Roles of the U5 snRNP in spliceosome dynamics and catalysis


MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, U.K.

Abstract

Most protein-coding genes in eukaryotes are interrupted by non-coding intervening sequences (introns), which must be precisely removed from primary gene transcripts (pre-mRNAs) before translation of the message into protein. Intron removal by pre-mRNA splicing occurs in the nucleus and is catalysed by complex ribonucleoprotein machines called spliceosomes. These molecular machines consist of several small nuclear RNA molecules and their associated proteins [together termed snRNP (small nuclear ribonucleoprotein) particles], plus multiple accessory factors. Of particular interest are the U2, U5 and U6 snRNPs, which play crucial roles in the catalytic steps of splicing. In the present review, we summarise our current understanding of the role played by the protein components of the U5 snRNP in pre-mRNA splicing, which include some of the largest and most highly conserved nuclear proteins.

Introduction

Nascent pre-mRNAs are packaged into spliceosomes in the nucleus by recruitment of several trans-acting RNA-protein complexes containing the U1, U2, U4, U5 and U6 snRNAs (small nuclear RNAs) and multiple protein factors. Additional protein splicing factors associate with precatalytic spliceosomes, which are extensively remodelled before the catalytic steps of splicing. Intron excision and exon ligation occur through two consecutive transesterification reactions, involving a branched or ‘lariat’ intermediate and product (Figure 1). Accurate recognition of the intron splice sites and branch site is achieved in part by sequential interactions with multiple spliceosome components. For example, the 5′-splice site is recognized initially by base-pairing with the U1 snRNA. This interaction is supplanted by base-pairing with the U6 snRNA as the spliceosome is remodelled for catalysis of the first transesterification.

Energetics of pre-mRNA splicing

Although the splicing reactions themselves do not directly require energy input, remodelling of a network of RNA–RNA and RNA–protein interactions in the spliceosome consumes multiple ATP molecules. These rearrangements are catalysed by several ATP-dependent RNA helicases and may also involve the activity of a GTPase related to the translation elongation factor EF-2. In fact, various motors, clocks and springs are required at multiple stages of spliceosome assembly, catalysis of the splicing reactions and release of the mRNA and ‘lariat’ intron products (reviewed in [1]). The largest class of energy-requiring proteins involved in splicing belongs to the DEXD/H box family: these proteins are commonly known as RNA helicases but in some cases may act as RNPases (ribonucleoproteinases), which disrupt RNA–protein complexes in the spliceosome [2].

Spliceosome remodelling factors in the U5 snRNP (small nuclear ribonucleoprotein)

Of the six U5 snRNP-specific proteins three are NTPases. Prp28p and Brr2p are members of the DEXD/H box family, whereas Snu114p is the sole GTPase identified in the spliceosome to date and is related to translation elongation factor EF-2 [3,4]. The NTPases of the U5 snRNP are involved in the critical switch in which U1 is replaced by U6 at the 5′-splice site (Figure 2). This is an important stage in spliceosome activation, which contributes to the fidelity of 5′-splice-site recognition. The two unwinding events that disrupt U1:5′-splice site base-pairing and U4:U6 base-pairing allow the 5′-splice site, U6 and U2 catalytic core structure to form [5] discussed below. It has been shown that Prp28p has a role in destabilizing the U1 snRNA interaction with the 5′-splice site. This DEXD/H box protein may unwind the helix formed between the 5′-splice site and U1 snRNA [6]. Alternatively, it is a prime candidate to act as an RNPase. Under normal conditions, Prp28p is an essential protein in yeast but if the U1C protein (a factor that stabilizes the U1:5′-splice site interaction) is mutated, Prp28p then becomes dispensable [7]. This implicates Prp28p in disrupting the U1C interaction, either directly acting against the protein or by disrupting the U1 snRNA:5′-splice site helix that forms its site of interaction.

Disruption of the U1:5′-splice site interaction is accompanied by the release of U6 from U4 (Figure 2). The U4/U6.U5 tri-snRNP contains the base-paired U4:U6 structure that prevents the formation of catalytic core structures until the correct time. In vitro work has shown that the human homologue of Brr2p is capable of unwinding RNA helices including a base-paired U4:U6 complex [8]. Furthermore,
Genes: Regulation, Processing and Interference

Figure 1 | Spliceosome dynamics and catalysis

(A) Conserved sequences in a typical yeast pre-mRNA. The intron consensus sequences direct spliceosome assembly and define the splice sites and branch site (Y denotes pyrimidine). The degree of sequence conservation is indicated by the size of the letters. (B) Two-step transesterification pathway of pre-mRNA splicing. In the first step, the 5′-splice site is attacked by the 2′-hydroxyl of the branch site adenosine, generating the Exon 1 intermediate and a branched intron ‘lariat’ still attached to Exon 2. In the second step, the 3′-splice site is attacked by the 3′-hydroxyl of the free Exon 1 intermediate. The final products are the spliced mRNA and the excised intron in a branched or ‘lariat’ form.

Figure 2 | Model of the 5′-splice site switch

During spliceosome assembly, the 5′-splice site is initially recognized by the U1 snRNP. This interaction is later replaced by U6, which must base-pair with the 5′-splice site before the first catalytic step. The unwinding of U4/U6 is also required for the spliceosome to become active. These events occur concurrently and allow the release of U1 and U4 snRNPs from the spliceosome. The roles that the U5 snRNP components, Prp28p, Brr2p and Snu114p, are hypothesized to perform during activation of the spliceosome are indicated: Prp28p displaces U1 (and the U1C protein, see text) from the 5′-splice site presumably utilizing ATP, whereas Brr2p and Snu114p are involved in the unwinding of U4 and U6 snRNAs, a process requiring ATP and GTP. This results in free U6 paired with the 5′-splice site. U6 is also capable of forming an intramolecular stem loop, a structure important for catalysis, after release from U4. The 5′-splice site consensus from Saccharomyces cerevisiae is used with base-pairing to U1 and U6 snRNAs indicated.

It has been shown that a mutation in the ATPase domain of Brr2p (brr2-1) inhibits the ATP-dependent disruption of U4/U6. U5 tri-snRNPs in yeast cell extracts [9]. These studies suggest that Brr2p is the motor that drives the release of U6 before formation of the U6:5′-splice site and U6:U2 structures. However, these results and other data show that DEXD/H box proteins have little intrinsic control over unwinding specificity or timing, so it is important to control these factors and prevent untimely activation of splicing complexes. Both Snu114p and Prp8p, a large U5 snRNP protein whose activities are discussed in detail below, are supposed to provide this control [10–12]. Elegant studies using a Snu114p mutant that switched specificity from GTP to XTP allowed dissection of this regulatory role of Snu114p [10]. It was shown that stalled complexes would only unwind the U4:U6 helices when supplied with hydrolysable XTP, implying that Snu114p has a role either in unwinding U4:U6 or more probably in controlling the action of Brr2p.

Molecular basis of catalysis in the spliceosome

The highly dynamic nature of the spliceosome has made it a particular challenge to understand the structural and mechanistic basis of the catalysis of pre-mRNA splicing. Studies performed in the 1990s revealed the formation of a network of crucial RNA–RNA interactions in the spliceosome’s core (reviewed in [13]). This array of contacts involving conserved sequences in the U2, U5 and U6 snRNAs and the pre-mRNA substrate is supposed to underlie recognition and positioning of the splice sites and intron branch site for catalysis (Figure 3). U2 and U6 snRNAs interact with the branch point and 5′-splice site respectively, and multiple base-pairing interactions between U2 and U6 then provide a structural basis for juxtaposing the branch site and 5′-splice site for the first catalytic step. U5 snRNA bears a highly conserved stem loop that is implicated in aligning the exons for the second catalytic step [14].
There are several striking similarities between the spliceosome’s core and a family of self-splicing introns (so-called Group II introns) found in bacteria and in organelle genomes in many lower eukaryotes. Group II introns splice through the same two-step transesterification pathway as the spliceosome, and the reactions proceed with the same stereochemistry. Several important secondary structures in Group II introns have clear counterparts in the spliceosome’s catalytic core. Both Group II introns and the spliceosome splice, and this activity is only observed at non-physiological concentrations, this protein-free RNA complex is unique in making extensive contacts with U5 and U6 snRNAs and with the pre-mRNA substrate at the splice sites and intron branch site [22–25]. Prp8p probably plays a role in stabilizing the interactions between the U5 snRNA loop sequence and the exons [23] and may also be responsible for juxtaposing the U5 loop with the rest of the catalytic core (Figure 3). Mutational studies of Prp8p [12,26–29] also support the view that this protein is intimately involved in the functions of the catalytic core of the spliceosome. Prp8p has been implicated in multiple aspects of spliceosome remodelling and activation in addition to putative cofactor activity in catalysis, including a central role in governing the activities of the Br2p and Prp28p RNA–dependent ATPases [11,12].

Despite its striking phylogenetic conservation, 62% identity between the yeast and human proteins throughout the 2400 amino acid sequence, Prp8p contains little in the way of recognizable sequence motifs, so that its domain structure is unclear and it is difficult to make testable predictions about the biochemical activities of the protein. One clear homology displayed by Prp8p is the presence, near the C-terminus, of a Jab1/MPN domain also found in some de-ubiquitinating enzymes, but the functional significance of this is currently unclear. Notwithstanding its extensive contacts with catalytic core RNAs, Prp8p apparently lacks homologies to classical RNA-binding domains, so it will be of great interest to locate and analyse its RNA-interaction surface(s).

Future prospects

High-resolution structural information on the spliceosome remains a distant prospect due to the highly dynamic and heterogeneous nature of the complexes. Cryo EM structures provide a view of the morphology of snRNP complexes and spliceosomes (see e.g. [30,31]). However, independent means of identifying the positions of splicing factors will be essential. Evidence of interaction networks between splicing factors from yeast two-hybrid analysis provides one means of identifying subcomplexes of functionally related factors and analysing the interactions between them [32,33]. Our understanding of RNA–RNA interactions has been helped greatly by site-specific photocross-linking data [34,35]. This technique has also highlighted important RNA–protein interactions, including the presence of Prp8p at the sites of chemistry in the spliceosome’s catalytic core.

A protein cofactor for the catalytic RNA core of the spliceosome

One particular spliceosomal protein has attracted a great deal of attention recently as the prime candidate for acting as a protein cofactor for RNA-based catalysis in the spliceosome. This is the Prp8 protein, a large, highly conserved component of the U5 snRNP. Prp8p is unique in making extensive contacts with U5 and U6 snRNAs and with the pre-mRNA substrate at the splice sites and intron branch site [22–25]. Prp8p probably plays a role in stabilizing the interactions between the U5 snRNA loop sequence and the exons [23] and may also be responsible for juxtaposing the U5 loop with the rest of the catalytic core (Figure 3). Mutational studies of Prp8p [12,26–29] also support the view that this protein is intimately involved in the functions of the catalytic core of the spliceosome. Prp8p has been implicated in multiple aspects of spliceosome remodelling and activation in addition to putative cofactor activity in catalysis, including a central role in governing the activities of the Br2p and Prp28p RNA–dependent ATPases [11,12].

Despite its striking phylogenetic conservation, 62% identity between the yeast and human proteins throughout the 2400 amino acid sequence, Prp8p contains little in the way of recognizable sequence motifs, so that its domain structure is unclear and it is difficult to make testable predictions about the biochemical activities of the protein. One clear homology displayed by Prp8p is the presence, near the C-terminus, of a Jab1/MPN domain also found in some de-ubiquitinating enzymes, but the functional significance of this is currently unclear. Notwithstanding its extensive contacts with catalytic core RNAs, Prp8p apparently lacks homologies to classical RNA-binding domains, so it will be of great interest to locate and analyse its RNA-interaction surface(s).
RNA–protein cross-linking provides a powerful method of identifying the sites of action of proteins such as Prp28p, Brr2p and Snu114p all of which are expected to contact RNA through their active sites. In the human in vitro system, Prp28p has already been shown to cross-link to the 5′-splice site consistent with its demonstrated role in the switch between U1 and U6 [36]. For Prp28p, the site of substrate RNA cross-linking was mapped to a small peptide sequence that included RNA helicase motif III, implicated in harnessing the energy of ATP hydrolysis to translocation. Such snapshots of proteins poised for action together with structural information and protein–protein interaction data should give us important new insights into the molecular basis of dynamics and catalysis in the spliceosome.

This work was supported by the Medical Research Council.

References
25 Vidal, V.P., Verdone, L., Mayes, A.E. and Beggs, J.D. (1999) RNA 5, 1470–1481

Received 23 June 2004