HHS Public Access

Author manuscript

Nat Struct Mol Biol. Author manuscript; available in PMC 2014 December 01.

Published in final edited form as:

Nat Struct Mol Biol. 2014 June ; 21(6): 544–551. doi:10.1038/nsmb.2832.

Core structure of the U6 snRNP at 1.7 Å resolution

Eric J. Montemayor[1,2], **Elizabeth C. Curran**[1,3], **Hong Hong Liao**[2], **Kristie L. Andrews**[1,3], **Christine N. Treba**[1], **Samuel E. Butcher**[2], and **David A. Brow**[1]

[1]Department of Biomolecular Chemistry, University of Wisconsin-Madison, Madison, WI, USA

[2]Department of Biochemistry, University of Wisconsin-Madison, Madison, WI, USA

Abstract

The spliceosome is a dynamic assembly of five small nuclear ribonucleoproteins (snRNPs) that removes introns from eukaryotic pre-mRNA. U6 is the most conserved of the spliceosomal snRNAs and participates directly in catalysis. Here, we report the crystal structure of the *Saccharomyces cerevisiae* U6 snRNP core, containing most of U6 snRNA and all four RRM domains of the Prp24 protein. It reveals a unique interlocked RNP architecture that sequesters the 5′ splice site-binding bases of U6 snRNA. RRMs 1, 2 and 4 of Prp24 form an electropositive groove that binds double-stranded RNA and may nucleate annealing of U4 and U6 snRNAs. Substitutions in Prp24 that suppress a mutation in U6 localize to direct RNA-protein contacts. Our results provide the most complete view to date of a multi-RRM protein bound to RNA, and reveal striking co-evolution of protein and RNA structure.

U6 snRNA endows substrate specificity and catalytic function to the spliceosome, and is thought to derive from domain 5 of group II self-splicing introns[1–4]. The U6 snRNP in the budding yeast *S. cerevisiae* contains the 112-nucleotide U6 snRNA, 51 kDa Prp24 protein, and 94 kDa Lsm2–8 heteroheptamer[5–9]. Incorporation of U6 into the spliceosome requires unwinding of the internal stem loop (ISL) within the U6 snRNP and pairing to U4 snRNA, forming a U4/U6 di-snRNP (Fig. 1a). Prp24 acts as a chaperone for annealing of the U4 and U6 snRNPs[10–15], and is displaced from U6 after U4/U6 pairing is complete[6,8,16]. The Lsm ring, which binds the uracil-rich 3′ end of U6, also promotes U4/U6 annealing but remains bound to U6 in the U4/U6 di-snRNP[7,17–19]. During spliceosome activation, U6 is transferred from U4 to U2 snRNA and the U6 ISL reforms (Fig. 1a), creating a structure that binds two catalytic metal ions required for the splicing reaction[3]. After intron excision, U6

Correspondence should be addressed to D.A.B. (dabrow@wisc.edu) or S.E.B. (sebutcher@wisc.edu).

[3]Present addresses: Department of Pharmacology, University of Washington, Seattle, WA, USA (E.C.C.) and Illinois College of Optometry, Chicago, IL, USA (K.L.A.).

Author Contributions: E.J.M and H.L. prepared crystallization samples; E.J.M. determined the structure; E.C.C., K.L.A., C.N.T. and D.A.B. performed the yeast genetic analysis; S.E.B and D.A.B supervised the work and wrote the manuscript with assistance from E.J.M.

Competing Financial Interests: The authors declare no competing financial interests.

Accession Code: Atomic coordinates and structure factors have been deposited with the Protein Data Bank under accession code 4N0T.

snRNA dissociates from U2 and reforms the U6 snRNP, which can enter another splicing cycle by re-annealing with U4 snRNA.

Yeast Prp24 contains four RNA recognition motifs (RRMs) and a C-terminal conserved sequence that interacts with the Lsm ring[13]. RRMs are ubiquitous, ~80 amino acid-long RNA-binding domains that typically recognize four single-stranded nucleotides[20–23]. Many RRM-containing proteins have multiple RRMs to enhance specificity and affinity for cognate RNAs *in vivo*[23], and can contain up to 6 RRMs[24]. In the case of Prp24, it is not clear how the multiple RRMs functionally cooperate to recognize U6 snRNA and promote annealing with U4 snRNA. However, we have shown that deletion of either the N- or C-terminal RRMs disrupts stable, stoichiometric binding of Prp24 to U6 RNA[25]. Furthermore, we previously determined the crystal and/or NMR structures of all four RRMs, including the non-canonical occluded RRM4 (oRRM4)[15], and identified the binding site of RRM2 in U6 RNA[14,26].

The strong temperature-dependence of RNA duplex stability makes cold-sensitive mutations a powerful tool for probing the dynamics of RNA base-pairing *in vivo*[1]. In previous work, we showed that the A62G mutation in U6, which stabilizes the base of the ISL (Fig. 1b), inhibits growth at low temperatures and decreases U4/U6 di-snRNP levels *in vivo*[12,27]. This cold-sensitive growth defect is suppressed by certain *cis*-acting mutations in U6 snRNA and *trans*-acting mutations in RRMs 2 and 3 of Prp24 (refs. 12,27). However, the mechanism of action of these suppressor mutations is poorly understood.

To investigate the mechanism by which Prp24 chaperones recycling of U6 snRNA, we determined the 1.7 Å-resolution crystal structure of the highly conserved central region of U6 RNA bound to 83% of the 444-residue Prp24 protein, including all four RRMs. The structure reveals a unique RNP architecture, with three RRMs encircling a large RNA loop to form interlocked protein and RNA rings. We also identified a large number of additional *trans*-acting U6-A62G-suppressor mutations in Prp24, most of which are predicted to disrupt specific protein-RNA contacts evident in the crystal structure. In addition to providing the first image of four contiguous RRMs bound to their RNA target and validating a genetic approach to defining an RNA-protein interface, the core U6 snRNP structure suggests how Prp24 may facilitate dissociation of U6 from U2 in post-catalytic spliceosomes and promote annealing of U6 with U4 RNA.

Results

Structure of U6 snRNA bound to Prp24

Based on our previously determined structures of domains of Prp24, and the locations of U6-A62G-suppressor mutations in U6 and Prp24 (see below), we designed a core U6 snRNP construct for crystallization. A complex consisting of U6-A62G nucleotides 30–101 (with U-to-C substitutions at positions 100 and 101) bound to wild-type Prp24 residues 34–400 was stable to purification by anion-exchange chromatography and formed crystals that diffracted X-rays to 1.7 Å resolution. The phase problem was solved by molecular replacement, using structures of individual RRM domains in Prp24[14] (Fig. 2, Table 1, **and**

Supplementary Figs. 1 and 2). This is the first reported crystal structure of U6 RNA, and the first structure of RNA bound to a protein with more than 2 RRMs (Supplementary Table 1).

The U6–Prp24 structure confirms the existence of the proposed telestem region in U6[12,19,28], spanning nucleotides 30–40 and 91–101 and including three non-canonical A-A or A-G pairs (Fig. 2 **and** Supplementary Movie). The ISL is highly similar to previous NMR structures (Supplementary Fig. 3a,b) and extends to include the invariant "AGC triad" (U6 residues 59-61)[29,30]. The telestem and ISL are roughly perpendicular to one another and are separated by an asymmetric internal loop or "bulge" spanning nucleotides A41–C58 (Fig. 2a). This bulge forms an extensive interface (~2,200 Å^2) with RRMs 2, 3 and oRRM4, as well as the region immediately preceding RRM1 in Prp24. This interface induces a highly distorted conformation of RNA that includes several novel ribonucleoprotein motifs that fall outside of the known 46 consensus clusters of RNA backbone suite conformations[31–33] (Fig. 3 **and** Supplementary Fig 3c). These novel motifs include a "skip-stack turn" (G50–A53) (Fig. 3a). This tight turn is in a region of the RNA that contains 4 consecutive C2′-endo sugar puckers (A49–G52). The skip-stack turn is located adjacent to the 5′ splice site-binding region of U6 (ref. 1) and is reminiscent of the "Z-anchor" motif that stabilizes RNA structure near the 5′ splice site of a group II self-splicing intron[34]. Both the skip-stack turn and Z-anchor have alternating stacked bases, but the former is protein-stabilized while the latter is stabilized by RNA. Another novel motif is the "dinucleotide bulge turn" (Fig. 3b), which bulges U57 and C58 to allow stacking of A56 and A59. The bulged U57 and C58, along with the 3′ side of the skip-stack turn, form a hydrophobic cage around Phe154 in RRM2 (Fig. 3c).

As observed in structures of the free protein[14,15,26], the beta-sheet of RRM1 is buried by extensive contacts with RRM2 and the beta-sheet of oRRM4 is occluded by a pair of flanking alpha-helices (Fig. 3d). Therefore, only RRM2 and RRM3 of Prp24 bind single-stranded RNA in the canonical RRM fashion[22]. RRM1 does not contact RNA within the crystallographic asymmetric unit (Fig. 2c), but does contact RNA belonging to a neighboring complex in the crystal (see below). RRM2 interacts with nucleotides 46–58, a region of U6 RNA that contains the highly conserved "ACAGA-box", which binds the intron 5′ splice site in assembled spliceosomes[1] (Fig. 4, Supplementary Movie). The RRM3 binding site spans U6 nucleotides 39–44, consistent with chemical modification studies predicting binding of Prp24 to U6 residues 40–43 (ref. 16). The occluding alpha-helices of oRRM4 contact both the ISL and telestem of U6, forming non-canonical contacts with double-stranded RNA and fixing the angle between the RNA helices (Figs. 2d **and** 3d).

The orientation of RRM3 is dramatically different than in the RRM1-3 crystal structure without U6 RNA[14], where an alpha-helix of RRM3 (residues 262-272) contacts the beta-sheet of RRM2 and partially occludes its canonical RNA-binding surface. Instead, in the U6–Prp24 structure, RRM3 undergoes an approximately 180° rotation and 20 Ångstrom displacement relative to RRM2, allowing RRM2 contacts with U6 RNA that would not be possible in the protein only structure. This finding is consistent with our earlier NMR studies, which showed that RRM2–RRM3 contacts present in the RNA-free crystal structure are not present in solution, and that RRM3 does not interfere with RNA binding to RRM2[26].

An interlocked RNP topology

The C-terminus of RRM3 is threaded through the asymmetric bulge of U6, placing oRRM4 on the opposite face of U6. Tertiary contacts between RRM2 and oRRM4 generate a “ring” of protein surrounding the U6 asymmetric bulge, which is also a closed ring due to the flanking telestem and ISL (schematic in Fig. 2c). The interlocked topology is further stabilized by formation of a novel “aspartate-bridged base pair” anchored by the last residue of RRM3 (Asp288), which forms hydrogen bonds with both A42 and G55 on opposite sides of the asymmetric bulge (Fig. 4a, see Fig. 6a). The aspartate bridge likely explains a previously reported UV cross-link between U6 nucleotide G55 and an unidentified amino acid in Prp24 (ref. 28).

The interlocked topology helps to explain the low nM apparent dissociation constant for U6 and Prp24 (ref. 25), and the stability of the complex to 2 M monovalent salt[6]. To our knowledge, this is the first reported observation of an interlocked RNP topology. However, multi-RRM proteins are very common[22,24,35,36]. Since the U6–Prp24 complex is the first determined structure of an RNP containing more than two linked RRM domains (Supplementary Table 1), it is likely that other examples of an interlocked RNP topology will be found.

The interlocked topology highlights the important role of cooperativity in RNA recognition by multi-domain proteins. For example, an NMR structure of isolated RRM2 bound to U6 nucleotides 49–54 (ref. 26) displays a binding mode that is shifted by 1 nucleotide relative to the crystal structure presented here. The shift can be explained by flanking RNA-protein interactions that were not present in the NMR study, including extensive contacts with the RRM2-3 linker, the R36 and R38 side chains that are N-terminal to RRM1, a tertiary contact between RRM2 and RRM4, and even the aspartate-bridged base pair (Figs. 2c, 3a,b,c, 4a and Supplementary Movie). Similarly, prior NMR chemical shift mapping studies of the isolated RRM3 domain indicated that this domain can bind to a double stranded U6 ISL construct with flanking single-stranded sequences[15]; however, we found that RRM3 contacts the telestem and adjacent 3′ single stranded region instead. Thus, it is clear that cooperativity plays a major role in determining binding mode, and structures of highly truncated proteins bound to RNA should be interpreted with caution.

Suppressors of U6-A62G map to the protein-RNA interface

We previously isolated 93 independent pseudo-revertant strains that suppress the cold-sensitive growth phenotype of the U6-A62G mutation[27]. Twenty-nine of these suppressor strains contain *cis*-acting mutations in U6 and have been described previously[27]. We now report that 32 of the remaining 64 suppressor strains harbor one of 30 different amino acid substitutions in Prp24 (Fig. 5a, Supplementary Table 2 and Supplementary Fig. 2). Sixteen of these substitutions were cloned, and all 16 were found to be sufficient for cold-resistant growth in the presence of either the A62G or more cold-sensitive A62U/C85A double mutation (referred to as “U6-UA”)[12,27] (Figs. 1b **and** 5b).

Mapping of the U6-A62G-suppressor mutations in both U6 (ref. 27) and Prp24 reveals a striking correspondence between the suppressor sites and protein-RNA contacts, indicating

that suppression is due to destabilization of the U6–Prp24 complex (Fig. 5c,d). For example, all four components of the aspartate bridge that anchors the RRM3–oRRM4 junction to the U6 asymmetric bulge (A42, C43, G55 and Asp288, Fig. 6a) are sites of suppressor substitutions. The observation that a conservative Asp288Glu substitution is sufficient for suppression (Fig. 5a) suggests that precise positioning of the side-chain carboxylate group relative to the protein backbone is required for the integrity of the aspartate bridge, consistent with the intricate network of hydrogen bonds that connect these residues. At the telestem–asymmetric bulge junction there are also numerous suppressor mutations in both RNA and protein residues (Fig. 5c,d). RRM3 residues Asn216, Asn253, and Ser283 form numerous hydrogen bonds with the last two base pairs of the telestem (Fig. 6b), and alterations to either the amino acid side chains or nucleotide bases suppresses cold-sensitivity.

RRM1 does not contact U6 RNA or contain suppressor mutations. However, Arg38, which is immediately adjacent to RRM1, contacts the U6 ACAGA-box (Fig. 6c) and is the site of a suppressor mutation. The absence of corresponding *cis*-acting suppressor mutations in this region of U6 (ref. 27) likely arises from the severe growth defects caused by mutation of the ACAGA-box[37]. Amino acid residues in the RRM2–oRRM4 interface that stabilize a "dinucleotide bulge turn" involving U6 residues U54 through A59 (Fig. 3b), including Phe154, Arg295, Ser350, and Asp351, are also altered in suppressor strains (Figs. 5a **and** 6d). All of the *cis*-acting suppressor substitution sites in the ISL, as well as U38 in the telestem, do not contact Prp24; these substitutions likely act by destabilizing their resident helix.

Together, these findings validate the use of high-density suppressor selections for predicting an RNA-protein interface. Additionally, co-localization of suppressor sites with the protein-RNA interface suggests that global de-stabilization of the U6 snRNP can compensate for local hyperstabilization of the U6 ISL. This notion is in accord with the hypothesis of Guthrie and colleagues that mutations in *PRP24* suppress a related cold-sensitive mutation in U4 snRNA (G14C, in U4/U6 Stem II) by weakening the interaction of Prp24 with U6 RNA, so that U6 can more readily pair with U4 (refs. 6,16). Indeed, the three Prp24 substitutions previously selected as suppressors of U4-G14C also suppress U6-A62G and U6-UA (Supplementary Fig. 4).

An electropositive groove in Prp24 binds double-helical RNA

Although RRM1 does not contact U6 in the U6–Prp24 complex, it is important for Prp24 function as alanine substitution of a strictly conserved phenylalanine in the RRM1 core (F87A; Supplementary Fig. 2) is lethal at 37 °C and slow-growing at 30 °C[25]. We therefore inspected crystal-packing contacts for clues to the functional role of RRM1. Interestingly, a roughly 20 Å-wide electropositive groove bordered by RRMs 1, 2 and oRRM4 contacts the U6 ISL from a neighboring RNP complex in the crystal lattice (Fig. 7). We previously showed by NMR that the same electropositive surface in a truncated Prp24 protein containing only RRMs 1 and 2 binds to a 12-nucleotide single-stranded RNA corresponding to U6 residues 49–60, and proposed that RRM1 might chaperone annealing of unwound U6 with U4 snRNA[26]. The presence of this surface of RRM1 on one face of a double-stranded

RNA-binding groove further supports the notion that the electropositive groove is a site for RNA annealing (see Discussion).

Discussion

Function of Prp24 as a U6 snRNA chaperone

U6 snRNA is at the catalytic center of the spliceosome[3]. Thus, its assembly into and disassembly from the spliceosome must be precisely executed to maintain the fidelity of pre-mRNA splicing. U6 snRNA is released from U2 snRNA at the completion of each splicing cycle, so that it can regenerate the U6 snRNP and allow subsequent assembly of the U4/U6 di-snRNP. The observed interactions of U6 nucleotides 39–44 with RRM3 of Prp24 and nucleotides 54–58 with RRM2 and oRRM4 are expected to be mutually exclusive with base pairing in the U2/U6 di-snRNA[38]. We therefore propose that Prp24 has two functions in chaperoning U6 through the splicing cycle: promoting dissociation of the U2/U6 RNA complex, and promoting association of the U4/U6 RNA complex (Fig. 1a). However, it is not yet clear if Prp24 binds to and actively unwinds the U2/U6 complex, or simply sequesters U6 RNA that has been displaced from U2 snRNA by other means, for example, Brr2 helicase activity[39,40]. The former mechanism is particularly attractive as U2/U6 base pairing does not preclude binding of Prp24 RRM2 to the U6 intron-binding site, raising the possibility that a Prp24–U2/U6 ternary complex acts to prevent reassociation of product RNA and reverse splicing[41]. Such a complex could be stabilized by binding of the U6 5′ stem into the electropositive groove delineated by RRMs 1, 2 and oRRM4, as it possesses electrostatic and spatial features compatible with binding of double-stranded RNA (Fig. 7).

The RRM1–2–4 electropositive groove is also a prime candidate for stabilization of U4/U6 duplex RNA. After dissociation from U2, interaction of the U6 ACAGA-box with RRM2 holds the U6 component of U4/U6 Stem I in close proximity to this groove, suggesting U4/U6 pairing is nucleated by stabilizing Stem I within this electropositive scaffold. However, Stem I nucleotides of U6 RNA (54-62) are not accessible for pairing with U4 in the structure presented here (Fig. 8a). We propose that the previously determined RRM1–RRM2 co-structure with a fragment of U6 RNA[26], in which nucleotides 54–60 are bound along one face of the electropositive groove (superposition used to construct Figure 8b), represents an on-pathway intermediate during U4/U6 annealing. Thus, as the ISL is unwound by thermal motion, reannealing of the ISL could be inhibited by favorable interactions between the separated strands and the electropositive surface of Prp24. These interactions could include specific binding of the CCCU sequence in the 5′ strand of the upper ISL, which was previously identified as a top SELEX target of Prp24[10]. In addition to unwinding of the ISL, transition from the ground-state U6 snRNP structure to the proposed intermediate would require breaking Prp24 contacts with the Stem I region of U6 RNA. The presence of U6-A62G-suppressor mutations in Prp24 residues that contact these U6 residues (nucleotides 54–58; Figs. 5c,d **and** 6d) raises the possibility that at least some suppressors operate by promoting the transition between the ground-state U6 snRNP and the intermediate proposed in Figure 8b.

Mutational perturbation of equilibria in the splicing cycle

The stable interaction of Prp24 with U6 snRNA is at odds with its function as a U4/U6 annealing chaperone, since it must relinquish U6 to U4 snRNA, just as U4 must subsequently relinquish U6 to U2 snRNA. The splicing cycle is made up of numerous such equilibria, for which a productive direction is favored *in vivo*, but can be reversed *in vitro*[41]. This property of the splicing cycle likely explains its temperature-dependent sensitivity to minor perturbations in RNA-RNA interactions (exemplified by the U6-A62G and U6-UA mutations) and RNA-protein interactions (exemplified by the U6-A62G-suppressor mutations in Prp24). The fact that 32 independently selected suppressors in Prp24 resulted in 30 unique substitutions illustrates both the multitude of contacts that stabilize a protein-RNA interface, and the utility of this genetic approach in identifying such contacts. We are actively investigating the identity of the 32 *trans*-acting suppressors that do not map to the *PRP24* gene. None of these suppressor strains contains mutations in the U2 or U4 genes, so the mutations most likely lie in another protein or proteins that influence U6 snRNA equilibria.

By identifying appropriate conditional mutations with which to select suppressors, this approach can be extended to other equlibria in the splicing cycle[42,43]. In addition to genome-wide spontaneous suppressor selections, as used here, gene-targeted selections are useful to probe specific interactions in detail. We previously conducted a targeted selection for mutations in the Prp8 protein that suppress a cold-sensitive mutation in U4 RNA that blocks spliceosome activation[44]. Most of these suppressors map to one face of a recent crystal structure of a large portion of Prp8 (ref. 45). It will be of interest to see if these suppressor mutations also lie on an RNA-protein interface.

Architecture of the U6 snRNP

We have determined the first crystal structure of U6 snRNA bound to the Prp24 protein, containing seventy nucleotides of U6 RNA and all four RRM domains of Prp24. The observed structure is consistent with previous chemical modification data[28]. For example, nucleotides 40-42, 49, 50, 52, and 55 are protected from chemical modification despite being in a single-stranded region of RNA, which is explained by the fact that these nucleotides are buried by interactions with Prp24 (Supplementary Fig. 5). Conversely, the Watson-Crick faces of nucleotides 45, 47, 51, 53 and 54 of the asymmetric bulge are all solvent exposed in the complex, and reactive to chemicals[28]. The close agreement with solution studies, along with the fact that the genetic suppressors localize to the RNA-protein interface, provides compelling evidence that the crystal structure accurately reflects the *in vivo* core U6 snRNP. This complex likely acts as a structural foundation for the remaining RNA and protein elements in the U6 snRNP.

The interlocked topology of the U6–Prp24 complex raises questions regarding the assembly pathway of the U6 snRNP, as the single-stranded asymmetric bulge of U6 is of insufficient width to allow threading of oRRM4 or RRMs 1–3. We propose that transient unwinding of the relatively unstable telestem allows wrapping of Prp24 around an unclosed loop of single-stranded RNA, after which the telestem reforms to generate the interlocked ribonucleoprotein structure (Supplementary Fig. 6). Notably, such a process would require

transient disruption of the RRM2–oRRM4 interaction as well, which is presumed to be considerably less stable in the absence of the U6 snRNA scaffold that mediates contacts between RRM2 and oRRM4 in the assembled snRNP. Another possible mechanism is direct recognition of the U2/U6 complex by Prp24 after release of the lariat intron, allowing U6 nucleotides 49–53 to bind to Prp24 RRM2. The telestem would then form upon dissociation of U2 RNA from U6, thereby generating the interlocked topology.

Extant biochemical data and the recently reported Lsm structure[9,13,28,46] can be used to generate a model of the complete U6 snRNP (Supplementary Fig. 7). In this model, the Lsm ring binds the base of the U6 telestem, consistent with observed crosslinks between U6 nucleotide G30 and Lsm2 (ref. 9) and proximity of the ring to the 3′ tail of U6 (ref. 46). This placement of the Lsm ring is also compatible with observed interaction of the extreme C-terminus of Prp24 with Lsm5, 7 and 8 (ref. 13), which could explain additional crosslinks between U6 nucleotides U28 and U29 with unidentified residues in Prp24 (ref. 28). Correct placement of the U6 5′ stem (missing from the model in Supplementary Fig. 7) is not apparent when analyzing the U6–Prp24 structure and extant biochemical data, although chemical protection studies strongly suggest the 5′ stem contacts Prp24 (ref. 28). We propose the 5′ stem could bind into the RRM1–2–4 electropositive groove, although this placement would likely require partial unwinding of the telestem. This proposal in turn raises the interesting possibility that binding of the U6 snRNP to the U4 snRNP could cause a conformational change that results in displacement of the U6 5′ stem from the electropositive groove, thereby allowing formation of U4/U6 Stem I (Fig. 8). The location of the U6 5′ stem will be resolved by the structure of the complete U6 snRNP, which will likely be elucidated in the near future since all individual components can now be made at yield and homogeneity suitable for crystallographic analysis.

Implications for ribonucleoprotein structure and function

An interesting feature of Prp24 revealed by our study is its ability to bind single- and double-stranded RNA. RRMs are generally considered single-stranded RNA binding domains, yet both RRM3 and oRRM4 bind the single-stranded asymmetric bulge as well as the adjacent helical regions. oRRM4 in particular contacts both the ISL and telestem, via its non-canonical N- and C-terminal alpha-helices. Thus, multi-RRM proteins are not restricted to direct read-out of adjacent 3-4 nucleotide base sequences, but rather can evolve to stabilize highly complex RNA secondary and tertiary structures. These mixed binding mechanisms could help direct the structural rearrangements necessary for U4/U6 di-snRNP assembly, which must involve unwinding of the U6 ISL[15] and may involve changes in U4 snRNA secondary structure as well. We predict that mixed single- and double-stranded RNA binding modes will be a common feature of proteins that act as RNA chaperones during spliceosome assembly, activation, and disassembly[47,48].

The U6–Prp24 complex reveals a striking degree of structural co-evolution of protein and RNA. For example, tertiary contacts between Prp24 RRM2 and oRRM4 are stabilized by the asymmetric bulge of U6 (Fig. 3c), the conformation of which is stabilized by extensive contacts with RRM2, RRM3 and oRRM4 (Figs. 3a,b **and** 4a). Furthermore, the RNA components of several novel structural motifs presented here, including the “skip-stack

turn", "dinucleotide bulge turn" and "aspartate bridge" (Fig. 3a,b and Fig. 6a) are all built upon a protein scaffold. The intricate and cooperative nature of these contacts implies a lengthy process of evolutionary co-adaptation. There is also a striking lack of tightly bound divalent cation along the protein-RNA interface, suggesting Prp24 has supplanted the stabilizing role of magnesium typically observed in ribonucleoproteins thought to have originated in a protein-free, RNA-only world, such as the ribosome and Ribonuclease P. Potentially, this characteristic could give Prp24 more control over the stability of helices in U6 snRNA, aiding its function as a chaperone for the structural transitions of U6.

Together, these findings advance our understanding of how multi-RRM proteins specifically recognize long RNAs and promote assembly of large ribonucleoproteins. The combined structural and genetic approach used here can serve as a model for future investigation of multi-RRM proteins.

Online Methods

Protein Production

E.coli STAR pLysS cells (Invitrogen) harboring a modified pET3a plasmid (Novagen) were used to synthesize residues 34-400 of Prp24 with a non-cleavable hexahistidine tag appended to the C-terminus of the protein. The cells were grown at 37 °C with shaking to mid-logarithmic phase and protein expression induced through addition of isopropyl β-D-thiogalactopyranoside (IPTG) to 1 mM. The induced cells were then grown for an additional 20 hours at 25 °C and harvested by centrifugation at 4,000 *x g* for 10 minutes. Five gram cell pellets were resuspended in 30 mL immobilized metal affinity chromatography (IMAC) buffer (500 mM NaCl, 50 mM HEPES acid, 50 mM imidazole base, 10 % (v/v) glycerol, 1 mM tris(2-carboxyethyl)phosphine (TCEP) HCl). Ten mg lysozyme and 0.25 mg DNase I (Sigma) were then added to 30 mL of the resuspended cells and the mixture stored at −80 °C. The cell suspension was thawed at room temperature and sonicated on ice prior to clarification of the lysate by centrifugation at 20,000 *x g* for 30 minutes. The soluble fraction was loaded onto a Ni^{2+}-charged nitrilotriacetic acid (NTA) agarose resin (Qiagen) that had been pre-equilibrated with IMAC buffer. The column was washed with 50 mL fresh IMAC buffer and Prp24 desorbed using IMAC buffer supplemented with 500 mM imidazole pH 7.0. The purified protein was dialyzed overnight at 4 °C against 1 L of cation-exchange chromatography buffer (100 mM NaCl, 10 mM HEPES acid, 10 mM sodium HEPES, 10 % glycerol, 1 mM TCEP HCl, pH 7.0) supplemented with 1 mM ethylenediaminetetraacetic acid (EDTA), and further purified using cation-exchange chromatography with salt gradient elution on an AKTA FPLC system equipped with a HiTrapSP column (GE Healthcare). The protein eluted at approximately 300 mM salt and was concentrated to approximately 10 mg/mL using a spin concentrator with a 10 kDa cutoff. The protein was stored at −80 °C. UV absorption was used to estimate the final protein concentration, using an extinction coefficient of 20,400 $M^{-1}cm^{-1}$ at 280 nm[49]. The measured A_{280}/A_{260} ratio was used to further estimate an extinction coefficient of 12,120 $M^{-1}cm^{-1}$ at 260 nm, which was subsequently used to estimate the concentration of the U6•Prp24 complex (see below).

RNA Production

In vitro transcription was used to synthesize nucleotides 30-101 of an A62G/U100C/U101C mutant form of the U6 snRNA, using recombinant his-tagged T7 RNA polymerase and synthetic DNA oligonucleotide templates (Integrated DNA Technologies). After transcription, the target RNA was resolved from abortive transcripts and "n+1" additions via denaturing 8 % polyacrylamide gel electrophoresis in 8 M Urea. The RNA was extracted from the gel by passive diffusion into a solution containing 100 mM sodium acetate, pH 5.6. Ethanol precipitation was used to concentrate the RNA prior to further purification by anion exchange chromatography (High Q column, Bio-Rad) in 20 mM Tris pH 7.6, 300 mM sodium chloride, with gradient elution up to 1.5 M sodium chloride. Finally, the RNA was dialyzed using a 10 kDa membrane into a storage buffer containing 100 mM sodium chloride, 10 mM monobasic potassium phosphate, 10 mM dibasic potassium phosphate, 1 mM disodium EDTA and 1 mM sodium azide. 6 % native PAGE analysis showed a small fraction of the purified RNA (ca. 5 %) was in the form an intermolecular dimer. The final yield from a 20 mL transcription reaction was approximately 5 mg of pure RNA. The RNA concentration was estimated using UV absorption and an extinction coefficient of 891,266 $M^{-1}cm^{-1}$ at 260 nm[50].

Reconstitution of the U6-Prp24 complex

Equimolar amounts of purified protein and RNA were mixed together (1 mL RNA at 10 μM in RNA storage buffer added to 0.05 mL protein at 200 μM in cation-exchange elution buffer) and loaded directly onto a MonoQ column that had been pre-equilibrated with MonoQ buffer at 4 °C (100 mM potassium chloride, 10 mM Tris base, 10 mM Tris HCl, 2 mM $MgCl_2$, 1 mM TCEP HCl and 5 % glycerol, pH ~ 8). Under these conditions, free protein flows through the column while ribonucleoprotein and free RNA bind to the column. A linear gradient to 2 M potassium chloride in MonoQ buffer was applied to the column to differentially desorb ribonucleoprotein and RNA. Non-denaturing polyacrylamide gel electrophoresis was used to assess homogeneity of the ribonucleoprotein, and sequential staining/destaining with toluidine blue and coomassie G-250 confirmed the presence of both protein and RNA in a single gel band. The eluted complex was concentrated using a spin concentrator with a 10 kDa cutoff to 5 mg/mL in MonoQ elution buffer, with an approximate concentration of potassium chloride in the sample of 400 mM being determined from measured conductivity. The concentration of the complex was estimated by UV absorption, using the sum of protein and RNA extinction coefficients at 260 nm (903,386 $M^{-1}cm^{-1}$).

Crystallization and Data Collection

Crystallization conditions for the purified U6•Prp24 complex were identified using a Mosquito high-throughput crystallization robot (TTP LabTech). Crystals were obtained by sitting drop vapor diffusion at 4 °C from an initial drop containing 0.2 uL concentrated U6•Prp24 complex and 0.2 uL crystallization reagent (100 mM lithium sulfate, 100 mM sodium citrate, pH 5.5 and 20 % PEG 1,000). In order to achieve cryogenic preservation, the crystal-containing drop was diluted against excess crystallization reagent supplemented with 20 % (v/v) glycerol. Quickly after dilution of the drop, a single crystal was harvested and

flash cooled in liquid nitrogen. Diffraction data were collected at 100 K on beamline 21-ID-F at the Advanced Photon Source. 360 images were recorded on a MarMosaic 225 detector (Rayonix) at 1 degree oscillations with X-rays tuned to a wavelength of 0.97872 Å. The diffraction data were indexed and scaled using HKL2000[51]. The space group was determined using POINTLESS[52] and Xtriage[53] was used to assay potential twinning of the data.

Structure Determination, Refinement and Analysis

Initial phases were obtained by molecular replacement in Phaser[54], using residues 41-196 and 206-291 from PDB 2GHP[14] as independent search models, which yielded an initial map of suitable quality for iterative, manual model building in Coot[55] with interspersed cycles of automated refinement and phase improvement in Phenix[53]. The final model contains approximately seven-hundred bound water molecules, many of which are visible along the single stranded RNA-protein interface and a single bound sulfate ion from the crystallization mixture. Typical magnesium bond lengths and geometries[56,57] were used to discriminate magnesium from water in the structure. There are residual densities in the final mF_o-DF_c maps associated with the backbone of U6 ISL (nucleotides 63-86) and the base of the telestem (nucleotides 30-36 and 97-100) that could not be readily modeled through alternate conformers of RNA or bound solutes. In these regions, a single conformer of RNA was fit into unbiased electron density derived from simulated-annealing omit maps and TLS[58] restraints were imposed to model apparent structural dynamics. The unusually high overall atomic displacement parameters in the final model are due in large part to regions of the RNA modeled under TLS restrains, which do not form extensive contacts with the Prp24 protein. The final model has an overall MolProbilty[33] score of 1.7 and over 98 % of all protein residues in allowed regions of a Ramachandran plot[59]. The interface area between single-stranded RNA (U6 nucleotides 41-56) and Prp24 was determined using the online PDBePISA[60] tool available at www.ebi.ac.uk/msd-srv/prot_int/pistart.html. All figures were generated in PyMOL[61]. The electrostatic surface was calculated using APBS[62] as implemented in PyMOL. A sample of the final electron density map is shown in Supplementary Figure 1.

Genetic Suppressor Selection and Analysis

Ninety-five independent spontaneous cold-resistant revertants of a yeast strain (DAB016) bearing a deletion of the chromosomal U6 RNA gene and containing the U6-A62G allele on a plasmid were previously selected[27]. Of these, 31 have a *de novo* mutation in the U6-A62G allele that is responsible for cold-resistance[27]. Genomic DNA was prepared from the remaining 64 strains and the *PRP24* locus was amplified by PCR with primers that generate a 1782 bp DNA fragment spanning from 265 bp upstream of the start codon to 183 bp downstream of the stop codon. The PCR product from each strain was subjected to Sanger sequencing that spanned the entire protein-coding region on at least one strand to identify mutations (Supplementary Table 2). (The wild-type *PRP24* coding region in DAB016 contains 16 single-nucleotide polymorphisms relative to the S288C sequence, 10 of which result in amino acid substitutions.)

Sixteen of the mutant alleles (Supplementary Table 2) were cloned by digestion of the PCR product with BsrGI and SnaBI, followed by ligation into pRS313[63]. Mutant *prp24* alleles in pRS313 were co-transformed with pRS314 containing wild-type *SNR6*, *snr6-A62G*, or *snr6-A62U/C85A* into the *SNR6/PRP24* double-disruption strain LL200 (*MATa*, *snr6ΔBsm*∷*LEU2*, *prp24ΔClaSnaB*∷*ADE2*, *ade2-1*, *can1-100*, *his3-11*, *leu2-3,112*, *lys2-Δ2*, *met2-Δ1*, *trp1-1*, *ura3-52;*[pUN50-PRP24] and [YCp50-SNR6][12]). Transformants were selected on –trp –his medium and plated to synthetic complete medium containing 5-fluoroorotate to select for loss of the *URA3*-marked plasmids containing wild-type alleles of *SNR6* and *PRP24*. Colonies were picked to 5 ml YEPD, grown overnight at 30°C and adjusted to an A_{600} = 1.0. The undiluted culture and successive 10-fold serial dilutions were spotted to YEPD plates using a 48-pin inoculation manifold and the plates were incubated at 16, 23, 30 and 37 °C.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We are grateful to J. Doudna, C. Guthrie, A. Hoskins, H. Noller, A. M. Pyle and members of the Brow and Butcher laboratories for helpful discussions and critical reading of the manuscript, C. Bingman for advice and technical assistance in performing crystallization screening, and B. Bhattacharyya and J. Keck for acquisition of diffraction data. Use of the Advanced Photon Source, an Office of Science User Facility operated for the U.S. Department of Energy (DOE) Office of Science by Argonne National Laboratory, was supported by the U.S. DOE under Contract No. DE-AC02-06CH11357. Use of the LS-CAT Sector 21 was supported by Grant 085P1000817. This work was supported by a grant from the U.S. National Institutes of Health (no. GM065166 to S.E.B and D.A.B).

References

1. Brow DA. Allosteric cascade of spliceosome activation. Annu Rev Genet. 2002; 36:333–60. [PubMed: 12429696]
2. Shukla GC, Padgett RA. A catalytically active group II intron domain 5 can function in the U12-dependent spliceosome. Mol Cell. 2002; 9:1145–50. [PubMed: 12049749]
3. Fica SM, et al. RNA catalyses nuclear pre-mRNA splicing. Nature. 2013; 503:229–34. [PubMed: 24196718]
4. Marcia M, Pyle AM. Visualizing group II intron catalysis through the stages of splicing. Cell. 2012; 151:497–507. [PubMed: 23101623]
5. Brow DA, Guthrie C. Spliceosomal RNA U6 is remarkably conserved from yeast to mammals. Nature. 1988; 334:213–218. [PubMed: 3041282]
6. Shannon KW, Guthrie C. Suppressors of a U4 snRNA mutation define a novel U6 snRNP protein with RNA-binding motifs. Genes Dev. 1991; 5:773–85. [PubMed: 1827420]
7. 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. [PubMed: 10428970]
8. Stevens SW, et al. Biochemical and genetic analyses of the U5, U6, and U4/U6 × U5 small nuclear ribonucleoproteins from *Saccharomyces cerevisiae*. RNA. 2001; 7:1543–53. [PubMed: 11720284]
9. Karaduman R, et al. Structure of yeast U6 snRNPs: arrangement of Prp24p and the LSm complex as revealed by electron microscopy. RNA. 2008; 14:2528–37. [PubMed: 18971323]
10. Ghetti A, Company M, Abelson J. Specificity of Prp24 binding to RNA: a role for Prp24 in the dynamic interaction of U4 and U6 snRNAs. RNA. 1995; 1:132–45. [PubMed: 7585243]
11. Raghunathan PL, Guthrie C. A Spliceosomal Recycling Factor That Reanneals U4 and U6 Small Nuclear Ribonucleoprotein Particles. Science. 1998; 279:857–860. [PubMed: 9452384]

12. Vidaver RM, Fortner DM, Loos-Austin LS, Brow DA. Multiple functions of *Saccharomyces cerevisiae* splicing protein Prp24 in U6 RNA structural rearrangements. Genetics. 1999; 153:1205–18. [PubMed: 10545453]
13. Rader SD, Guthrie C. A conserved Lsm-interaction motif in Prp24 required for efficient U4/U6 di-snRNP formation. RNA. 2002; 8:1378–92. [PubMed: 12458792]
14. Bae E, et al. Structure and interactions of the first three RNA recognition motifs of splicing factor Prp24. J Mol Biol. 2007; 367:1447–58. [PubMed: 17320109]
15. Martin-Tumasz S, Richie AC, Clos LJ 2nd, Brow DA, Butcher SEA. novel occluded RNA recognition motif in Prp24 unwinds the U6 RNA internal stem loop. Nucleic Acids Res. 2011; 39:7837–47. [PubMed: 21653550]
16. Jandrositz A, Guthrie C. Evidence for a Prp24 binding site in U6 snRNA and in a putative intermediate in the annealing of U6 and U4 snRNAs. EMBO J. 1995; 14:820–32. [PubMed: 7882985]
17. Achsel T, et al. 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; 18:5789–802. [PubMed: 10523320]
18. Verdone L, Galardi S, Page D, Beggs JD. Lsm proteins promote regeneration of pre-mRNA splicing activity. Curr Biol. 2004; 14:1487–91. [PubMed: 15324666]
19. Ryan DE, Stevens SW, Abelson J. The 5′ and 3′ domains of yeast U6 snRNA: Lsm proteins facilitate binding of Prp24 protein to the U6 telestem region. RNA. 2002; 8:1011–33. [PubMed: 12212846]
20. Dreyfuss G, Swanson MS, Pinol-Roma S. Heterogeneous nuclear ribonucleoprotein particles and the pathway of mRNA formation. Trends Biochem Sci. 1988; 13:86–91. [PubMed: 3072706]
21. Kenan DJ, Query CC, Keene JD. RNA recognition: towards identifying determinants of specificity. Trends Biochem Sci. 1991; 16:214–20. [PubMed: 1716386]
22. Daubner GM, Clery A, Allain FH. RRM-RNA recognition: NMR or crystallography…and new findings. Curr Opin Struct Biol. 2013; 23:100–8. [PubMed: 23253355]
23. Lunde BM, Moore C, Varani G. RNA-binding proteins: modular design for efficient function. Nat Rev Mol Cell Biol. 2007; 8:479–90. [PubMed: 17473849]
24. Punta M, et al. The Pfam protein families database. Nucleic Acids Res. 2012; 40:D290–301. [PubMed: 22127870]
25. Kwan SS, Brow DA. The N- and C-terminal RNA recognition motifs of splicing factor Prp24 have distinct functions in U6 RNA binding. RNA. 2005; 11:808–20. [PubMed: 15811912]
26. Martin-Tumasz S, Reiter NJ, Brow DA, Butcher SE. Structure and functional implications of a complex containing a segment of U6 RNA bound by a domain of Prp24. RNA. 2010; 16:792–804. [PubMed: 20181740]
27. Fortner DM, Troy RG, Brow DA. A stem/loop in U6 RNA defines a conformational switch required for pre-mRNA splicing. Genes Dev. 1994; 8:221–33. [PubMed: 8299941]
28. Karaduman R, Fabrizio P, Hartmuth K, Urlaub H, Luhrmann R. RNA structure and RNA-protein interactions in purified yeast U6 snRNPs. J Mol Biol. 2006; 356:1248–62. [PubMed: 16410014]
29. Huppler A, Nikstad LJ, Allmann AM, Brow DA, Butcher SE. Metal binding and base ionization in the U6 RNA intramolecular stem-loop structure. Nat Struct Biol. 2002; 9:431–5. [PubMed: 11992125]
30. Sashital DG, Cornilescu G, McManus CJ, Brow DA, Butcher SE. U2-U6 RNA folding reveals a group II intron-like domain and a four-helix junction. Nat Struct Mol Biol. 2004; 11:1237–42. [PubMed: 15543154]
31. Duarte CM, Wadley LM, Pyle AM. RNA structure comparison, motif search and discovery using a reduced representation of RNA conformational space. Nucleic Acids Res. 2003; 31:4755–61. [PubMed: 12907716]
32. Richardson JS, et al. RNA backbone: consensus all-angle conformers and modular string nomenclature (an RNA Ontology Consortium contribution). RNA. 2008; 14:465–81. [PubMed: 18192612]
33. Chen VB, et al. MolProbity: all-atom structure validation for macromolecular crystallography. Acta Crystallogr D. 2010; 66:12–21. [PubMed: 20057044]

34. Toor N, Keating KS, Taylor SD, Pyle AM. Crystal structure of a self-spliced group II intron. Science. 2008; 320:77–82. [PubMed: 18388288]
35. Ray D, et al. A compendium of RNA-binding motifs for decoding gene regulation. Nature. 2013; 499:172–7. [PubMed: 23846655]
36. Clery A, Blatter M, Allain FH. RNA recognition motifs: boring? Not quite Curr Opin Struc Biol. 2008; 18:290–8.
37. Madhani HD, Bordonne R, Guthrie C. Multiple roles for U6 snRNA in the splicing pathway. Genes Dev. 1990; 4:2264–77. [PubMed: 2149118]
38. Burke JE, Sashital DG, Zuo X, Wang YX, Butcher SE. Structure of the yeast U2/U6 snRNA complex. RNA. 2012; 18:673–83. [PubMed: 22328579]
39. Xu D, Nouraini S, Field D, Tang SJ, Friesen JD. An RNA-dependent ATPase associated with U2/U6 snRNAs in pre-mRNA splicing. Nature. 1996; 381:709–13. [PubMed: 8649518]
40. Small EC, Leggett SR, Winans AA, Staley JP. The EF-G-like GTPase Snu114p regulates spliceosome dynamics mediated by Brr2p, a DExD/H box ATPase. Mol Cell. 2006; 23:389–99. [PubMed: 16885028]
41. Tseng CK, Cheng SC. Both catalytic steps of nuclear pre-mRNA splicing are reversible. Science. 2008; 320:1782–4. [PubMed: 18583613]
42. Wells SE, et al. CUS1, a suppressor of cold-sensitive U2 snRNA mutations, is a novel yeast splicing factor homologous to human SAP 145. Genes Dev. 1996; 10:220–32. [PubMed: 8566755]
43. Kuhn AN, Li Z, Brow DA. Splicing factor Prp8 governs U4/U6 RNA unwinding during activation of the spliceosome. Mol Cell. 1999; 3:65–75. [PubMed: 10024880]
44. Kuhn AN, Brow DA. Suppressors of a cold-sensitive mutation in yeast U4 RNA define five domains in the splicing factor Prp8 that influence spliceosome activation. Genetics. 2000; 155:1667–82. [PubMed: 10924465]
45. Galej WP, Oubridge C, Newman AJ, Nagai K. Crystal structure of Prp8 reveals active site cavity of the spliceosome. Nature. 2013; 493:638–43. [PubMed: 23354046]
46. Zhou L, et al. Crystal structures of the Lsm complex bound to the 3′ end sequence of U6 small nuclear RNA. Nature. 2014; 506:116–20. [PubMed: 24240276]
47. Staley JP, Guthrie C. Mechanical devices of the spliceosome: motors, clocks, springs, and things. Cell. 1998; 92:315–26. [PubMed: 9476892]
48. Ares M Jr, Weiser B. Rearrangement of snRNA structure during assembly and function of the spliceosome. Prog Nucleic Acid Re. 1995; 50:131–59.
49. Wilkins MR, et al. Protein identification and analysis tools in the ExPASy server. Methods Mol Biol. 1999; 112:531–52. [PubMed: 10027275]
50. Kibbe WA. OligoCalc: an online oligonucleotide properties calculator. Nucleic Acids Res. 2007; 35:W43–6. [PubMed: 17452344]
51. Otwinowski Z, Minor W. Processing of X-ray diffraction data collected in oscillation mode. Method Enzymol. 1997; 276:307–326.
52. Evans P. Scaling and assessment of data quality. Acta Crystallogr D. 2006; 62:72–82. [PubMed: 16369096]
53. Adams PD, et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr D. 2010; 66:213–21. [PubMed: 20124702]
54. McCoy AJ, et al. Phaser crystallographic software. J Appl Crystallogr. 2007; 40:658–674. [PubMed: 19461840]
55. Emsley P, Cowtan K. Coot: model-building tools for molecular graphics. Acta Crystallogr D. 2004; 60:2126–32. [PubMed: 15572765]
56. Petrov AS, Bowman JC, Harvey SC, Williams LD. Bidentate RNA-magnesium clamps: on the origin of the special role of magnesium in RNA folding. RNA. 2011; 17:291–7. [PubMed: 21173199]
57. Klein DJ, Moore PB, Steitz TA. The contribution of metal ions to the structural stability of the large ribosomal subunit. RNA. 2004; 10:1366–79. [PubMed: 15317974]

58. Winn MD, Isupov MN, Murshudov GN. Use of TLS parameters to model anisotropic displacements in macromolecular refinement. Acta Crystallogr D. 2001; 57:122–133. [PubMed: 11134934]

59. Ramachandran GN, Ramakrishnan C, Sasisekharan V. Stereochemistry of polypeptide chain configurations. J Mol Biol. 1963; 7:95–9. [PubMed: 13990617]

60. Krissinel E, Henrick K. Detection of Protein Assemblies in Crystals. Lect Notes Comput Sc. 2005; 3695:163–174.

61. Schrodinger LLC. The PyMOL Molecular Graphics System, Version 1.3r1. 2010

62. Baker NA, Sept D, Joseph S, Holst MJ, McCammon JA. Electrostatics of nanosystems: application to microtubules and the ribosome. Proc Natl Acad Sci U S A. 2001; 98:10037–41. [PubMed: 11517324]

63. Sikorski RS, Hieter P. A system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in Saccharomyces cerevisiae. Genetics. 1989; 122:19–27. [PubMed: 2659436]

64. McManus CJ, Schwartz ML, Butcher SE, Brow DA. A dynamic bulge in the U6 RNA internal stem-loop functions in spliceosome assembly and activation. RNA. 2007; 13:2252–65. [PubMed: 17925343]

65. Yean SL, Wuenschell G, Termini J, Lin RJ. Metal-ion coordination by U6 small nuclear RNA contributes to catalysis in the spliceosome. Nature. 2000; 408:881–4. [PubMed: 11130730]

66. Venditti V, Clos L 2nd, Niccolai N, Butcher SE. Minimum-energy path for a U6 RNA conformational change involving protonation, base-pair rearrangement and base flipping. J Mol Biol. 2009; 391:894–905. [PubMed: 19591840]

67. Kennedy SD, Kierzek R, Turner DH. Novel conformation of an RNA structural switch. Biochemistry. 2012; 51:9257–9. [PubMed: 23134175]

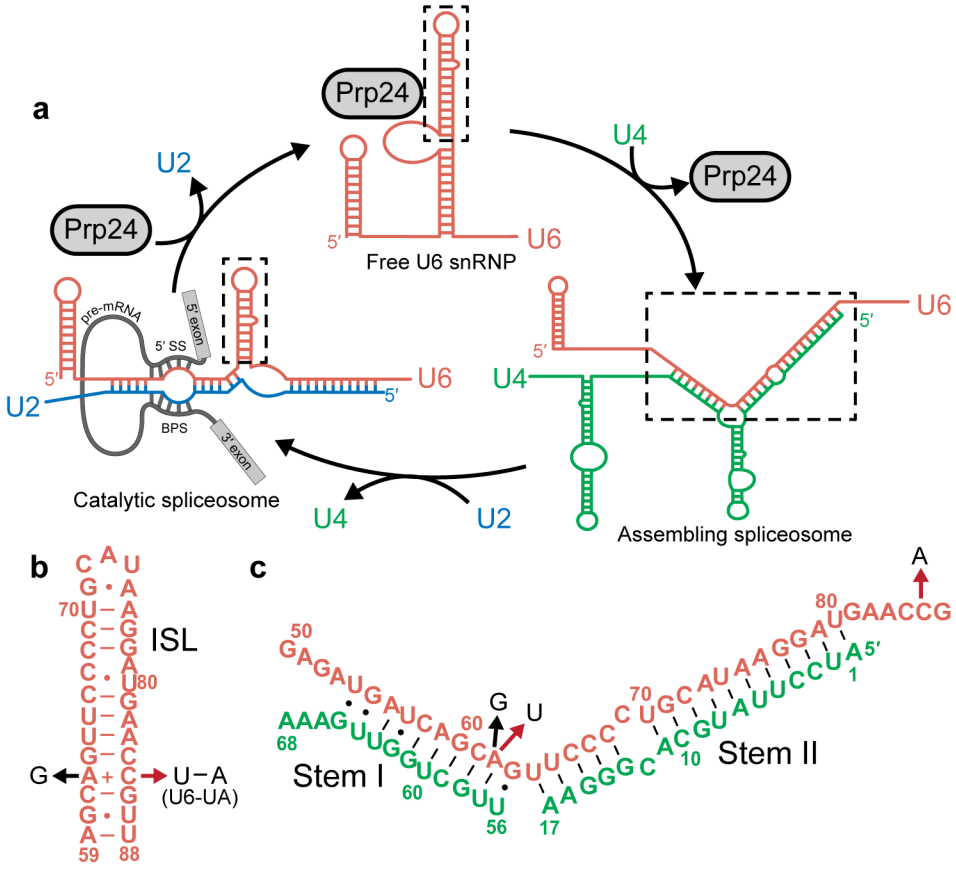

Figure 1.

Conformational changes in U6 snRNA during the splicing cycle. (**a**) Current models of secondary structure in free U6, U4/U6 and U2/U6 snRNAs. A pre-mRNA is shown base-paired to U2/U6. Prp24 is thought to stably bind only free U6 snRNA. Boxes indicate structures shown in detail below. Dashes and circles represent Watson-Crick and non-Watson-Crick base-pairs, respectively, and "+" denotes protonation of A62 in the A62–C85 base-pair. (**b**) Sequence of the yeast U6 ISL present in the U6 snRNP. A slightly truncated version of the U6 ISL is present in U2/U6. Black arrow denotes the A62G mutation. Red arrow denotes the A62U,C85A double mutation in the "U6-UA" mutant. (**c**) The U4/U6 base-paired region is shown in detail. The arrows denote the positions of mutations shown in panel **b**.

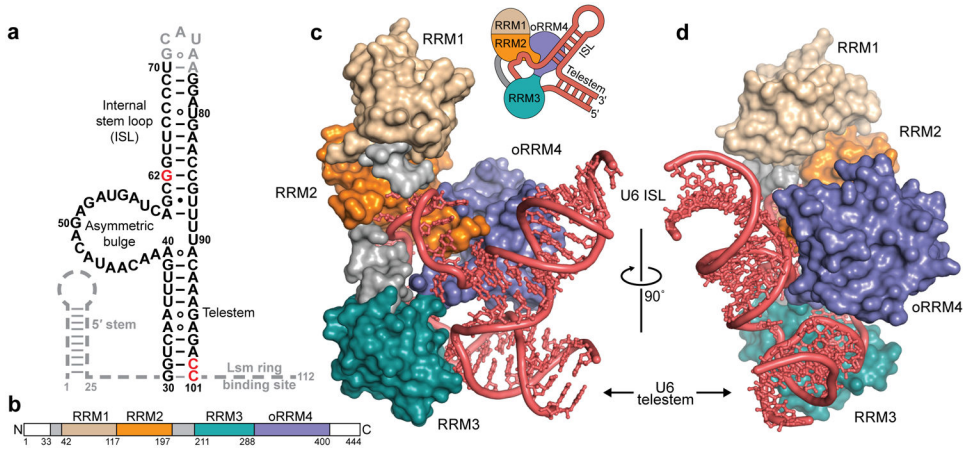

Figure 2.

Structure of the yeast U6–Prp24 complex. (**a**) Secondary structure of *S. cerevisiae* U6(30-101)-A62G,U100C,U101C mutant snRNA bound to Prp24, as observed in the crystal structure. Dashed gray lines indicate regions of the RNA that were deleted to facilitate crystallization. Nucleotides 71–76 and 101 are disordered in the crystal, and red nucleotides are mutated relative to the wild-type U6 sequence. Dashes represent Watson-Crick base-pairing, while open and closed circles denote non-Watson-Crick pairing. (**b**) Domain architecture of the Prp24 protein from *S. cerevisiae*. The first 33 and last 44 amino acids of Prp24 (in white) were deleted from the construct used in crystallization trials. (**c,d**) Two views of the crystal structure of the U6-A62G–Prp24 complex, rotated 90° relative to one another. U6 snRNA is colored salmon and the Prp24 domains are colored as in panel **b**. A cartoon schematic of the entwined protein/RNA topology is shown in (**c**). A stereo view of the structure is presented in Supplementary Figure 1.

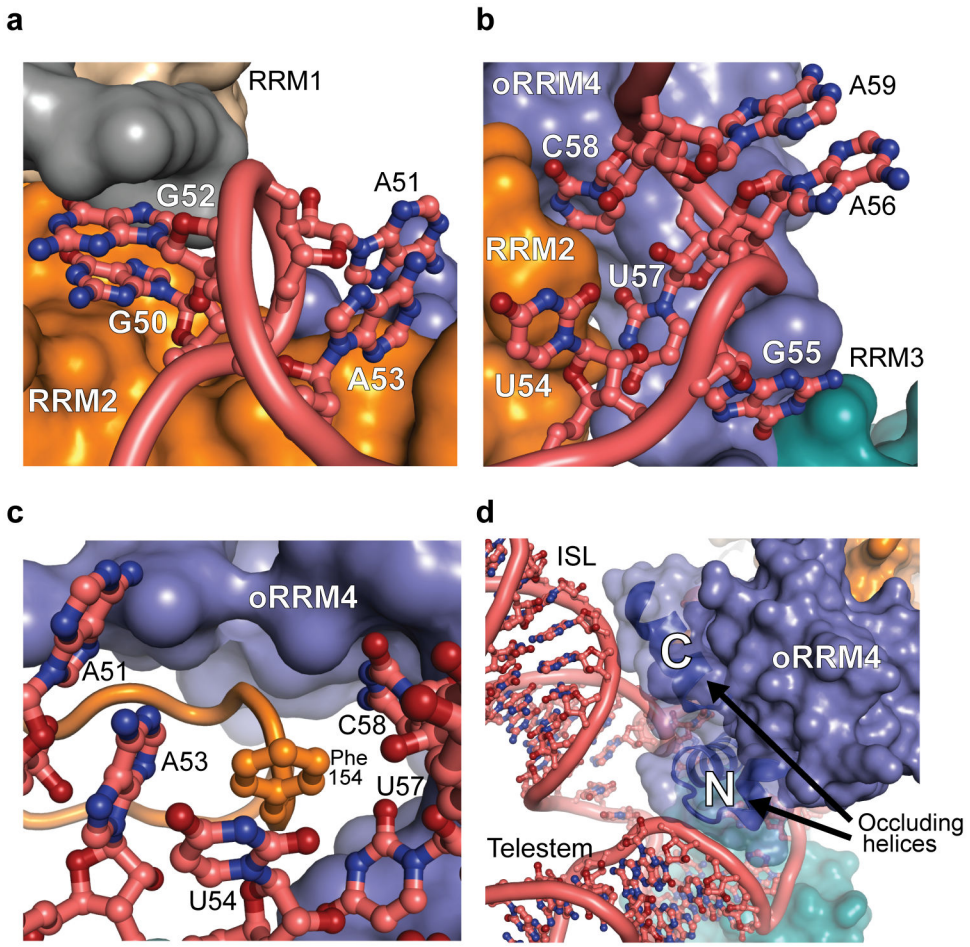

Figure 3.

Novel structural motifs in the U6–Prp24 complex. RNA and protein are colored as in Figure 2. (**a**) A "skip-stack turn" motif stabilized by contacts with residues N-terminal to RRM1 (gray), RRM2 (orange), and oRRM4 (purple). In this motif, U6 nucleotides 50-53 form alternating *i* to *i*+2 and *i*+1 to *i*+3 stacking interactions, resulting in a tight turn. (**b**) A "dinucleotide bulge turn" at the base of the ISL spans U6 nucleotides 56-59. The first and last bases stack on one another, while the central two bases point the opposite direction and are bound by Prp24. The remainder of the ISL and telestem are omitted for clarity. (**c**) A hydrophobic cage of nucleotides surrounds Phe154 in RRM2 and mediates tertiary contacts between RRM2 and oRRM4. (**d**) The occluding alpha-helices (cartoon under semi-transparent surface) at the N- and C-termini of oRRM4 bridge the ISL and telestem.

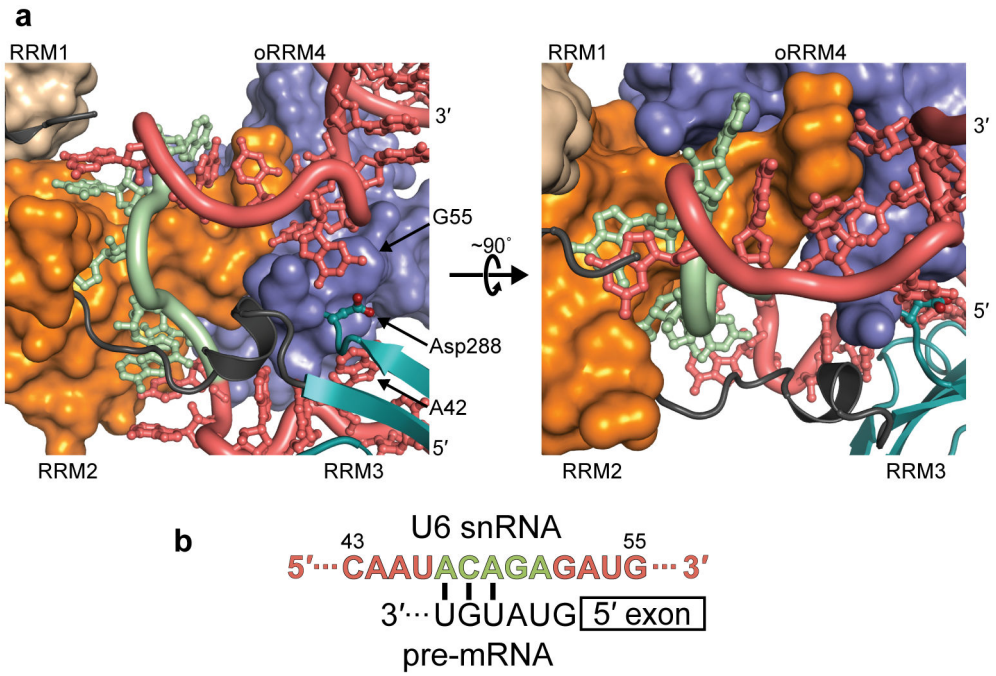

Figure 4.

Prp24 interactions with the U6 asymmetric bulge. (**a**) Two views of the U6-A62G–asymmetric bulge, rotated 90° relative to one another. U6 nucleotides that comprise the ACAGA-box motif are colored green. The remaining U6 nucleotides are colored salmon and Prp24 is colored as in Figure 2. The RRM2-3 linker and RRM3 are shown in cartoon representation to allow visualization of the RNA, including the “aspartate bridge” motif involving Asp288. (**b**) Base-pairing of the U6 snRNA ACAGA-box (nucleotides 47-51) with the intron 5′ splice site in the activated spliceosome[1].

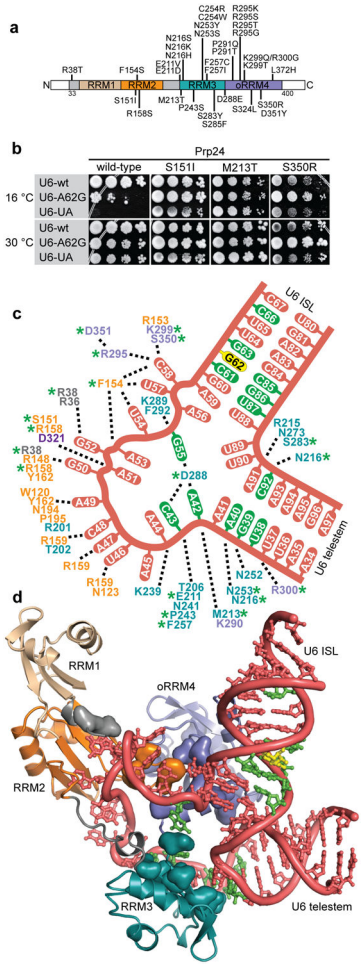

Figure 5.

Suppressors of U6-A62G cold-sensitivity map to the RNA-protein interface. (**a**) Substitutions in Prp24 selected as suppressors of U6-A62G cold-sensitivity. R158S and F257I were identified previously[12]. (**b**) Suppression of the cold-sensitivity of U6-A62G and U6-UA by cloned alleles of *PRP24*. An OD_{600} = 1 culture and three successive 10-fold serial dilutions are shown for each. All strains grow well at 30 °C. (**c**) Schematic of non-covalent interactions between amino acids in Prp24 and nucleotides in U6 snRNA (dotted lines; not intended to differentiate between backbone or side-chain interactions). Positions of suppressor substitutions in Prp24 are denoted by green asterisks, and in U6 by green residues. (**d**) Positions of suppressor substitutions in the U6-A62G–Prp24 complex. All substituted amino acids are shown in surface representation. All substituted nucleotides are shown in green. The U6-A62G mutation is shown in yellow.

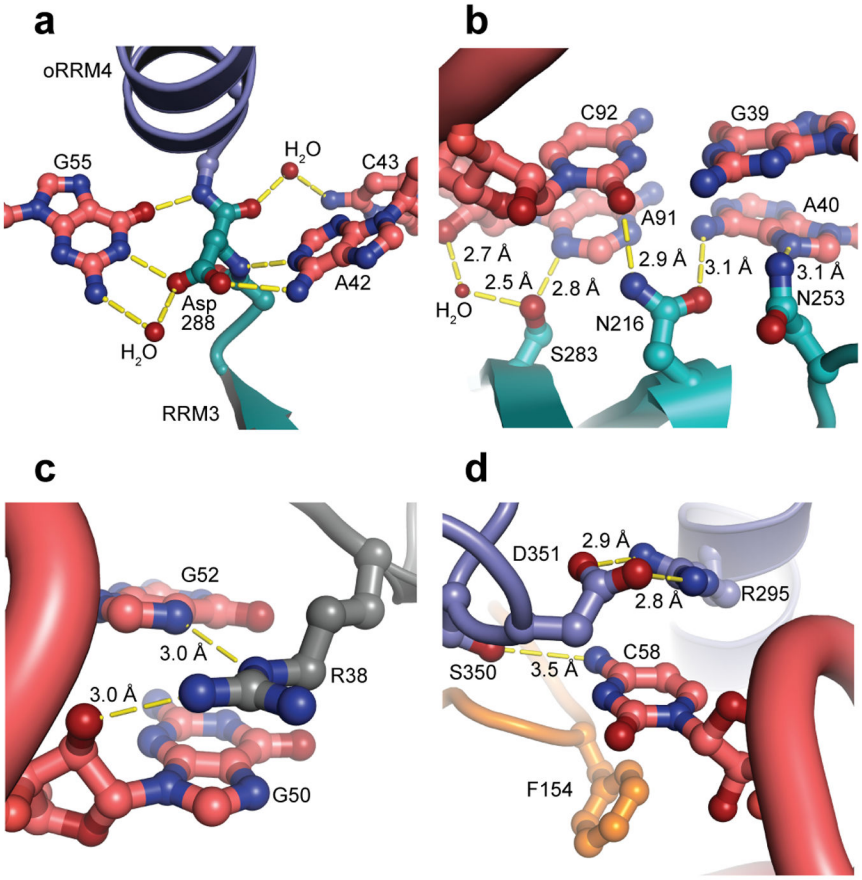

Figure 6.

U6-A62G-suppressor mutations are expected to destabilize the interaction between Prp24 and U6 snRNA. RNA and protein are colored as in Figure 2. (**a**) Hydrogen-bonds between residues comprising an “aspartate bridge” between U6 nucleotides A42 and G55. Asp288, C43, A42 and G55 are all sites of suppressor mutations, which should disrupt the protein-mediated base pair. (**b**) Interactions between RRM3 and the telestem. Substitutions at Asn216, Asn253 and Ser283, and U6 nucleotides G39, A40 and C92 should disrupt this H-bond network. (**c**) Interaction of the Prp24 N-terminal domain with the U6 ACAGA-box. Arg38 interacts with G50 and G52; suppressor mutation R38T should disrupt both these interactions. (**d**) Intersection of RRM2, oRRM4, and the U6 asymmetric loop. Suppressor substitutions at Phe154, Arg295, Ser350, Asp351, and U6 nucleotide C58 should disrupt H-bonds and/or stacking interactions.

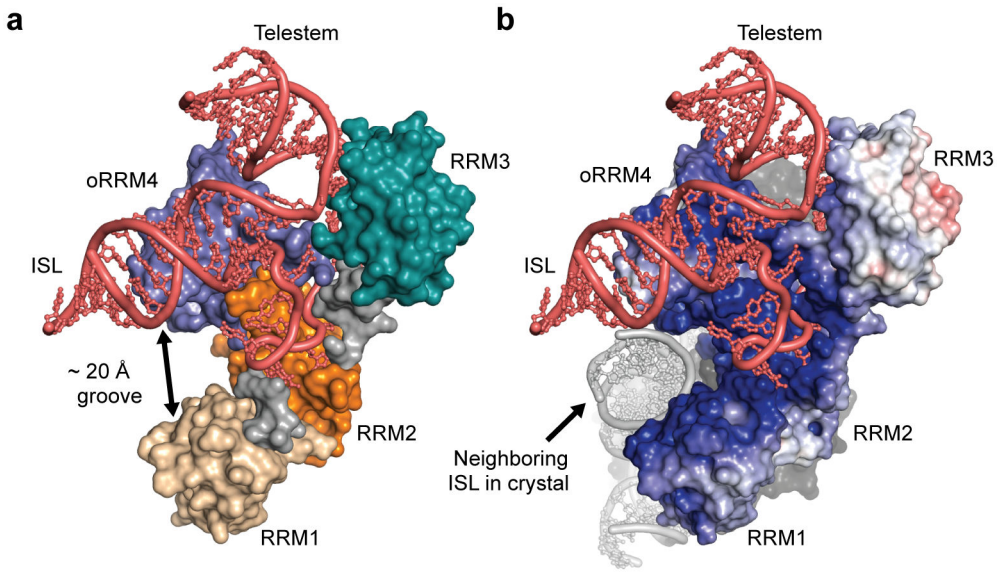

Figure 7.

An electropositive groove in Prp24 binds double-stranded RNA. (**a**) RRMs 1, 2 and oRRM4 delineate a surface groove of approximately 20 Å in width. RNA and protein are colored as in Figure 2. (**b**) In the crystal, the electropositive groove of one U6 snRNP binds the ISL from an adjacent U6 snRNP (shown in gray). The electrostatic surface of Prp24 is contoured from +8 kT/e (blue) to -8 kT/e (red).

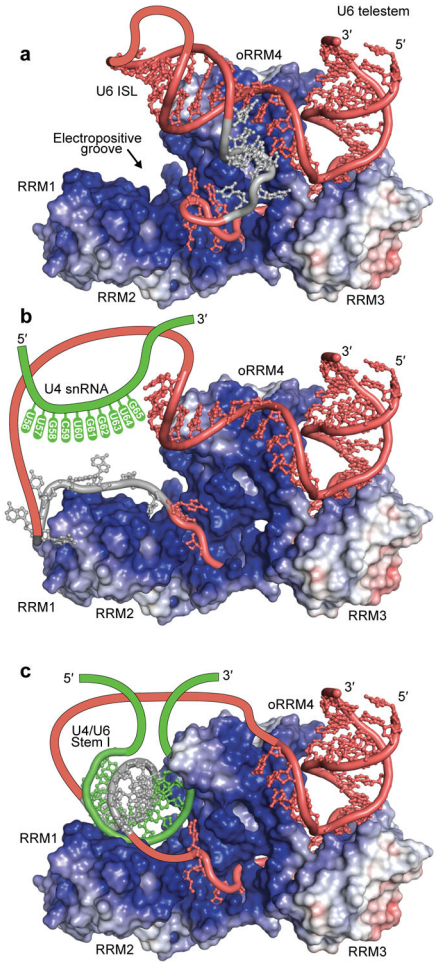

Figure 8.

Proposed mechanism for Prp24-mediated annealing of U4 and U6 snRNAs. (**a**) Ground state structure of the U6–Prp24 complex. The electrostatic surface of Prp24 is contoured from +8 kT/e (blue) to -8 kT/e (red). The portion of U6 that forms U4/U6 Stem I is highlighted in gray. (**b**) An early U4/U6 RNA annealing intermediate. We propose transient unwinding of the U6 ISL allows the groove to bind single-stranded U6 snRNA in an orientation compatible with recognition by U4 snRNA nucleotides 56–65 (in green, see Fig. 1). The conformation of U6 nucleotides 49–60 is taken from the solution structure of RRMs 1 and 2 (PDB: 2GO9) in complex with the same region of RNA[26]. The remainder of non-cartoon RNA is taken from the crystal structure presented here. (**c**) A later U4/U6 RNA annealing intermediate. Stem I of the U4/U6 di-snRNA has formed as is shown in the Prp24 electropositive groove.

Table 1
Data collection and refinement statistics (molecular replacement)

	U6-A62G-Prp24
Data collection	
Space group	$P2_1$
Cell dimensions	
a, *b*, *c* (Å)	61.8, 71.4, 82.1
α, β, γ (°)	90, 109.6, 90
Resolution (Å)	50-1.7(1.73-1.70)*
R_{sym}	0.06(0.90)
$I / \sigma I$	28.3(2.15)
Completeness (%)	99.9(100)
Redundancy	7.6(7.5)
Refinement	
Resolution (Å)	26.3-1.70
No. reflections	74201
R_{work} / R_{free}	18.4(27.0)/21.1(30.1)
No. atoms	5032
Protein	2942
RNA	1389
Ligand/ion	5
Water	696
B factors (Å^2)	49.5
Protein	41.8
RNA	69.7
Ligand/ion	57.1
Water	41.8
r.m.s. deviations	
Bond lengths (Å)	0.006
Bond angles (°)	1.03

All diffraction data were obtained from a single crystal.

*Values in parentheses are for highest-resolution shell.