FUNCTIONS OF U-SNRNPS

JOAN A. STEITZ
Department of Molecular Biophysics & Biochemistry
Howard Hughes Medical Institute
Yale University School of Medicine
New Haven, Connecticut

The U-snRNPs of mammalian cells include the U1-, U2-, U5-, and U4/U6- containing particles of the nucleoplasm, the U3 RNP of the nucleolus, and a number of less abundant particles which are presumed to be nucleoplasmic. These last snRNPs are related to the abundant nucleoplasmic particles in that they are immunoprecitable with anti-Sm antibodies and have a 5' m_3G-cap on their RNA moieties. Their RNAs range in size from about 60 to 200 nucleotides. The U1-U6 snRNPs are present in about 10^6 copies per cell, whereas the least abundant Sm snRNP particle so far identified is probably in the 10^4 range.

U-snRNPs appear to be present in all eucaryotic species. Very recently, yeast (S. cerevisiae) analogs of mammalian U1-U6 RNAs have been identified (Ares, Cell (1986) 47, 49-59; Patterson and Guthrie, Cell (1987) 49, 613-624; Hughes et al., EMBO J. (1987) 6, 2145-2155; C. Guthrie, M. Rosbash, pers. commun.). Each of these, especially yeast U1 and U2, is distinctly longer than its mammalian counterpart, but exhibits striking secondary structure and some primary structure homology. U6 is the most conserved at the sequence level. The abundance of U-snRNPs in yeast is about 1,000-fold lower than in mammalian cells, and again these particular particles comprise only a subset of the snRNPs present.

snRNPs in Splicing

So far, functional assignments have been easiest to make for the most abundant mammalian snRNP particles. The U1, U2, U5 and U4/U6 snRNPs all appear to participate in pre-mRNA splicing. Their task is to recognize distant splice sites in a pre-mRNA, bring them into proximity, and align them for the precise cutting and ligation that occurs during the splicing reaction. A large (40-60S) active splicing complex, or "spliceosome" (Brody and Abelson, Science (1985) 228, 903-967; Grabowski et al., Cell (1985) 42, 345-353; Frendewey and Keller, Cell (1985) 42, 355-367), forms on each intron. Its assembly requires the participation of many protein factors in addition to snRNPs, including the hnRNP proteins that package the pre-mRNA (Choi et al., Science (1986) 231, 1534-1538).

Intron excision is a two-step process that involves an unusual branched RNA intermediate in both the mammalian and yeast systems (Ruskin et al., Cell (1984) 38, 317-331; Konarska et al., Nature (1985) 313, 552-557; Lin et al., J Biol. Chem. (1985) 260, 14780-92). In the first step, cleavage at the 5' splice site occurs coordinately with the formation of a 2',5'-phosphodiester bond between an A residue located some 30 nucleotides upstream from the 3' splice site and the conserved G at the 5' end of the intron. In the second step, a cut is made at the 3' splice site, the two exons are ligated, and the intron is released still in the form of a lariat. Both steps require ATP, although its phosphates are incorporated into neither the intermediates nor products of the reaction. Short consensus sequences have been developed for both the 5' ($^{C}_{A}AG/GU^{A}_{G}AGU$) and 3' ($Y_{>11}N^{C}_{U}AG/G$) splice sites. The branch is formed within an essential, conserved heptanucleotide (UACUAAC) sequence in yeast; a much weaker, more limited consensus has been derived for mammalian branch points (YNRAY), but the splicing apparatus can utilize another A residue

if this sequence is deleted.

Starting with the original hypothesis (1980) that U1 snRNPs function in splicing, experimental evidence implicating all four abundant mammalian snRNPs in this RNA processing reaction has slowly accrued. Data resulting from three types of experimental approaches have been most decisive: (1) Targeted destruction of individual snRNP RNAs in an active splicing extract using RNase H and a complementary deoxyoligonucleotide was first elegantly applied to U1, showing that both the snRNP and the 5' end of U1 RNA are essential for splicing (Krämer et al., Cell (1984) 38, 299-307). Subsequently, the U2 snRNP (Black et al., Cell (1985) 42, 737-750; Krainer and Maniatis, Cell (1985) 42, 725-736) and the U4/U6 particle (Black and Steitz, Cell (1986) 46, 697-704; Berget and Robberson, Cell (1986) 46, 691-696) have been implicated in the same manner. Results obtained in this type of experiment have further demonstrated that separate snRNPs contribute independently to the splicing reaction and are required early, before the appearance of splicing intermediates. (2) RNase protection experiments utilizing anti-(U1)RNP, anti-(U2)RNP, anti-Sm, or anti-m_3G antibodies have localized the binding sites for specific snRNPs on the pre-mRNA. This type of analysis has shown that U1 snRNPs bind 5' splice sites in normal (Mount et al., Cell (1983) 33, 509-518; Black et al., op.cit.) and 5' cryptic splice sites in mutated (Chabot and Steitz, Mol. Cell. Biol. (1987) 7, 698-707) transcripts; U2 snRNPs bind the intron region surrounding the branchpoint both before (Black et al., op.cit.) and after (Chabot and Steitz, Mol. Cell. Biol. (1987) 7, 281-293) the branch has been formed; and the U5 snRNP associates with the 3' splice site (Chabot et al., Science (1985) 230, 1344-1349) via a 70-100kd protein that itself exhibits selective binding (Gerke and Steitz, Cell (1986) 47, 973-984; Tazi et al., Cell (1986) 47, 755-766). (3) Genetic suppression experiments have revealed that splicing defective mutations in pre-mRNAs can be rescued by engineering compensatory base changes into U RNAs. Thus, base-pairing can be deduced to be critical to U1 snRNP recognition of the 5' splice site (Zhuang and Weiner, Cell (1986) 46, 827-835) and to yeast U2 interaction with the intron branch point (Parker et al., Cell (1987) 49, 229-239). Finally, finding snRNPs in splicing complexes isolated in various ways has further confirmed their participation (Konarska and Sharp, Cell (1986) 46, 845-855; Grabowski and Sharp, Science (1986) 233, 1294-1298).

Currently, the pathway of snRNP assembly into spliceosomes and the content of various pre-splicing complexes is under intense investigation. What is generally agreed is that the initial complexes formed at the branch point/3' splice site, minimally involving the U2 snRNP in both mammalian (Konarska and Sharp, op.cit.) and yeast (Pikielny et al., Nature (1986) 324, 341-345) systems, are more stable than those formed at the 5' splice site. Several observations hint that ATP hydrolysis may be required for the association/dissociation of snRNPs as the complexes build and change during the splicing process. It is not yet known whether any particular snRNP is directly responsible for holding the excised 5' exon in the spliceosome pending completion of the reaction; however, protected fragments from the 5' exon are recovered with anti-Sm antibodies (Chabot and Steitz, op.cit., 281-293). Direct interactions between the pre-mRNA and three of the four snRNPs involved in splicing provides a pleasing explanation for minimum intron size: the sum of the lengths of intron regions protected by the U1, U2 and U5 snRNPs is about 60 nucleotides. Simple geometry may therefore preclude use of smaller "introns."

It remains unclear whether any individual snRNP or other splicing component possesses catalytic (cutting or ligation) activity; alternatively, only the assembled spliceosome as a whole may be catalytically active in that it arranges intron sequences into a proper configuration for "RNA

catalysis" to occur. It is most intriguing that the branch-site A residue is predicted to bulge from an RNA double helix in the U2/intron interaction (Parker et al., Cell (1987) 49, 229), as well as in mitrochondrial introns that self-splice via lariat intermediates (Peebles et al., Cell (1986) 44, 213-224; Van der Veen et al., Cell (1986) 44, 225-236). Also unknown is whether changes in snRNP RNAs (each of which is encoded by multiple genes in mammals) or proteins contribute to the regulation of splicing observed in various tissues or at different developmental stages.

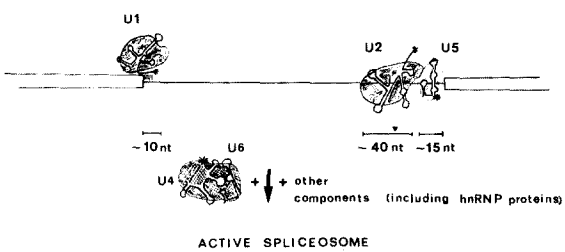

FIGURE 1. Known interactions of mammalian snRNPs with a splicing substrate. Lengths of intron regions protected from RNase are shown. The caret indicates the branchpoint. An order of assembly is not meant to be implied.

U3 RNP

The U3 RNP has long been postulated to be involved in some step of rRNA processing simply on the basis of its nucleolar location. Recently, the human U3 RNP has been shown to possess at least 5 proteins (Parker and Steitz, Mol. Cell. Biol. (1987) 7, 2899-2913) in addition to a previously identified 36kd protein (Lischwe et al., J. Biol. Chem. (1985) 260, 14304-14310). U3 has been suggested to participate in the cleavage that separates 5.8S rRNA from the internal transcribed spacer II (Crouch et al., Mol. Biol. Rep. (1983) 9 75-78; Bachellerie et al., Mol. Biol. Rep. (1983) 9, 79-86; Tague and Gerbi, J. Mol. Evol. (1984) 20, 362-367). However, recent analyses of U3 sequences that are available in the RNP (Parker and Steitz, op.cit.) make this proposal less attractive than the possibility that the U3 RNP contributes to some event occurring near the 3' end of 28S rRNA.

Minor Sm snRNPs

The only minor snRNP which has been assigned a function is sea urchin U7. It is essential for the correct 3' end maturation of sea urchin histone pre-mRNAs, in which an endonucleolytic cut is made between an upstream hairpin loop structure and a downstream conserved CAAGAAAGA sequence. The existence of the U7 snRNP was discovered in experiments where fractions from sea urchin cells were used to complement defective 3' end maturation of sea urchin transcripts in Xenopus oocytes (Birchmeier et al., Proc. Natl. Acad. Sci. USA (1984) 81, 1057-1061). Later, base-pairing between the sea urchin U7 RNA and the downstream element was demonstrated by genetic studies in which a 3' processing-deficient mutation in a histone pre-mRNA was suppressed by a compensatory base change in U7 (Schaufele et al., Nature (1986) 323, 777-781). Mammalian in vitro systems are also capable of accurate histone mRNA 3' end generation (Gick et al., EMBO J. (1986) 5, 1319-1326; Mowry and Steitz, Mol. Cell. Biol. (1987) 7, 1663-1673). An Sm snRNP is implicated by the results of both fractionation (Gick et al., op.cit.) and RNase protection experiments (Mowry and Steitz, op.cit.); in the latter, two interactions--early with the hairpin loop, later extending to include the downstream element--are seen. The mammalian counterpart of sea urchin U7 should soon be identified.

Polyadenylation is another nuclear RNA processing event that is suspected to involve an Sm snRNP. In vitro polyadenylation is inhibited by addition of anti-Sm antibodies (Moore and

Sharp, Cell (1985) 41, 845-855); in protection experiments using anti-Sm and anti-m_3G cap antibodies, recovered fragments included the hexanucleotide (AAUAAA) signal located just upstream from the cut site in four different polyadenylation substrates (Hashimoto and Steitz, Cell (1986) 46, 581-591). Despite purification attempts in several laboratories, this snRNP has so far eluded identification.

"Trans"-splicing, if it occurs in mammalian cells, is another process which could involve minor Sm snRNPs. The phenomenon of trans-splicing--the joining of exons from two separate RNA transcripts--is the mechanism of production of trypanosome mRNAs. A 35-nucleotide "spliced-leader" is added to the 5' end of every mRNA so far analyzed in this organism (Murphy et al., Cell (1986) 47, 517-525; Sutton and Boothroyd, Cell (1986) 47, 527-535). The recently determined sequence of trypanosome U2 RNA (Tschudi et al., Nucl. Acids Res. (1986) 14, 8893-8903) is striking in that it diverges from those of mammalian and yeast U2 within a highly conserved region, to produce extensive complementarity to the spliced leader. This could represent a mechanism for juxtaposing two separate transcripts before they become joined by branch formation during the first step of splicing. Although hypothetical, this further suggests that two different U2 (and perhaps other) snRNPs would have to exist in cells that carry out trans-splicing as well as the (normal) cis-splicing. The recent discovery that some transcripts in C. elegans undergo both trans- and cis-splicing (Krause and Hirsh, Cell (1987) 49, 763-761) means that such a phenomenon is not inconceivable in mammalian (or yeast) cells.

Exactly how many minor U-snRNPs exist in mammalian cells is not known. Their functions most likely relate to the processing of pre-mRNA or pre-rRNA, but participation in any other step along the pathway of gene expression (or in genome maintenance) should not be overlooked. Wherever specific recognition of a single-stranded nucleic acid sequence is required, the RNA moiety of a snRNP could well be involved.