

Secondary structure of U6 small nuclear RNA: implications for spliceosome assembly

Elizabeth A. Dunn and Stephen D. Rader¹

Department of Chemistry, University of Northern British Columbia, Prince George, British Columbia, Canada V2N 4Z9

Abstract

U6 snRNA (small nuclear RNA), one of five RNA molecules that are required for the essential process of pre-mRNA splicing, is notable for its high level of sequence conservation and the important role it is thought to play in the splicing reaction. Nevertheless, the secondary structure of U6 in the free snRNP (small nuclear ribonucleoprotein) form has remained elusive, with predictions changing substantially over the years. In the present review we discuss the evidence for existing models and critically evaluate a fundamental assumption of these models, namely whether the important 3' ISL (3' internal stem-loop) is present in the free U6 particle, as well as in the active splicing complex. We compare existing models of free U6 with a newly proposed model lacking the 3' ISL and evaluate the implications of the new model for the structure and function of U6's base-pairing partner U4 snRNA. Intriguingly, the new model predicts a role for U4 that was unanticipated previously, namely as an activator of U6 for assembly into the splicing machinery.

U6 snRNA (small nuclear RNA) and pre-mRNA splicing

Nuclear pre-mRNA splicing involves the removal of pre-mRNA regions (introns) that do not code for functional molecules and the subsequent joining of the coding regions (exons) to produce a mature mRNA transcript. Splicing is complex, requiring dynamic interactions between five snRNAs, U1, U2, U4, U5 and U6, and over 100 proteins. A number of these proteins associate with a specific snRNA in an snRNP (small nuclear ribonucleoprotein) particle. These particles are thought to assemble on each new pre-mRNA transcript in an ordered and step-wise fashion to generate a large RNA-protein complex, known as the spliceosome, which is responsible for catalysing splicing reactions.

Mechanistic and structural similarities to group II self-splicing introns have led to the proposal that, like self-splicing introns, pre-mRNA splicing may be an RNA-catalysed event [1,2]. Of the five snRNAs, U6 is the most likely to play a direct role in catalysis since it associates with the 5' splice site of the pre-mRNA in the active core of functional spliceosomes [3] (Figure 1). Furthermore, the highly conserved AGC triad has been shown to base-pair to U2 in a structure that has been implicated in exon ligation [4], and the adjacent U6 3' ISL (3' internal stem-loop), which bears structural resemblance to domain V of group II introns [1,2,5], co-ordinates a catalytically essential Mg²⁺ during the splicing reactions [6]. Lastly, U6 is unusually highly conserved, both in size and sequence, with approx. 60% sequence identity between yeast and human U6 snRNAs, increasing to 80% identity across the middle third of these sequences [7,8].

Despite the high level of sequence conservation, many mutations in U6 snRNA result in only weak conditional growth phenotypes, or show no observable growth phenotype at all. One stretch of nucleotides, however, known as the ACAGAGA box, is unexpectedly intolerant of mutation [9–11]. Point mutations within this stretch of seven nucleotides result in lethality *in vivo* and severe inhibition of splicing *in vitro* with various levels of first- and/or second-step blocks [9,10]. The ACAGAGA sequence has been shown to genetically interact with, as well as to cross-link to, the 5' splice site region of the pre-mRNA and has also been proposed to play an as-yet-undefined role throughout the splicing reaction [9,12–15].

In order for U6 snRNA to enter into functional spliceosomes, it must first interact with U4 snRNA through an extensive base-pair interaction in a di-snRNP (Figure 1). The functional relevance of this particle, or the large conformational rearrangements required in order to accommodate its formation, has yet to be revealed. The role of U4 in splicing is even less clear as U4 snRNA is thought to leave the spliceosome following formation of the di-snRNP, but prior to any catalytic event [16,17]. It is clear, however, that the U4/U6 complex can form either via a biogenesis pathway, from newly assembled U4 and U6 snRNPs, or via a 'recycling' pathway, using free U4 and U6 snRNPs that have been disassembled from previous rounds of splicing [18,19]. Both of these pathways are thought to require the splicing factor Prp24 to catalyse U4/U6 formation [20].

Models of the U6 secondary structure in free U6 snRNP

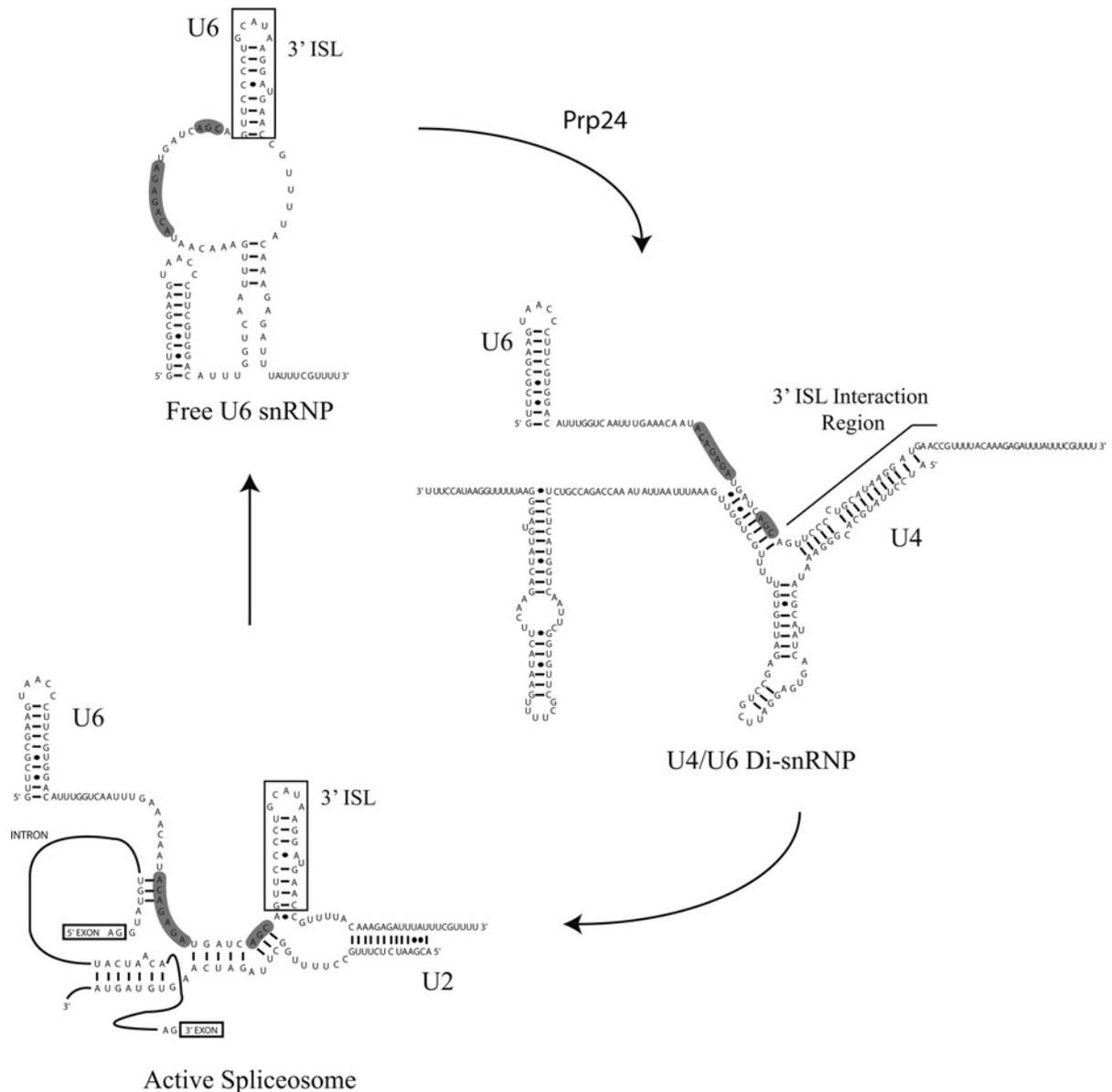
Over the last 15 years, a number of models for the yeast U6 snRNA secondary structure in free U6 snRNP have been proposed, yet they offer little insight into the structural rearrangements leading to the U4/U6 di-snRNP formation,

Key words: pre-mRNA splicing, Prp24, small nuclear ribonucleoprotein biogenesis (snRNP biogenesis), spliceosome assembly, U4 small nuclear RNA, U6 small nuclear ribonucleoprotein.
Abbreviations used: 3' ISL, 3' internal stem-loop; snRNA, small nuclear RNA; snRNP, small nuclear ribonucleoprotein.

¹To whom correspondence should be addressed (email rader@unbc.ca).

Figure 1 | U6 snRNA undergoes a number of structural rearrangements during the splicing cycle

Proposed secondary structure in: free U6 snRNP [23]; the U4/U6 di-snRNP interaction [7]; and in the active spliceosome [1,3,12–15,41,42]. The nearly invariant ACAGAGA and AGC sequences are shaded grey and the catalytically important 3' ISL is boxed or overlined.



or into the transition of U6 from an inactive species to an active catalytic element at the heart of the functional spliceosome. Common to each of these models is the catalytically important 3' ISL and an adjacent region of more ambiguous structure. In an early model, the adjacent region was proposed to form a third stem structure that later evolved into the 'telestem' [21,22] (Figure 2). In the most recent model, the upper telestem interaction has been disrupted, generating a large asymmetric bulge consisting of 23 nucleotides on the 5' side and seven nucleotides on the 3' side [23] (Figure 2). It is unusual to find such large unstructured

segments in RNA molecules; however, Karaduman et al. [23] argue that this large bulge accommodates binding of the U6-snRNP-specific protein Prp24, since, in the absence of protein, chemical modification experiments have suggested a much more structured RNA molecule.

Although the 3' ISL was first proposed in the mammalian free U6 snRNP in 1980 [24], the yeast U6 3' ISL was first described in the splicing literature in 1992 in the context of the active spliceosome [1]. Two years later, this structure was proposed for the yeast free U6 snRNP [21] and alternative models for free U6 base-pairing interactions that are also

mechanism where bond breaking in the free U4 and U6 intramolecular stems provides the energy required for formation of the intermolecular interactions.

Chemical structure probing of the free U6 snRNP

Investigations of the secondary structure of U6 snRNA by analysis of purified free U6 snRNP particles is limited to two chemical modification studies [23,33]. Notably, neither study demonstrated that their snRNPs were functional. These experiments were carried out on U6 snRNP species isolated and probed under very different conditions: Jandrositz and Guthrie [33] used native free U6 snRNPs that were purified from other U6-containing species on a glycerol gradient, whereas Karaduman et al. [23] introduced a TAP (tandem affinity purification)-tag at the C-terminal end of Prp24 to allow further purification of U6. The highly conserved C-terminal domain of Prp24 has been shown to interact with the Lsm complex [34], and it is not clear whether introducing a large tag at this location has altered the snRNP structure. Consequently, in comparing the two data sets to assess the validity of the models in Figure 2, we only consider those strong protections and strong modifications that were consistent between the two studies, i.e. we explicitly ignore any modifications or protections that were not seen in both studies. This allows us to be confident that the protection pattern observed was not an artefact of the experimental conditions used. We evaluate our new model alongside the two most recently proposed models from the literature, Vidaver et al. [22] and Karaduman et al. [23].

When only the consistent protections and modifications were considered, there were 21 data points that map to the region of U6 modelled with different secondary structures (nucleotides 29–102) (Figure 2). Three of these positions, Ala⁴⁰–Ala⁴², are strongly modified in the naked RNA controls in both studies, and are strongly protected in the presence of protein, consistent with them serving as a single-stranded protein-binding site, as proposed on the basis of genetic observations [22,35]. This is consistent with the Karaduman et al. [23] and Dunn–Rader models, but not the Vidaver et al. [22] model. Of the remaining 18 data points, 14 fit the Vidaver et al. [22] and Karaduman et al. [23] models, whereas 16 fit the Dunn–Rader model. All eight strongly modified positions are predicted not to base-pair in both the Karaduman et al. [23] and Dunn–Rader models, whereas one of these nucleotides is predicted to be base-paired in the Vidaver et al. [22] model. Of the ten strongly protected positions, seven, six and eight are predicted to be base-paired in the Vidaver et al. [22], Karaduman et al. [23] and Dunn–Rader models respectively. Crucially, two reproducible protections at positions Ala⁴⁹ and Gly⁵⁰ that are inconsistent with both the Vidaver et al. [22] and Karaduman et al. [23] models are explained by stem-loop A in the Dunn–Rader model (Figure 2). The two positions that are inconsistent with the Dunn–Rader model, Gly⁵⁵ and Gly⁶⁰, are also inconsistent with the Karaduman et al. [23] model. These nucleotides are predicted

to lie in the stem-loop A bulge and loop respectively in the Dunn–Rader model, and in the large bulge in the Karaduman et al. [23] model (Figure 2). UV cross-linking of Gly⁵⁵ to Prp24, and protection of these positions from hydroxyl radical cleavage, suggest that these nucleotides may be protected from modification through RNA–protein or tertiary RNA–RNA interactions rather than through base-pairing [23].

A model for allosteric activation of U6 snRNA through interaction with U4 snRNA

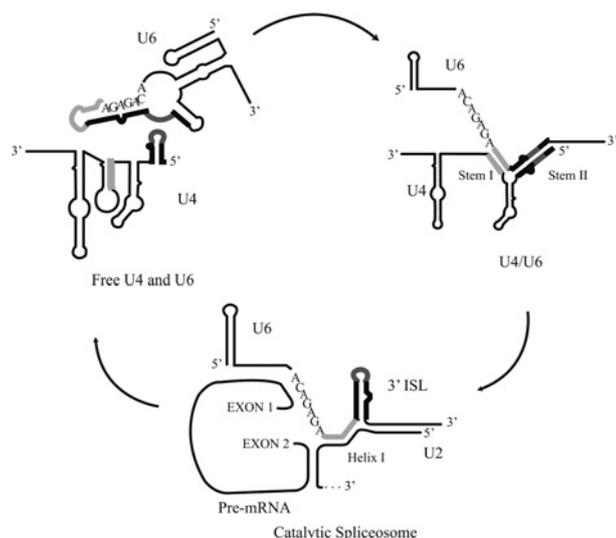
U4 snRNA is an essential gene product; however, its function has remained ambiguous since it is not tightly associated with the active spliceosome and may even dissociate from it following U4/U6 unwinding [16,17]. Brow and Guthrie [36] proposed that U4 might act as an antisense negative regulator of U6 activity, masking catalytically important residues while U6 adopts a conformation that is more favourable for its incorporation into the spliceosome. Alternatively, Stevens et al. [37] suggest that the sub-stoichiometric levels of U4 snRNP, relative to other splicing snRNPs, imply that it is a limiting factor in spliceosome assembly. Thus interaction with U4 might serve as a regulatory point for inclusion of U6 into functional spliceosomes. Although these proposals address the function of U4, they do not explain the large structural rearrangements required in U6 to accommodate the U4/U6 interaction, particularly as the 3' ISL of U6 has been proposed to be nearly identical in the free snRNP and the active spliceosome.

Our model of the U6 snRNA secondary structure is consistent with these proposed roles for U4 snRNA/RNP in splicing; however, we suggest that U4 snRNA performs an additional role as an activator of U6 prior to its role in the di-snRNP (Figure 3). All other proposed yeast and mammalian models of the U6 secondary structure in the free U6 snRNP predict that the U4/U6 interaction domain base-pairs with itself in the 3' ISL. Our model proposes instead that this region of U6 base-pairs with regions of U6 that flank the U4/U6 interaction domain. Importantly, the ACAGAGA sequence is inaccessible in our model due to base-pairing with the U4/U6 interaction domain (Figure 3). We propose that base-pair formation between U4 and U6 is required to release the ACAGAGA sequence from the intramolecular interaction so that it can subsequently interact with the 5' splice site of a pre-mRNA transcript. Hence U4 snRNA can be viewed as an allosteric activator of U6.

The mechanism of U4/U6 di-snRNP formation is not yet understood; however, the U6-snRNP-specific protein Prp24 catalyses base-pairing of U4 and U6 snRNAs [20,34,38,39]. Chemical modification/interference experiments on mammalian U6 snRNA suggest that nucleotides 65–70 are important for initiating base-pairing between U4 and U6 [27]. In yeast, the strong chemical modification of the corresponding residues 72–75, located in the single-stranded junction connecting stem-loops A and B in our model, indicate that these residues are indeed accessible in both the naked RNA and in the presence of protein in free U6 snRNP

Figure 3 | A model for allosteric activation of U6 snRNA by U4 snRNA

Mutually exclusive interaction regions in U4 and U6 are shown in light grey, dark grey and black.



[23,33]. Thus Prp24 probably interacts with U6 snRNA in such a way as to present these nucleotides to U4. Base-pairing between U4 and U6 may be initiated by contacts between the loop residues of a short stem-loop, the U4 kissing loop, located at the 5' end of U4 snRNA (T. Wong and S. Rader, unpublished work) and U6 residues 72–75 (Figure 3).

Following initiation of base-pairing through the loop-bulge interaction, extension of the intermolecular interaction would result in the complete disruption of stem-loops A and B of U6, releasing the ACAGAGA sequence from intramolecular base-pairing. According to our model, the only U6 present in the cell with an exposed ACAGAGA sequence would be found in the di-snRNP. Simultaneously, the U2/U6 helix I and the U6 3' ISL regions of U6 would be engaged in interactions with U4, preventing premature formation of these functionally important structures (Figure 3). The ACAGAGA sequence is predicted to lie outside of the U4/U6 interaction domain, and strong chemical modification of this sequence in U4/U6 di-snRNPs suggests that the ACAGAGA sequence is in fact accessible in this context [7,33]. Cross-links observed between U4 snRNA and the 5' splice site in early stages of spliceosome formation imply that a proofreading mechanism may exist to ensure that correct base-pairing between U6 and the 5' splice site is established prior to U4/U6 disruption [40]. Dissociation of U4 following establishment of the correct U6/pre-mRNA interaction would then allow the formation of U2/U6 helix I and the U6 3' ISL within the catalytically active spliceosome (Figure 3).

Summary

We have reviewed existing models of U6 snRNA in the free snRNP particle in the light of some of the existing experimental results and functional considerations.

Importantly, whereas previous models assumed the existence of the stable 3' ISL, our model proposes instead two weaker stems, consistent with the low-activation-energy observed for U4/U6 di-snRNP formation. In replacing the 3' ISL with two smaller stems, our model removes the structurally unprecedented asymmetric bulge in favour of the well-known structural motif of a three-way helical junction, which results in a modestly better fit to the most consistent chemical modification data. Finally, our model sequesters the functionally critical ACAGAGA sequence, suggesting that an important role for the U4 snRNP is to open up U6 to activate it for spliceosome assembly.

Funding

Work in our laboratory was supported by a Health Research Project Award [grant numbers 31426-1 and 31426-2 (to E.A.D.)]; and by a National Science and Engineering Research Council of Canada Discovery Grant [grant number 298521 (to S.D.R.)]. S.D.R. is a Cottrell College Scholar of the Research Corporation.

References

- Madhani, H.D. and Guthrie, C. (1992) A novel base-pairing interaction between U2 and U6 snRNAs suggests a mechanism for the catalytic activation of the spliceosome. *Cell* **71**, 803–817
- Peebles, C., Zhang, M., Perlman, P. and Fanzen, J. (1995) Catalytically critical nucleotides in domain 5 of a group II intron. *Proc. Natl. Acad. Sci. U.S.A.* **92**, 4422–4426
- Kim, C.H. and Abelson, J. (1996) Site-specific crosslinks of yeast U6 snRNA to the pre-mRNA near the 5' splice site. *RNA* **2**, 995–1010
- Hilliker, A.K. and Staley, J.P. (2004) Multiple functions for the invariant triad of U6 snRNA. *RNA* **10**, 921–928
- Sashital, D.G., Cornilescu, G. and Butcher, S.E. (2004) U2–U6 RNA folding reveals a group II intron-like domain and a four-helix junction. *Nat. Struct. Mol. Biol.* **11**, 1237–1242
- Yean, S.-L., Wuenschell, G., Termini, J. and Lin, R.-J. (2000) Metal-ion coordination by U6 small nuclear RNA contributes to catalysis in the spliceosome. *Nature* **408**, 881–884
- Brow, D.A. and Guthrie, C. (1988) Spliceosomal U6 is remarkably conserved from yeast to mammals. *Nature* **334**, 213–218
- Guthrie, C. and Patterson, B. (1988) Spliceosomal snRNAs. *Annu. Rev. Genet.* **22**, 387–419
- Fabrizio, P. and Abelson, J. (1990) Two domains of yeast U6 small nuclear RNA required for both steps of nuclear precursor messenger RNA splicing. *Science* **250**, 404–409
- Madhani, H.D., Bordonne, R. and Guthrie, C. (1990) Multiple roles for U6 snRNA in the splicing pathway. *Genes Dev.* **4**, 2264–2277
- Wolff, T., Messen, R., Hammel, J. and Bindereif, A. (1994) Splicing function of mammalian U6 small nuclear RNA: conserved positions in central domain and helix I are essential during the first and second step of pre-mRNA splicing. *Proc. Natl. Acad. Sci. U.S.A.* **91**, 903–907
- Sawa, H. and Abelson, J. (1992) Evidence for a base-pairing interaction between U6 small nuclear RNA and the 5' splice site during the splicing reaction in yeast. *Proc. Natl. Acad. Sci. U.S.A.* **89**, 11269–11273
- Wasserman, D.A. and Steitz, J.A. (1992) Interaction of small nuclear RNAs with precursor messenger RNA during *in vitro* splicing. *Science* **257**, 1918–1925
- Kandels-Lewis, S. and Séraphin, B. (1993) Roles of U6 snRNA in 5' splice site selection. *Science* **262**, 2035–2039
- Lesser, C.F. and Guthrie, C. (1993) Mutations in U6 snRNA that alter splice site specificity: implications for the active site. *Science* **262**, 1982–1988
- Siliciano, P.G., Brow, D.A., Roiha, H. and Guthrie, C. (1987) An essential snRNA from *S. cerevisiae* has properties predicted for U4, including interaction with a U6-like snRNA. *Cell* **50**, 585–592
- Yean, S.-L. and Lin, R.-J. (1991) U4 small nuclear RNA dissociates from yeast spliceosomes and does not participate in the subsequent splicing reaction. *Mol. Cell. Biol.* **11**, 5571–5577

- 18 Neuman de Vegvar, H.E. and Dahlberg, J.E. (1990) Nucleocytoplasmic transport and processing of small nuclear RNA precursors. *Mol. Cell. Biol.* **10**, 3365–3375
- 19 Stanek, D., Rader, S.D., Klingauf, M. and Neugebauer, K.M. (2003) Targeting of U4/U6 small nuclear RNP assembly factor SART3/p110 to Cajal bodies. *J. Cell Biol.* **160**, 505–516
- 20 Raghunathan, P.L. and Guthrie, C. (1998) A spliceosomal recycling factor that reanneals U4 and U6 small nuclear ribonucleoprotein particles. *Science* **279**, 857–860
- 21 Fortner, D.M., Troy, R.G. and Brow, D.A. (1994) A stem/loop in U6 RNA defines a conformational switch required for pre-mRNA splicing. *Genes Dev.* **8**, 221–233
- 22 Vidaver, R.M., Fortner, D.M., Loos-Austin, L.S. and Brow, D.A. (1999) Multiple functions of *Saccharomyces cerevisiae* splicing protein Prp24 in U6 RNA structural rearrangements. *Genetics* **153**, 1205–1218
- 23 Karaduman, R., Fabrizio, P., Hartmuth, K., Urlaub, H. and Lührmann, R. (2006) RNA structure and RNA-protein interactions in purified yeast U6 snRNPs. *J. Mol. Biol.* **356**, 1248–1262
- 24 Epstein, P., Reddy, R., Henning, D. and Busch, H. (1980) The nucleotide sequence of nuclear U6 (4.7S) RNA. *J. Biol. Chem.* **255**, 8901–8906
- 25 Owczarzy, R., Tataurov, A.V., Wu, Y., Manthey, J.A., McQuisten, K.A., Almabrazi, H.G., Pedersen, K.F., Lin, Y., Garretson, J., McEntaggart, N.O. et al. (2008) IDT SciTools: a suite for analysis and design of nucleic acid oligomers. *Nucleic Acids Res.* **36**, W163–W169
- 26 Raghunathan, P.L. and Guthrie, C. (1998) RNA unwinding in U4/U6 snRNPs requires ATP hydrolysis and the DEIH-box splicing factor Brr2. *Curr. Biol.* **8**, 847–855
- 27 Wolff, T. and Bindereif, A. (1993) Conformational changes of U6 RNA during the spliceosome cycle: an intramolecular helix is essential both for initiating the U4–U6 interaction and for the first step of splicing. *Genes Dev.* **7**, 1377–1380
- 28 Brow, D.A. and Vidaver, R.M. (1995) An element in human U6 RNA destabilizes the U4/U6 spliceosomal RNA complex. *RNA* **1**, 122–131
- 29 Achsel, T., Brahm, H., Kastner, B., Bachi, A., Wilm, M. and Lührmann, R. (1999) A doughnut-shaped heteromer of human Sm-like proteins binds to the 3'-end of U6 snRNA, thereby facilitating U4/U6 duplex formation *in vitro*. *EMBO J.* **18**, 5789–5802
- 30 Huppler, A., Nikstad, L.J., Allmann, A.M., Brow, D.A. and Butcher, S.E. (2002) Metal binding and base ionization in the U6 RNA intramolecular stem-loop structure. *Nature* **9**, 431–435
- 31 Reiter, N.J., Nikstad, L.J., Allmann, A.M., Johnson, R.J. and Butcher, S.E. (2003) Structure of the U6 RNA intermolecular stem-loop harboring an SP-phosphorothioate modification. *RNA* **9**, 533–542
- 32 Reference deleted
- 33 Jandrositz, A. and Guthrie, C. (1995) Evidence for a Prp24-binding site in U6 snRNA and in a putative intermediate in the annealing of U6 and U4 snRNAs. *EMBO J.* **14**, 820–832
- 34 Rader, S.D. and Guthrie, C. (2002) A conserved Lsm-interaction motif in Prp24 required for efficient U4/U6 di-snRNP formation. *RNA* **8**, 1378–1392
- 35 Ryan, D.E., Stevens, S.W. and Abelson, J. (2002) The 5' and 3' domains of yeast U6 snRNA: Lsm proteins facilitate binding of Prp24 protein to the U6 telomere region. *RNA* **8**, 1011–1033
- 36 Brow, D.A. and Guthrie, C. (1989) Splicing a spliceosomal RNA. *Nature* **337**, 14–15
- 37 Stevens, S.W., Barta, I., Ge, H.Y., Moore, R.E., Young, M.K., Lee, T.D. and Abelson, J. (2001) Biochemical and genetic analyses of the U5, U6 and U4/U6 · U5 small nuclear ribonucleoproteins from *Saccharomyces cerevisiae*. *RNA* **7**, 1543–1553
- 38 Shannon, K.W. and Guthrie, C. (1991) Suppressors of a U4 snRNA mutation define a novel U6 snRNP protein with RNA-binding motifs. *Genes Dev.* **5**, 773–785
- 39 Ghetti, A., Company, M. and Abelson, J. (1995) Specificity of Prp24 binding to RNA: a role for Prp24 in the dynamic interaction of U4 and U6 snRNAs. *RNA* **1**, 132–145
- 40 Johnson, T.L. and Abelson, J. (2001) Characterization of U4 and U6 interactions with the 5' splice site using a *S. cerevisiae in vitro* trans-splicing system. *Genes Dev.* **15**, 1957–1970
- 41 Field, D.J. and Freisen, J.D. (1996) Functionally redundant interactions between U2 and U6 spliceosomal RNAs. *Genes Dev.* **10**, 489–501
- 42 Parker, R., Siliciano, P.G. and Guthrie, C. (1987) Recognition of the TACTAAC box during mRNA splicing in yeast involves base pairing to the U2-like snRNA. *Cell* **49**, 229–239

Received 8 February 2010
doi:10.1042/BST0381099