å resolution. **(a)** Side view and **(b)** top view of the U1 snRNP obtained by cryoelectron microscopy. The major protuberances contain the 70K and A proteins, whereas the smaller two, which are connected by a neck, arise from the U1 snRNA's stem-loop IV and four-way junction (4W), as indicated. The main body is marked by a circle in (b). **(c)** Secondary and **(d)** 3D structural models of the U1 snRNA. The four stem-loops, four-way junction and Sm site are color coded. **(e)** Model of the 3D arrangement of the RNA and proteins in the human U1 snRNP. A surface representation of the Sm protein ring was fitted into the 3D structure. The precise positions of stem-loops I and II of the U1 snRNA are not known. Reprinted with permission from *Nature* (**409**:539-542) copyright 2001 Macmillan Magazines Ltd.

G–E–F–D2–D1–B–D3, with the Sm site RNA–Sm protein contacts occurring on the inner surface of the proposed ring ([48]; Figure 1). LSm proteins also form a related structure; a complex of LSm2–8 lacking RNA appears highly similar to the snRNP Sm core in the electron microscope [40]. The structure of the Sm core appears to be evolutionarily conserved between humans, *S. cerevisiae* and trypanosomal UsnRNPs [49].

A number of studies in the past year have provided experimental evidence for the proposed model of the Sm core. *In vitro* studies performed with an apparent Sm protein progenitor from the archaeon *Archeoglobus fulgidus* indicated that the propensity to form multimeric complexes and bind to short U-rich stretches of RNA is an ancient property of Sm-motif-containing proteins [50•]. EM of the RNP complex formed upon interaction of this Sm-related protein with oligo-uridine revealed a ring-like structure similar to the mammalian UsnRNP Sm core. Significantly, multivariate statistical analysis of these EM images revealed that the observed ring exhibits a seven-fold symmetry, demonstrating for the first time that an Sm protein ring indeed consists of seven subunits.

In a separate study, several Sm proteins were shown to contact specific nucleotides of the Sm site within the mammalian Sm core [51•]. In addition, Sm protein amino acids contacting the Sm site were identified and, significantly, these amino acids were found to be present in a region predicted to protrude into the hole of the Sm protein ring, namely loop L3 of the Sm1 motif. These data provided the first evidence that multiple Sm protein–RNA contacts occur on the inner surface of the UsnRNP's Sm protein ring. Experimental support for the proposed UsnRNP Sm core model was also provided by cryo-EM studies of the human U1 snRNP, which revealed the 3D structure of the Sm core at 10 Å resolution (see below).

Elucidation of the three-dimensional structure of the human U1 snRNP

Owing to its relatively simple composition, the human U1 snRNP has been the best characterized of the UsnRNPs at the biochemical and structural level. Together with the Sm proteins, it contains three specific proteins, denoted 70K, A and C, that associate with the 164 nucleotide long U1 snRNA (reviewed in [2]). The U1–70K and U1–A proteins, which contain classical RNA recognition motifs (RRMs), bind directly to stem-loops I and II of the U1 snRNA, respectively, whereas the U1–C protein associates via protein–protein interactions with U1–70K and/or one or more of the Sm proteins (reviewed in [2]). Information about the tertiary structure of this UsnRNP was previously limited to the atomic structure of a fragment of U1–A bound to the U1 snRNA stem-loop II that was determined by X-ray crystallography [52].

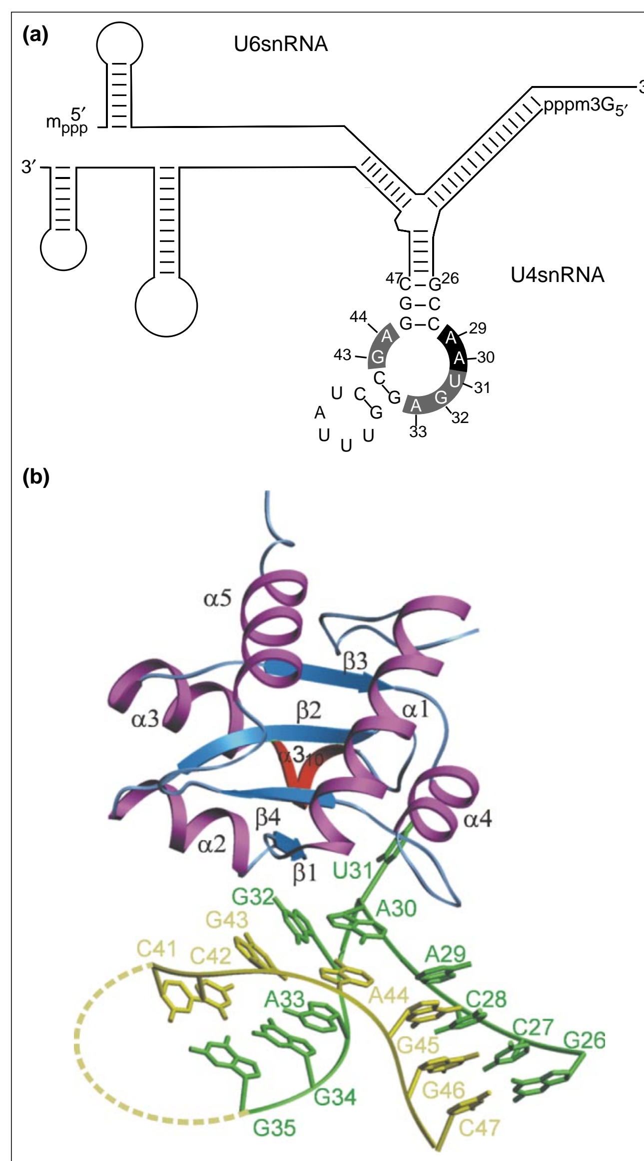
By performing single-particle electron cryomicroscopy, Stark *et al.* [53••] have determined the 3D structure of the human U1 snRNP at 10 Å resolution; this is the first reported 3D structure of a UsnRNP. Two different views of the U1 structure obtained after particle averaging, angular reconstitution and 3D reconstruction are shown in

Figures 2a and b. A model of the 3D molecular architecture of the U1 snRNP, based on the constraints provided by the 3D reconstruction and all available structural and biochemical data, is shown in Figure 2e. The major structural features of the 3D map are: a ring-shaped main body with an outer diameter of 70–80 Å, within which a heptameric Sm protein ring fits well and through which the Sm site RNA appears to thread, thus supporting the proposed Sm core model; two large and one small density element protruding from the top of the ring, which can be attributed to the U1–70K and U1–A proteins, and stem–loop IV of the U1 snRNA, respectively; and lastly, a fourth protuberance extending radially from the main body that can be assigned as the four-way helical junction of the U1 snRNA (see Figures 2c and d for 2D and space filling models of the U1 snRNA). U1–C and the carboxy-terminal domain of the SmB protein could be placed within the density located opposite the 70K and A protuberances, in front of the ring. These data provide exciting insights into the higher order structure of the U1 snRNP, cryo-EM thus represents a highly useful tool for future investigation of the 3D structure of the remaining UsnRNPs, as well as the spliceosome and its assembly intermediates.

Crystal structure of a U4 snRNA–protein complex

In the past year much has been learned about the morphology of the human U4/U6 snRNP. The U4 and U6 snRNAs are extensively base paired in the U4/U6 snRNP, and both Sm and LSm heptameric complexes are components of this particle. At least five additional proteins associate with the U4/U6 snRNP, including a 15.5 kDa protein, polypeptides of 20, 60 and 90 kDa that form a complex with one another [54,55] and a newly characterized 61 kDa protein that is homologous to the *S. cerevisiae* Prp31p protein (O Makarova, E Makarov, R Lührmann, unpublished data). The 20 kDa protein (now termed cyclophilin H) belongs to the cyclophilin family of peptidyl-prolyl isomerases; recently, its atomic structure was determined by X-ray crystallography and shown to resemble that of cyclophilin A, with notable differences seen only in their surface properties [56]. A 13S U4/U6 snRNP containing the aforementioned proteins has now been isolated by dissociation of human 25S U4/U6.U5 tri-snRNPs (O Makarova, E Makarov, R Lührmann, unpublished data). Electron micrographs revealed that this U4/U6 particle exhibits a bicycle-like structure, with two round domains (attributed to the U4–Sm and U6–LSm complexes) separated by what are probably parts of the two snRNAs and U4/U6-specific proteins.

Figure 3



Structure of the human 15.5 kDa–U4 5′ stem-loop complex. (a) Secondary structure model of the U4 and U6 snRNAs. The nucleotide sequence of the fragment of the 5′ stem-loop used for crystallization studies is indicated. Nucleotides previously shown to be essential for binding of the 15.5 kDa protein that are 100% conserved are shaded grey, whereas those shaded black are preferentially purines. (b) 3D ribbon diagram of the 15.5 kDa protein (alpha helices and beta sheets are shown in purple and blue, respectively) bound to a fragment of the U4 snRNA 5′ stem-loop. Green and yellow are used to distinguish the two RNA strands. Reprinted with permission from [58••].

The 15.5 kDa U4/U6 protein binds to the 5′ stem-loop of the U4 snRNA, and it has been suggested that it serves as a platform for the subsequent association of the remaining U4/U6 proteins [57]. Interestingly, the crystal structure of the 15.5 kDa–U4 5′ stem-loop complex has now been resolved at 2.9 Å resolution [58••]. The 15.5 kDa protein folds into a single globular domain that exhibits a so-called α–β–α sandwich structure, one of the most commonly observed protein folds (Figure 3). Only one side of the protein contacts the RNA, interacting predominantly with a purine-rich (5+2) internal loop formed within the U4 snRNA 5′ stem-loop, which adopts a stem–internal-loop–stem structure. The RNA folds into an unusual structure, with the double helical stems converging at an angle of approximately 30° and one helix containing two tandem sheared GA base pairs. U31 of the internal loop

protrudes from the rest of the RNA and fits into a pocket of the 15.5 kDa protein. As the 15.5 kDa protein belongs to a larger family of RNP-associated proteins that possess a novel RNA-binding motif, the determination of the atomic structure of the 15.5 kDa–U4 5′ stem-loop complex also provides insight into how this family generally recognizes RNA ([57] and references therein).

Surprisingly, the 15.5 kDa protein is also a component of box C/D snoRNPs, binding directly to the snoRNA's box C/D motif; this motif can theoretically adopt a structure similar to that of the U4 snRNA's 5′ stem-loop [59••]. Interestingly, the U4/U6-61 kDa protein (Prp31p) shares homology with both Nop56p and Nop58p, two proteins, that, like the 15.5 kDa protein, are common to all box C/D snoRNPs. This suggests that structural similarities between these snoRNPs and the U4 snRNP may extend beyond the 15.5 kDa protein–RNA complex ([60]; O Makarova, E Makarov, R Lührmann, unpublished data). In addition to structural similarities, the U4 snRNA and box C/D snoRNAs are thought to function in an analogous manner, acting as chaperones via their extensive base pairing with the U6 snRNA and pre-rRNA, respectively. These results thus raise the interesting possibility that box C/D snoRNPs and the U4 snRNP evolved from a common RNP ancestor.

Structural similarities between the major and minor snRNPs

Our knowledge of the structure of the 18S U11/U12 and U4atac/U6atac minor spliceosomal snRNPs is in its infancy. In addition to eight U11/U12-specific proteins and two sets of canonical Sm proteins, the human 18S U11/U12 snRNP also contains four subunits of splicing factor (SF) 3b (SF3b155, SF3b145, SF3b130 and SF3b49) that were initially identified as U2 snRNP proteins [61]. Recent studies have revealed that a 14 kDa protein (p14), previously shown in mammals to interact with the pre-mRNA's branch site adenosine, is also a component of both 17S U2 snRNPs and 18S U11/U12 snRNPs (CL Will, R Lührmann, C Query, unpublished results). Thus, consistent with their analogous functions in splicing, some structural characteristics might be conserved between the U2 and U12 snRNPs.

The U4/U6 and U4atac/U6atac snRNPs also appear to be structurally similar. Not only do their snRNAs form a similar Y-shaped base pairing interaction ([62]; Figure 3), but preliminary immunoprecipitation studies indicate that, at least several, if not all of the proteins found in the U4/U6 snRNP, are also present in the U4atac/U6atac snRNP (C Schneider, C Will, R Lührmann, unpublished data). Correspondingly, the U4/U6 15.5 kDa protein binds to the 5′ stem-loop of the U4atac snRNA *in vitro*, suggesting that a complex similar to that formed in U4/U6 is also found in the U4atac/U6atac snRNP [57]. The apparent conservation of a large number of proteins between the major and minor UsnRNPs suggests that one class of these snRNPs may have evolved in the presence of the other.

Novel insights into the spliceosome assembly pathway

Two of the main functions of the spliceosomal UsnRNPs are to recognize the 5′ and 3′ intron/exon boundaries and to assemble onto these sites the macromolecular RNP enzyme that catalyzes the splicing reaction. In mammals, four distinct spliceosomal complexes, which form in the temporal order E, A, B and C, have been detected (reviewed in [63]). Assembly of the major spliceosome is initiated by the ATP-independent recognition of the 5′ splice site by the U1 snRNP, which leads to the formation of the spliceosomal complex E. This interaction is mediated by base pairing of the U1 snRNA with the 5′ splice site, as well as by protein–protein and protein–pre-mRNA interactions involving the U1–70K (in conjunction with SR proteins) and U1–C proteins, respectively [64].

The *S. cerevisiae* U1 snRNP contains nine specific proteins (as opposed to three in mammals), some of which (Prp40p, Nam8p and Luc7p) facilitate the U1/5′ splice site interaction [65–68]. Nam8p interacts with nonconserved pre-mRNA nucleotides downstream of the 5′ splice site and is the only stably associated U1 protein known to directly influence 5′ splice site choice [68]. Recent studies have demonstrated that it is also required for Mer1p-activated splicing, one of the few examples of regulated splicing in budding yeast, which occurs during meiosis [69•]. Thus, Nam8p forms part of the constitutive splicing machinery and also functions in discriminating between 5′ splice sites during regulated splicing.

Interestingly, a human homologue of the Nam8p protein, TIA-1, which is known to promote apoptosis, regulates the alternative splicing of the human Fas receptor, a protein involved in programmed cell death, and the *Drosophila* msl-2, a protein involved in X chromosome dosage compensation [70••]. TIA-1 binds to U-rich sequences of these pre-mRNAs and, like Nam8p, promotes the association of the U1 snRNP with 5′ splice sites located upstream of its binding site. Thus, despite differences in their U1 association behaviour (i.e. Nam8p is stably associated whereas TIA-1 is not), these proteins appear to be orthologues. The fact that several proteins are more stably associated with the U1 snRNP in yeast versus vertebrates may be related to the fact that splicing is highly regulated in the latter but constitutive in essentially all yeast introns. That is, in vertebrates the association of certain proteins with U1 appears to be more flexible, enhancing the potential for regulating the affinity of the U1 snRNP for different 5′ splice sites.

Recent studies have challenged several aspects of the currently accepted model of spliceosome assembly. The U2 snRNP has long been thought to first associate in an ATP-dependent step subsequent to E complex formation, recognizing the pre-mRNA's branch site (reviewed in [63]). The U2–branch site interaction is facilitated by a combination of base pairing and protein–pre-mRNA interactions, involving SF3b and SF3a subunits, as well as the

non-snRNP splicing factors U2AF and SF1/BBP, and this interaction results in the formation of splicing complex A (i.e. the pre-spliceosome) [64]. The recent identification of p14, a branch site adenosine interacting protein, as a U2 snRNP component suggests that this protein may also facilitate the U2–branch site interaction (CL Will, R Lührmann, C Query, unpublished data). Interestingly, in contrast to the longstanding model, 17S U2 snRNPs have now been shown to be functionally associated with the pre-mRNA at the time of E complex formation [71••]. The intermolecular interactions mediating the association of the 17S U2 snRNP with this complex are presently not clear. However, its integration into the E complex does not require the branch site, and thus the stable interaction of U2 with this region of the pre-mRNA appears to occur in a subsequent step when splicing complex A is formed. These results, together with previous studies indicating that at least a small percentage of U1 and U2 snRNPs are associated with one another in splicing extracts [72,73], also suggest that the initial steps of major and minor spliceosome assembly may be more similar than originally thought. In the minor spliceosome, the U11 and U12 snRNPs, which form an 18S complex before association with the pre-mRNA, bind simultaneously to the 5′ splice site and branch site, respectively [74].

The U4/U6•U5 tri-snRNP also functionally associates with the pre-mRNA at a much earlier stage of spliceosome assembly than previously thought. According to the consensus model of spliceosome assembly, the U4/U6•U5 tri-snRNP first interacts with the pre-mRNA after the pre-spliceosome (complex A) has been formed, generating the spliceosome (complex B). At the time of complex B formation, many UsnRNP structural rearrangements occur, leading to the formation of a complex RNA–RNA network within the spliceosome (reviewed in [75]). For example, the U4/U6 base pairing interaction is disrupted and the U6 snRNA base pairs with the U2 snRNA and also the 5′ splice site. Furthermore, U5 snRNA loop I base pairs with exon sequences at the 5′ splice site, and later with exon sequences at the 3′ splice site. The 5′ and 3′ splice sites are also contacted by the U5-220 kDa protein (Prp8p in yeast), which is thought to help tether the U5 snRNP to the spliceosome and also to contribute to 5′ and 3′ splice site recognition (reviewed in [76]). Recent studies have indicated that the 5′ splice site is also recognized by the U4/U6•U5 tri-snRNP, together with U1, at the earliest stages of spliceosome assembly (i.e. prior to A complex assembly) both in nematode and HeLa cell extracts [77••]. Surprisingly, the tri-snRNP contacts the 5′ splice site in an ATP-dependent manner, even in the absence of a stable U2 snRNP–branch site interaction. This early interaction appears to represent a functionally important, initial recruitment of the tri-snRNP to the pre-mRNA, with subsequent stabilization of tri-snRNP association occurring at the time of complex B formation, presumably via the formation of the aforementioned RNA–RNA interactions.

An additional function for UsnRNP DExD/H-box proteins?

Major structural rearrangements lead to the formation of spliceosomal complex C, in which the pre-mRNA has undergone the first catalytic step of splicing. After the second step, the spliceosome disassembles and the mRNA and excised intron are released. Thus, the spliceosome is highly dynamic, with multiple RNA–RNA and RNA–protein rearrangements occurring during spliceosome assembly, the catalytic steps of splicing and subsequent spliceosome disassembly. Many of these conformational changes are thought to be mediated by UsnRNP proteins. These include known or suspected RNA unwindases (the U5-200 and U5-100 kDa DExD/H-box proteins) and a putative GTPase (the ribosomal elongation factor 2 (EF2)-like U5-116 kDa protein) (reviewed in [75]). Although the *in vivo* targets of most of these proteins have not been conclusively identified, the U5-200 kDa protein (Brr2/Snu246/Slt22/Rss1p in yeast) has been shown to unwind U4/U6 duplexes *in vitro* [78,79], and the yeast orthologue of U5-100 kDa (Prp28p) has been implicated in disrupting the U1 snRNA-5′ splice site interaction [80].

In addition to potentially catalyzing RNA strand displacement, the U5-200 and U5-100 kDa proteins, as well as other spliceosomal DExD/H-box proteins, have also been proposed to disrupt RNA–protein interactions [75]. The recent demonstration that a member of the DExH-box protein family, the Vaccinia virus protein NPH-II, can actively displace the U1–A protein from its extremely stable interaction with RNA *in vitro*, adds credence to the idea that the U5-200 and U5-100 kDa proteins may directly catalyze RNA–protein rearrangements within the spliceosome [81••]. Indeed, *in vivo* studies in yeast now suggest that Prp28p may function by actively disrupting an RNA–protein interaction that helps to stabilize the U1 snRNA–5′ splice site base pairing interaction, namely the interaction between the U1-C protein and the 5′ splice site [82••].

A novel RNA–RNA interaction pathway in the minor U12-dependent spliceosome

The dynamic RNA–RNA network that guides the formation of the spliceosome appears to be highly similar in the major and minor spliceosomes (reviewed in [3]). Upon stable association of the U5 and U4atac/U6atac snRNPs with the U12-dependent pre-spliceosome, RNA rearrangements occur that mimic those observed in the major spliceosome.

The nature and dynamics of RNA–RNA interactions within the U12-dependent spliceosome have now been analyzed in detail, using crosslinking and an elegant oligonucleotide blocking/release approach [83••]. Interestingly, in addition to the well established pathway of RNA interactions that leads to catalytic core formation in the major spliceosome, an alternative interaction pathway was revealed. In the latter, a structural intermediate is formed in which the U6atac remains partially bound to

U4atac (via stem II) and at the same time forms the well characterized helix Ia (but not Ib) with the U12 snRNA, but does not yet base pair with the 5′ splice site. Previously, it was thought that the U6/5′ splice site (or U6atac/5′ splice site) interaction had to occur before the unwinding of the U4/U6 (or U4atac/U6atac) duplex and the formation of U2/U6 (or U12/U6atac) helices Ia and Ib. These studies thus indicate that, at least in the minor spliceosome, the complex intermolecular RNA structure that contributes to the spliceosome's active site can be generated by alternative pathways. Whether or not this novel structural intermediate also forms in the major spliceosome is presently not known.

Growing evidence that UsnRNAs catalyze splicing

A longstanding debate in the pre-mRNA splicing field is whether splicing is ultimately catalyzed by RNA, proteins or both. The elucidation of similarities between the functionally important intramolecular structures formed in self-splicing group II introns and intermolecular structures formed by the pre-mRNA and the U2, U5 and U6 snRNAs, as well as the fact that both types of introns are spliced by the same chemical pathway, has bolstered the idea that pre-mRNA splicing is catalyzed by RNA. Much evidence points to a central, if not exclusive, role for the U6 snRNA in splicing catalysis.

U6 is clearly a component of the spliceosome's catalytic core and is the most highly phylogenetically conserved UsnRNA, and minor modifications in its sequence or phosphodiester backbone completely block the catalytic steps of splicing (reviewed in [84]). Significantly, Yean *et al.*, [85••] now demonstrate that the U6 snRNA binds to a divalent metal ion that is required for the first step of splicing. The spliceosome is a metalloenzyme [86], and thus the ability of U6 to coordinate a metal ion required for catalysis strongly supports the idea that splicing is catalyzed by this UsnRNA. However, these results do not rule out the possibility that other snRNAs, such as U2, or even proteins also directly contribute to splicing catalysis. Spliceosomal proteins, such as the U5-220 kDa protein (Prp8p), also appear to be present at the catalytic center of the spliceosome [87,88]. An unambiguous answer to this question awaits the determination of the atomic structure of the spliceosome's active sites or the establishment of a protein-free *in vitro* splicing system.

Conclusions

UsnRNPs play a central role in pre-mRNA splicing and, accordingly, much effort has been placed on understanding their biogenesis, structure and function. Although progress has been made in the past year, many aspects of UsnRNP biogenesis remain poorly understood. These include the precise order of assembly and modification events, the molecular details of many maturation steps and, in some instances, their cellular location. Nonetheless, a molecular picture of this process is now beginning to emerge. Our understanding of UsnRNP structure is also far from complete. However, the growing use of X-ray crystallography and now cryo-EM, coupled with classical biochemical techniques, has generated much-needed information about UsnRNP 3D structure and promises to continue to do so in the future. UsnRNP structural information should ultimately prove invaluable for gaining insight into the structure and, as a consequence, the function of the spliceosome itself.

Data gathered in the past year have revealed new roles for the U2 and U4/U6•U5 tri-snRNPs, namely that they function at the very early stages of splicing complex formation. Refinements in splicing complex isolation techniques may in the future lead to additional changes in the currently accepted model of spliceosome assembly. Finally, evidence in favor of UsnRNA-based catalysis of splicing continues to mount. Nonetheless, the unequivocal identification of those spliceosomal components directly responsible for catalysis remains a major challenge in the pre-mRNA splicing field.

Acknowledgements

We thank Elisa Izaurralde, Iain Mattaj, Tom Tuschl and members of our laboratory (Tilmann Achsel, Patrizia Fabrizio, Claudia Schneider and Evgeny Makarov) for helpful comments on the manuscript. We are grateful to Holger Stark and Ralf Ficner for kindly providing Figures 2 and 3, respectively. We apologize to those of our colleagues whose work, due to space limitations, was not cited. Work from our laboratory was supported by grants from the Gottfried Wilhelm Leibniz Program and from SFB397 and 523 to RL.

References and recommended reading

Papers of particular interest, published within the annual period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. Moore MJ, Query CC, Sharp PA: **Splicing of precursors to mRNA by the spliceosome.** In *RNA World*. Edited by Gesteland RF, Atkins JF. New York: Cold Spring Harbor Laboratory Press; 1993:303-357.

2. Will CL, Lührmann R: **snRNP structure and function.** In *Eukaryotic mRNA Processing*. Edited by Krainer AR. Oxford: IRL Press; 1997:130-173.

3. Burge CB, Tuschl T, Sharp PA: **Splicing of precursors to mRNAs by the spliceosomes.** In *The RNA World*, edn 2. Edited by Gesteland RF, Cech TR, Atkins JF. New York: Cold Spring Harbor Laboratory Press; 1999:525-560.

4. Hamm J, Mattaj IW: **Monomethylated cap structures facilitate RNA export from the nucleus.** *Cell* 1990, **63**:109-118.

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

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

7. Izaurralde E, Kutay U, von Kobbe C, Mattaj IW, Görlich D: **The asymmetric distribution of the constituents of the Ran system is essential for transport into and out of the nucleus.** *EMBO J* 1997, **16**:6535-6547.

8. •• Ohno M, Segref A, Bachi A, Wilm M, Mattaj IW: **PHAX, a mediator of U snRNA nuclear export whose activity is regulated by phosphorylation.** *Cell* 2000, **101**:187-198.
A specific mediator of vertebrate UsnRNP export, designated PHAX, was purified and characterized via *in vitro* reconstitution and oocyte injection experiments. A phosphorylated form of PHAX that is found in the nucleus, together with other previously characterized factors, mediates UsnRNA

export and is subsequently dephosphorylated in the cytoplasm. These studies point to the PHAX phosphorylation/dephosphorylation cycle as a means to ensure the directionality of UsnRNA export and greatly enhance our understanding of the molecular mechanism of UsnRNA export.

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

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

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

12. Pellizzoni L, Charroux B, Dreyfuss G: **SMN mutants of spinal muscular atrophy patients are defective in binding to snRNP proteins.** *Proc Natl Acad Sci USA* 1999, **96**:11167-11172.

13. Charroux B, Pellizzoni L, Perkinson RA, Yong J, Shevchenko A, Mann M, Dreyfuss G: **Gemin4. A novel component of the SMN complex that is found in both gems and nucleoli.** *J Cell Biol* 2000, **148**:1177-1186.

14. Friesen WJ, Dreyfuss G: **Specific sequences of the Sm and Sm-like (Lsm) proteins mediate their interaction with the spinal muscular atrophy disease gene product (SMN).** *J Biol Chem* 2000, **275**:26370-26375.

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

16. Brahms H, Raymackers J, Union A, de Keyser F, Meheus L, Lührmann R: **The C-terminal RG dipeptide repeats of the spliceosomal Sm proteins D1 and D3 contain symmetrical dimethylarginines, which form a major B-cell epitope for anti-Sm autoantibodies.** *J Biol Chem* 2000, **275**:17122-17129.

17. Noble SM, Guthrie C: **Identification of novel genes required for yeast pre-mRNA splicing by means of cold-sensitive mutations.** *Genetics* 1996, **143**:67-80.

18. • Xue D, Rubinson DA, Pannone BK, Yoo CJ, Wolin SL: **U snRNP assembly in yeast involves the La protein.** *EMBO J* 2000, **19**:1650-1660.

A combination of genetic and biochemical studies in budding yeast reveal that, in addition to binding and stabilizing nascent polymerase III transcripts, the La protein binds to U1, U2, U4 and U5 pre-snRNAs and apparently, at least in the case of the U4 snRNA, facilitates Sm protein binding and thus formation of the UsnRNP Sm core particle.

19. Pannone BK, Xue D, Wolin SL: **A role for the yeast La protein in U6 snRNP assembly: evidence that the La protein is a molecular chaperone for RNA polymerase III transcripts.** *EMBO J* 1998, **17**:7442-7453.

20. Mattaj IW: **Cap trimethylation of U snRNA is cytoplasmic and dependent on U snRNP protein binding.** *Cell* 1986, **46**:905-911.

21. Seipelt RL, Zheng B, Asuru A, Rymond BC: **U1 snRNA is cleaved by RNase III and processed through an Sm site-dependent pathway.** *Nucleic Acids Res* 1999, **27**:587-595.

22. Neuman de Vegvar HE, Dahlberg JE: **Nucleocytoplasmic transport and processing of small nuclear RNA precursors.** *Mol Cell Biol* 1990, **10**:3365-3375.

23. Chanfreau G, Elela SA, Ares M Jr, Guthrie C: **Alternative 3′-end processing of U5 snRNA by RNase III.** *Genes Dev* 1997, **11**:2741-2751.

24. Allmang C, Kufel J, Chanfreau G, Mitchell P, Petfalski E, Tollervey D: **Functions of the exosome in rRNA, snoRNA and snRNA synthesis.** *EMBO J* 1999, **18**:5399-5410.

25. van Hoof A, Lennertz P, Parker R: **Yeast exosome mutants accumulate 3′-extended polyadenylated forms of U4 small nuclear RNA and small nucleolar RNAs.** *Mol Cell Biol* 2000, **20**:441-452.

26. • van Hoof A, Lennertz P, Parker R: **Three conserved members of the RNase D family have unique and overlapping functions in the processing of 5S, 5.8S, U4, U5, RNase MRP and RNase P RNAs in yeast.** *EMBO J* 2000, **19**:1357-1365.

Evidence is provided that 3′end processing of the U4 and U5 snRNAs in yeast is not solely mediated by enzymes present in the exosome; it also requires the activity of members of the RNase D family of 3′ exonucleases.

27. Fischer U, Sumpter V, Sekine M, Satoh T, Lührmann R: **Nucleo-cytoplasmic transport of U snRNPs: definition of a nuclear location signal in the Sm core domain that binds a transport receptor independently of the m3G cap.** *EMBO J* 1993, **12**:573-583.

28. Palacios I, Hetzer M, Adam SA, Mattaj IW: **Nuclear import of U snRNPs requires importin beta.** *EMBO J* 1997, **16**:6783-6792.

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

30. Bordonné R: **Functional characterization of nuclear localization signals in yeast Sm proteins.** *Mol Cell Biol* 2000, **20**:7943-7954.

31. Massenet S, Mougin A, Branlant C: **Posttranscriptional modifications in the U small nuclear RNAs.** In *The Modification and Editing of RNA*. Edited by Grosjean H, Benne R. Washington DC: ASM Press; 1998: 201-227.

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

33. •• Jády BE, Kiss T: **A small nucleolar guide RNA functions both in 2′O-ribose methylation and pseudouridylation of the U5 spliceosomal RNA.** *EMBO J* 2001, **20**:541-551.

In vitro and *in vivo* pseudouridylation and 2′-O-methylation assays demonstrate for the first time that a snoRNA (U85) directs the modification of an RNA polymerase II-transcribed UsnRNA, namely U5. Furthermore, unlike all other snoRNPs characterized to date, the U85 snoRNP catalyzes both pseudouridylation and 2′-O-methylation of RNA.

34. Tollervey D, Kiss T: **Function and synthesis of small nucleolar RNAs.** *Curr Opin Cell Biol* 1997, **9**:337-342.

35. Massenet S, Motorin Y, Lafontaine DL, Hurt EC, Grosjean H, Branlant C: **Pseudouridine mapping in the *Saccharomyces cerevisiae* spliceosomal U small nuclear RNAs (snRNAs) reveals that pseudouridine synthase pus1p exhibits a dual substrate specificity for U2 snRNA and tRNA.** *Mol Cell Biol* 1999, **19**:2142-2154.

36. Hetzer M, Mattaj IW: **An ATP-dependent, Ran-independent mechanism for nuclear import of the U1A and U2B′′ spliceosome proteins.** *J Cell Biol* 2000, **148**:293-303.

37. Tycowski KT, You ZH, Graham PJ, Steitz JA: **Modification of U6 spliceosomal RNA is guided by other small RNAs.** *Mol Cell* 1998, **2**:629-638.

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

39. Lange TS, Gerbi SA: **Transient nucleolar localization of U6 small nuclear RNA in *Xenopus laevis* oocytes.** *Mol Biol Cell* 2000, **11**:2419-2428.

40. Achsel T, Brahms H, Kastner B, Bachi A, Wilm M, Lührmann R: **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* 1999, **18**:5789-5802.

41. Vidal VP, Verdone L, Mayes AE, Beggs JD: **Characterization of U6 snRNA-protein interactions.** *RNA* 1999, **5**:1470-1481.

42. Mayes AE, Verdone L, Legrain P, Beggs JD: **Characterization of Sm-like proteins in yeast and their association with U6 snRNA.** *EMBO J* 1999, **18**:4321-4331.

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

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

Affinity purification of proteins associated with Lsm3 in yeast revealed the presence of two distinct Lsm complexes consisting of Lsm2–Lsm8 or Lsm1–Lsm7. The latter complex also contained a known exoribonuclease and subsequent genetic studies indicated that it indeed participates in mRNA degradation.

45. •• Tharun S, He W, Mayes AE, Lennertz P, Beggs JD, Parker R: **Yeast Sm-like proteins function in mRNA decapping and decay.** *Nature* 2000, **404**:515-518.

A combination of genetic and biochemical studies provide evidence that, in addition to the Lsm2–Lsm8 complex that is associated with the U6 snRNA, a different complex consisting of Lsm1–Lsm7 exists and plays a role in cytoplasmic mRNA degradation.

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

47. Séraphin B: **Sm and Sm-like proteins belong to a large family: identification of proteins of the U6 as well as the U1, U2, U4 and U5 snRNPs.** *EMBO J* 1995, **14**:2089-2098.

48. Kambach C, Walke S, Young R, Avis JM, de la Fortelle E, Raker VA, Lührmann R, Li J, Nagai K: **Crystal structures of two Sm protein complexes and their implications for the assembly of the spliceosomal snRNPs.** *Cell* 1999, **96**:375-387.

49. Palfi Z, Lucke S, Lahm HW, Lane WS, Kruft V, Bragado-Nilsson E, Séraphin B, Bindereif A: **The spliceosomal snRNP core complex of *Trypanosoma brucei*: cloning and functional analysis reveals seven Sm protein constituents.** *Proc Natl Acad Sci USA* 2000, **97**:8967-8972.

50. Achsel T, Stark H, Lührmann R: **The Sm domain is an ancient RNA**
• **binding motif with oligo[U] specificity.** *Proc Natl Acad Sci USA* 2001, **98**:3685-3689.
In vitro binding studies with an apparent Sm protein progenitor from the archaeon A fulgidus, coupled with electron microscopy, demonstrate for the first time that, upon interacting with RNA, proteins containing an Sm motif indeed form a seven-membered ring (as previously proposed for proteins comprising the UsnRNP Sm core).

51. Urlaub H, Raker VA, Kostka S, Lührmann R: **Sm protein-Sm site RNA**
• **interactions within the inner ring of the spliceosomal snRNP core structure.** *EMBO J* 2001, **20**:1-10.
UV crosslinking experiments coupled with mass spectrometry, reveal that, within the UsnRNP Sm core domain, the UsnRNA's Sm site is contacted by several Sm proteins in a site-specific manner. These contacts occur on the inner surface of the ring formed by the Sm proteins, providing experimental evidence for the previously proposed structure of the mammalian Sm core domain.

52. Oubridge C, Ito N, Evans PR, Teo CH, Nagai K: **Crystal structure at 1.92 Å resolution of the RNA-binding domain of the U1A spliceosomal protein complexed with an RNA hairpin.** *Nature* 1994, **372**:432-438.

53. Stark H, Dube P, Lührmann R, Kastner B: **Arrangement of RNA and**
•• **proteins in the spliceosomal U1 small nuclear ribonucleoprotein particle.** *Nature* 2001, **409**:539-542.
Single particle electron cryomicroscopy reveals for the first time the 3D structure, at 10Å resolution, of a spliceosomal UsnRNP. On the basis of the electron cryomicroscopy images, as well as previous biochemical data, a model of the 3D architecture of the human U1 snRNP is presented.

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

55. Teigelkamp S, Achsel T, Mundt C, Gothel SF, Cronshagen U, Lane WS, Marahiel M, Lührmann R: **The 20kD protein of human [U4/U6.U5] tri-snRNPs is a novel cyclophilin that forms a complex with the U4/U6-specific 60 kD and 90 kD proteins.** *RNA* 1998, **4**:127-141.

56. Reidt U, Reuter K, Achsel T, Ingelfinger D, Lührmann R, Ficner R: **Crystal structure of the human U4/U6 small nuclear ribonucleoprotein particle-specific SnuCyp-20, a nuclear cyclophilin.** *J Biol Chem* 2000, **275**:7439-7442.

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

58. Vidovic I, Nottrott S, Hartmuth K, Lührmann R, Ficner R: **Crystal**
•• **structure of the spliceosomal 15.5 kDa protein bound to a U4 snRNA fragment.** *Mol Cell* 2000, **6**:1331-1342.
The 15.5 kDa U4/U6 snRNP and box C/D snoRNP protein belongs to a novel family of RNA-binding proteins. The atomic structure of the U4 snRNA/15.5 kDa protein complex at 2.9 Å resolution is presented, providing insight into the molecular mechanism whereby this family of proteins recognize RNA.

59. Watkins NJ, Ségault V, Charpentier B, Nottrott S, Fabrizio P, Bachi A,
•• Wilm M, Rosbash M, Branlant C, Lührmann R: **A common core RNP structure shared between the small nucleolar box C/D RNPs and the spliceosomal U4 snRNP.** *Cell* 2000, **103**:457-466.
A component of the U4/U6.U5 tri-snRNP, the 15.5 kDa protein, which binds directly to the U4 snRNA is shown to also be a common component of box C/D snoRNPs. These studies, coupled with the fact that the U4 snRNP and box C/D snoRNPs exhibit functional similarities, raise the intriguing possibility that these RNP particles may have evolved from a common ancestor.

60. Gautier T, Berges T, Tollervey D, Hurt E: **Nucleolar KKE/D repeat proteins Nop56p and Nop58p interact with Nop1p and are required for ribosome biogenesis.** *Mol Cell Biol* 1997, **17**:7088-7098.

61. Will CL, Schneider C, Reed R, Lührmann R: **Identification of both shared and distinct proteins in the major and minor spliceosomes.** *Science* 1999, **284**:2003-2005.

62. Tarn WY, Steitz JA: **Highly diverged U4 and U6 small nuclear RNAs required for splicing rare AT-AC introns.** *Science* 1996, **273**:1824-1832.

63. Reed R, Palandjian L: **Spliceosome assembly.** In *Eukaryotic mRNA processing*. Edited by Krainer AR. Oxford: IRL Press; 1997:103-129.

64. Will CL, Lührmann R: **Protein functions in pre-mRNA splicing.** *Curr Opin Cell Biol* 1997, **9**:320-328.

65. Gottschalk A, Tang J, Puig O, Salgado J, Neubauer G, Colot HV, Mann M, Séraphin B, Rosbash M, Lührmann R, Fabrizio P: **A comprehensive biochemical and genetic analysis of the yeast U1 snRNP reveals five novel proteins.** *RNA* 1998, **4**:374-393.

66. Abovich N, Rosbash M: **Cross-intron bridging interactions in the yeast commitment complex are conserved in mammals.** *Cell* 1997, **89**:403-412.

67. Fortes P, Bilbao-Cortes D, Fornerod M, Rigaut G, Raymond W, Séraphin B, Mattaj IW: **Luc7p, a novel yeast U1 snRNP protein with a role in 5′ splice site recognition.** *Genes Dev* 1999, **13**:2425-2438.

68. Puig O, Gottschalk A, Fabrizio P, Séraphin B: **Interaction of the U1 snRNP with nonconserved intronic sequences affects 5′ splice site selection.** *Genes Dev* 1999, **13**:569-580.

69. Spingola M, Ares M Jr: **A yeast intronic splicing enhancer and**
• **Nam8p are required for Mer1p-activated splicing.** *Mol Cell* 2000, **6**:329-338.
The constitutively expressed yeast U1 snRNP protein Nam8p, together with the cell-type-specific splicing factor Mer1p, is shown to regulate the meiosis-specific splicing of a subset of yeast pre-mRNAs.

70. Forch P, Puig O, Kedersha N, Martinez C, Granneman S, Séraphin B,
•• Anderson P, Valcarcel J: **The apoptosis-promoting factor TIA-1 is a regulator of alternative pre-mRNA splicing.** *Mol Cell* 2000, **6**:1089-1098.
In vitro splicing studies demonstrate that the mammalian apoptosis factor TIA-1 is not only structurally homologous to the yeast U1 snRNP Nam8p protein but also functions in an analogous manner. Specifically, TIA-1 is shown to be involved in 5′ splice site selection, promoting the association of the U1 snRNP with 5′ splice sites that are followed by U-rich sequences.

71. Das R, Zhou Z, Reed R: **Functional association of U2 snRNP with**
•• **the ATP-independent spliceosomal complex E.** *Mol Cell* 2000, **5**:779-787.
Affinity purification of the mammalian spliceosomal E complex reveals that in addition to the U1 snRNP, the functional form of the U2 snRNP is also present in this early complex. These studies suggest that, contrary to the previously accepted model of spliceosome assembly, the U2 snRNP functionally associates with the pre-mRNA at an early, ATP-independent step together with the U1 snRNP.

72. Mattaj IW, Habets WJ, van Venrooij WJ: **Monospecific antibodies reveal details of U2 snRNP structure and interaction between U1 and U2 snRNPs.** *EMBO J* 1986, **5**:997-1002.

73. Bringmann P, Lührmann R: **Antibodies specific for N6-methyladenosine react with intact snRNPs U2 and U4/U6.** *FEBS Lett* 1987, **213**:309-315.

74. Frilander MJ, Steitz JA: **Initial recognition of U12-dependent introns requires both U11/5′ splice-site and U12/branchpoint interactions.** *Genes Dev* 1999, **13**:851-863.

75. Staley JP, Guthrie C: **Mechanical devices of the spliceosome: motors, clocks, springs, and things.** *Cell* 1998, **92**:315-326.

76. Umen JG, Guthrie C: **The second catalytic step of pre-mRNA splicing.** *RNA* 1995, **1**:869-885.

77. Maroney PA, Romfo CM, Nilsen TW: **Functional recognition of 5′**
•• **splice site by U4/U6.U5 tri-snRNP defines a novel ATP-dependent step in early spliceosome assembly.** *Mol Cell* 2000, **6**:317-328.
Employing a novel *in vitro* cis/trans-splicing competition assay evidence is provided that factors in addition to the U1 snRNP contribute to early 5′ splice site recognition. UV crosslinking and physical protection studies further demonstrate that the U4/U6•U5 tri-snRNP contacts the 5′ splice site in an

ATP-dependent manner, even in the absence of a stable U2 snRNP/branch site interaction. These results indicate that the 5′ splice site is also recognized by the U4/U6•U5 tri-snRNP, together with U1, already at the earliest stages of spliceosome assembly (i.e. prior to A complex assembly).

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

79. Raghunathan PL, Guthrie C: **RNA unwinding in U4/U6 snRNPs requires ATP hydrolysis and the DEIH-box splicing factor Brr2.** *Curr Biol* 1998, **8**:847-855.

80. Staley JP, Guthrie C: **An RNA switch at the 5′ splice site requires ATP and the DEAD box protein Prp28p.** *Mol Cell* 1999, **3**:55-64.

81. Jankowsky E, Gross CH, Shuman S, Pyle AM: **Active disruption of •• an RNA-protein interaction by a DExH/D RNA helicase.** *Science* 2001, **291**:121-125.
The Vaccinia virus DExH/D NPH-II protein is shown to displace the U1A protein from RNA in an active ATP-dependent manner *in vitro*. These studies provide the first evidence that DExH/D proteins, which were originally thought to act as RNA helicases, can also (or alternatively) catalyze the disruption of protein–RNA interactions.

82. Chen JY-F, Stands L, Staley JP, Jackups RR, Latus LJ, Chang T-H: •• **Specific alterations of U1-C protein or U1 small nuclear RNA can eliminate the requirement of Prp28p, an essential DEAD box splicing factor.** *Mol Cell* 2001, **7**:227-232.
Genetic studies in yeast demonstrate that the lethal phenotype of a Prp28p deletion mutant can be suppressed by a mutation in the U1-C protein that reduces its affinity for the pre-mRNA. These studies suggest that one potential function of the DEAD box protein Prp28p is to actively disrupt the interaction of the U1-C protein with the pre-mRNA. These studies thus suggest that the *in vivo* function of some DExH/D box proteins may be to displace proteins from RNA.

83. Frilander MJ, Steitz JA: **Dynamic exchanges of RNA interactions •• leading to catalytic core formation in the U12-dependent spliceosome.** *Mol Cell* 2001, **7**:217-226.
Photochemical crosslinking and an oligonucleotide displacement method are used to analyze the nature and dynamics of RNA–RNA interactions in the minor, U12-dependent spliceosome. These studies reveal an alternative RNA–RNA interaction pathway that leads to catalytic core formation in the minor spliceosome.

84. Nilsen TW: **RNA-RNA interactions in nuclear pre-mRNA splicing.** In *RNA Structure and Function*. Edited by Simons RW, Grunberg-Manago M. New York: Cold Spring Harbor Press; 1998:279-307.

85. Yean SL, Wuenschell G, Termini J, Lin RJ: **Metal-ion coordination by •• U6 small nuclear RNA contributes to catalysis in the spliceosome.** *Nature* 2000, **408**:881-884.
Incorporation of a single phosphorothioate at position U80 of the yeast U6 snRNA allows spliceosome formation but blocks the first step of pre-mRNA splicing *in vitro* when magnesium is the sole source of divalent metal ions; however, in the presence of manganese the first step of splicing is restored. These studies indicate that the U6 snRNA binds a divalent metal ion that is required for splicing catalysis, thereby providing strong evidence that pre-mRNA splicing is catalyzed by a UsnRNA.

86. Sontheimer EJ, Sun S, Piccirilli JA: **Metal ion catalysis during splicing of premessenger RNA.** *Nature* 1997, **388**:801-805.

87. Siatecka M, Reyes JL, Konarska MM: **Functional interactions of Prp8 with both splice sites at the spliceosomal catalytic center.** *Genes Dev* 1999, **13**:1983-1993.

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