Rearrangement of snRNA Structure during Assembly and Function of the Spliceosome[1]

MANUEL ARES, Jr.[2] AND
BRYN WEISER

Biology Department
Sinsheimer Laboratories
University of California, Santa Cruz
Santa Cruz, California 95064

I. General Features of Splicing 132
II. Dynamic RNA: Technical Considerations 136
III. RNA–RNA Interactions Early in Spliceosome Assembly 139
IV. RNA–RNA Interactions in the Assembling Spliceosome Prior to the First Chemical Step 143
V. Rearrangement of Spliceosomal RNA Structure during the Catalytic Steps of Splicing 146
VI. Conserved Residues in the Core of the Assembled Spliceosome 150
VII. Interactions during the Second Catalytic Step 151
VIII. Modified Nucleotides in Splicing 154
IX. Release of Spliced mRNA and Regeneration of snRNA Structures ... 156
X. Conclusions and Perspectives 157
References 157

In the past 15 years, an RNA *Weltanschauung* has evolved, progressing quickly from an initial astonished disbelief at the catalytic ability of naked RNA to the jaded assumption that there is little that RNA cannot be made to do with a clever *in vitro* evolution experiment. Interest in RNA is natural; in the two central processes of the eukaryotic gene expression pathway, splicing and translation, RNA is both the substrate and part of the enzyme. These "enzymes," the ribosome and the spliceosome, are large ribonucleoprotein complexes (RNPs), assembled by ordered and regulated means to achieve specific recognition of RNA substrates and carry out multiple chemical reactions. Our preoccupation with the RNA moieties of these enzymes stems

[1] Abbreviations: RNP, ribonucleoprotein particle; snRNA, small nuclear RNA; snRNP, small nuclear ribonucleoprotein particle; 4-sU, 4-thiouridine; SL-RNA, spliced leader RNA.

[2] To whom correspondence may be addressed.

Progress in Nucleic Acid Research and Molecular Biology, Vol. 50

from the hypothesis that they compose intrinsically, if not exclusively, the catalytic elements.

The chemical reactions of nuclear pre-mRNA splicing are not beyond the reach of RNA: the autocatalytic group-II intron RNAs perform them without the benefit of protein. Yet the spliceosome meets the task with dramatic stepwise changes in snRNA composition and conformation, escorted by an entourage of protein factors. Why does the spliceosome operate like a Rube Goldberg machine, performing such complex and unlikely events? The complexity of the spliceosome may be necessary to accommodate the vast numbers of different substrates and regulatory influences that complicate its work.

The core of the spliceosome contains a highly conserved RNA structure, suggesting that the chemical events of splicing are based heavily on RNA–RNA interactions. Yet, spliceosomes from different organisms are not exactly the same. Most studies on splicing have used human (HeLa) cells or the yeast *Saccharomyces cerevisiae*. Differences in splicing, and the intrinsic strengths and weaknesses of the experimental materials, have restricted some findings to one or the other system, but the conservation of snRNA and splicing factor structure and function means that the yeast and human spliceosomes are very similar. Some RNA–RNA interactions are essential in both systems. Some make more important contributions to function in one system than in the other, at least as assayed in the laboratory. Others interactions have not yet been shown to contribute to function, but are inferred to be critical for success in the ultimate bioassay, evolution. In this review, the idea that spliceosomal RNA works in fundamentally the same way in all systems is used to develop a cohesive representation of the rearrangement of the most conserved RNA–RNA interactions during splicing.

I. General Features of Splicing

In this section, we briefly review general features of splicing necessary as a context in which to place the RNA–RNA interactions described so far. More extensive reviews on splicing that contain discussion about the roles of RNA and protein factors, the effects of substrate mutations on splicing, the cellular localization of splicing factors, and the structure and function of snRNPs are recommended (*1–8*, *81–83*).

A. Two Transesterifications

With the development of cell-free splicing extracts, it was determined that two chemical reactions are performed by the splicing machinery on the substrate pre-mRNA (*1*). The first chemical reaction is a transesterification

consisting of an in-line nucleophilic attack on the phosphate between the 3′ nucleotide of the first exon and the 5′ nucleotide of the intron (the 5′ splice site) by a 2′ hydroxyl within the intron (the branchpoint) with the first exon as the leaving group (*1*, *9*; see Fig. 1). This "first-step" reaction produces the two splicing intermediates: free exon 1 and an intron–exon 2 molecule in

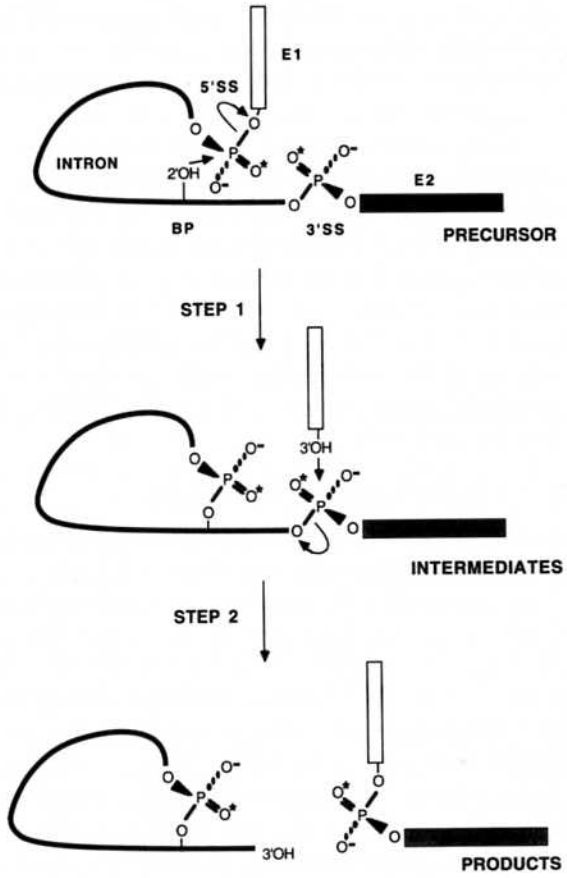

FIG. 1. Stereochemical course of splicing at the reactive phosphates. In the first step of splicing, the branchpoint (BP) 2′ hydroxyl (2′OH) attacks the phosphate at the 5′ splice site (5′SS) to form the lariat intermediate, and the 5′ exon (E1) is the leaving group. In the second step, the 3′ hydroxyl (3′OH) of E1 attacks the phosphate at the 3′ splice site (3′SS) to form the spliced exons, with the intron lariat product as the leaving group. Both reactions occur with the Sp phosphorothioate isomer, but not with the Rp isomer (*9*, *80*). Exon 1 (E1), white bar; intron, thick black line; exon 2 (E2), black bar; O^-, oxygen substituted by S in phosphorothioate Rp isomer, not tolerated; O^*, oxygen substituted by S in the functional Sp isomer. (Redrawn from *9*.)

"lariat" form, so called because the 5′ phosphate of the intron is now covalently attached to the branchpoint nucleotide, forming a closed circle of RNA with a 3′ tail (*1*).

The second chemical reaction is a similar transesterification, but this time the in-line attack is by the 3′ hydroxyl of exon 1 (the leaving group in the first reaction) on the phosphate between the last nucleotide of the intron and the first nucleotide of exon 2 (the 3′ splice site). This "second-step" reaction produces the two splicing products: the spliced exons (with the phosphate conserved from the first nucleotide of exon 2) and the intron in lariat form with its 5′ phosphate joined to the 2′ hydroxyl of the branchpoint (*1*, Fig. 1).

The number of phosphodiester bonds broken (one at the 5′ splice site of the pre-mRNA and one at the 3′ splice site of the lariat–exon 2 intermediate) is equal to the number of phosphodiester bonds formed (one at the branchpoint and one between the spliced exons, Fig. 1). The phosphorothioate stereoselectivity of the active site(s) for the two reactions is similar, and both reactions proceed with phosphate inversion (*9*, Fig. 1). This suggests that the second reaction is not carried out by removing the branchpoint from the active site, replacing it with the 3′ splice site and carrying out a reaction similar to the reverse of the first step (*9*). These and other data suggest that the spliceosome contains two distinct (but possibly overlapping) catalytic sites (discussed in *9*).

B. The Enzyme Is Built on the Substrate

The spliceosome is built of small nuclear ribonucleoprotein particles (snRNPs). The snRNPs are named after the snRNA they contain (U1–U22 . . .), and are classified by the common core proteins they contain (*6–8*). Not all snRNPs are involved in splicing, but those that are (U1, U2, U4/U6, and U5) share a common "core" set of proteins (the "Sm" proteins) that are tightly associated with the snRNA through a shared RNA sequence motif (*6–8*, the "Sm-binding site": $AU_{4-6}G$). Additional proteins may be found associated only with a particular snRNP; these are the "snRNP-specific proteins" (*6–8*). The protein composition of the different snRNPs remains an active area of research. As methods for the purification of snRNPs are improved, more and more snRNP-specific proteins are being identified.

The chemical reactions of splicing occur with great efficiency and accuracy, although the reactive phosphates and hydroxyls may be as many as tens of thousands of nucleotides apart in the primary transcript. The identification of splice sites in constitutive splicing and the regulation of alternative splicing are constrained by spliceosome assembly events that bring the regions of the pre-mRNA transcript together and into the catalytic structure. Spliceosome assembly follows an ordered pathway of snRNP binding, and is

influenced by the activity of numerous splicing factors (*1–6*). First, the U1 snRNP delimits the 5′ splice-site region. Then U2 binds near where the branchpoint will be to form the "prespliceosome." Next, a "tri-snRNP" containing U4, U5, and U6 snRNAs joins the complex. In the tri-snRNP complex, U4 is extensively base-paired to U6, and U5 association is mediated by proteins. The association of U1 with the complex is weak, and special measures must be taken to detect it (*1*). Prior to the first chemical step, the association of U4 also becomes weaker (*1–5*). The extent of the base-pairing between U4 and U6 suggests that an active process must take place to disrupt tight U4 association with the complex (*1–5*). The chemical steps of splicing follow.

In the yeast system, progression of the complex through these functional states is mediated by a set of proteins with amino-acid sequence similarity to ATP-dependent RNA helicases (*1–5*). The splicing protein members of the family consist of two subgroups, the "DEAD[3]-box" proteins (PRP5, PRP28, SPP81) and the related DEAH proteins (PRP2, PRP16, PRP22), named for a conserved amino-acid motif that is part of the nucleotide-binding domain. None has yet been shown to have helicase activity; however, many have RNA-dependent ATPase activity (*1,2*). The proteins are required at specific stages: PRP5 and PRP28 during spliceosome assembly, PRP2 after assembly but before the first step, PRP16 after the first step and before the second step, and PRP22 after the second step before release of spliced mRNA from the complex (*1–5*). Presumably the events catalyzed by each are unique and use ATP to drive the complex forward, possibly by initiating conformational changes. They may also play a role in the fidelity of splicing (*10*). The nature of their substrates remains mysterious.

C. U1 and U4 snRNPs Have Essential Functions Limited to Spliceosome Assembly

U1 and U4 are weakly associated with the spliceosome by the time the chemical steps of splicing begin. Recent evidence indicates that they are dispensable for the chemical steps of splicing (*11*; R. J. Lin, personal communication) and shows that U1 and U4 are not essential parts of the catalytic apparatus of the spliceosome. As described in Sections IV,B, V,A and V,B in more detail, interactions between U1 and the 5′ splice site as well as U4 and U6 are dissolved in favor of other important, mutually exclusive interactions. Thus U1 and U4 can be considered spliceosome assembly factors with roles in the construction of the catalytic apparatus of the spliceosome.

[3] DEAD = Asp-Glu-Ala-Asp; H = histidine; PRP = pre-RNA processing mutants; SPP81 = suppressor of *prp8-1*.

D. *trans*-Splicing: A Modified Spliceosome Built with an Exon–snRNA "Hybrid" snRNP

Certain protozoa and nematode worms carry out *trans*-splicing whereby an RNA, called the spliced leader (SL) RNA, containing a 5′ splice site is assembled into a snRNP, enters the spliceosome, and participates in the splicing reactions (reviewed in *12*). The SL-RNA donates its 5′ nucleotides as though they were exon 1, to a transcript carrying a 3′ splice site and a downstream exon. Branched molecules equivalent to the lariat of *cis*-splicing are observed, and the process requires U2, U4, and U6. In organisms where both *cis*- and *trans*-splicing take place, some of the same snRNAs are used for both processes, suggesting that the *trans*-spliceosome shares fundamental properties with the *cis*-spliceosome (*12*). In *trans*-splicing organisms not known to perform *cis*-splicing, neither U1 nor U5 has been found, suggesting that the snRNP component of the spliced leader may supply any necessary functions carried out by these snRNPs in *cis*-splicing (*12*).

II. Dynamic RNA: Technical Considerations

A. How Are Interactions Identified and Placed in Time?

Conformational rearrangements in the spliceosome are deduced or inferred from data that identify temporal changes in the interaction between different sets of snRNA and pre-mRNA nucleotides during splicing. Two main technical issues arise: how are the RNA–RNA interactions identified, and how are they placed in time? Identification of an RNA–RNA interaction in the spliceosome usually involves one or more of the following observations: (1) phylogenetic variation in sequence that conserves the potential for equivalent kinds of base–base interactions, (2) genetic experiments that identify functional combinations of non-wild-type nucleotides, (3) chemical or enzymatic probes of RNA secondary structure, and (4) cross-linking studies that position specific RNA residues near each other in splicing complexes.

Placing the requirement for a particular dynamic RNA–RNA interaction within the sequence of events during splicing is more difficult. An interaction may be required to get to or pass a particular landmark, or may form only after a particular landmark is passed. Failing definitive placement, correlation of the appearance and disappearance of the interaction relative to other splicing events must be used. A number of landmarks are available. The timing of events relative to the first and second chemical steps is informative. In yeast, temperature-sensitive mutations in splicing-factor genes provide numerous landmarks in the process (*2–4*).

For the many assembly events prior to the first step, time is marked with respect to the formation of presplicing complexes, including the commitment complex and the prespliceosome. In addition, spliceosome assembly can be blocked at specific stages by various treatments, such as ATP depletion, EDTA addition, mild heat treatment, or removal of specific snRNPs (*1*). Thus far, the number of discrete steps in the splicing pathway is unknown. In many cases the order in which certain landmarks occur relative to each other is also unclear. To generate images of the changes in RNA–RNA interactions in the spliceosome, we have reviewed data that can be interpreted to constrain RNA–RNA interactions in space and time. Because these data originate from distinct experimental approaches, we summarize and compare below the strengths and limitations of phylogenetic, genetic, structure-mapping, and cross-linking approaches.

B. Phylogenetic Variation and Dynamic RNA

Of the means for detecting interactions listed above, phylogenetic variation consistent with conservation of potential for base–base interaction comprises the bulk of the data used to determine the secondary structures of the snRNAs, much of which has been reviewed elsewhere (*13*). Such variation is useful for modeling the structure of regions of the spliceosome where the functional demands on RNA structure are sufficiently flexible that variation in sequence is allowed. Phylogenetic variation alone is inadequate for complete modeling of dynamic RNAs for two reasons. First, an extreme lack of variation characterizes regions of the RNA where multiple functional constraints are at play. Informative sequence variation will occur much less frequently in an RNA strand that interacts with two other strands at different times. Second, where variation does exist in dynamic RNA, it will be consistent with multiple structures, some of which may not be relevant to function. For example, phylogenetic data cannot be used to distinguish between a static requirement for a pseudoknot (*14*) and a dynamic requirement for both of the two component stem loops that may be derived from it. The limitations of the phylogenetic approach seem especially severe in the spliceosome, where large-scale changes in secondary structure occur during function.

C. Genetic Analysis of RNA–RNA Interactions in Splicing

In instances where strong conclusions elude the phylogenetic approach, reverse genetic techniques can be used to test models based on specific interactions. The power of this approach is based on *in vivo* function rather

than physical association. An interaction is detected when a phenotype caused by a mutation in one region of RNA can be suppressed by a compensating mutation in the interacting region of RNA (*1–5*). Thus, double-mutant combinations that function more like wild type than the single mutations are taken as evidence for a functional interaction between the altered nucleotides. The argument against indirect modes of suppression rather than suppression by direct restoration of RNA–RNA interactions is strengthened when allele specificity of suppression can be demonstrated. If several functional combinations of nucleotides are consistent with Watson–Crick (or other) base–base interactions, specify hydrogen bonds may be predicted.

For studying the timing of RNA rearrangements in splicing, the genetic approach is only partly satisfying. First, as with phylogenetics, genetic suppression results are most often displayed as *in vivo* phenotypes. Unless the interaction is specifically required for the second step, it will be difficult to place in time. Many conserved nucleotides have both an early and a late function, making mutational approaches to testing their late function cumbersome or impossible. Finally, there remains the nagging uncertainty that the observed suppression results from new or unusual compensating activities that wild-type nucleotides do not carry out during normal function. Nonetheless, this approach has been extremely successful, and has been used to study the U1–5′ splice site, U2–branchpoint, U4–U6, U2–U6, U5–exon, 5′ splice site–U6, and U1–3′ splice site interactions, as well as interactions that promote conserved internal structure in U2 and U6 (*1*, *5*).

D. Cross-linking and Chemical Accessibility of RNA in the Spliceosome

Detection of RNA–RNA interactions by UV cross-linking and UV-induced psoralen cross-linking, as well as detection of higher order structure by chemical cross-linking, has often been the first clue identifying interacting partners in the spliceosome. Cross-linking data can be very useful because the cross-linked partners must be very close.[4] The yield of cross-linked material is often low, and a constant worry is that a small percentage of aberrantly structured RNA complex is cross-linking very efficiently to produce a signal with little functional meaning. Occasionally such difficulties can be overcome by demonstrating functionality of cross-linked material, but more often other kinds of evidence supporting the interaction must be obtained to solidify the importance of a particular cross-linking result.

Other methods of probing RNA structure have been developed, but application to the study of RNA structural rearrangements during splicing has not been common. The main difficulty in the application of these methods to nuclear pre-mRNA splicing has been in the preparation of sufficient amounts of splicing complexes at discrete, known steps in the splicing process. Interpretation of the data where multiple conformations are present is difficult, and often the most interesting complexes are not abundant. So far, chemical probing has been useful in analyzing the structure of snRNAs within free snRNPs or snRNP complexes as they await entry into the spliceosome, but little of the dynamics of spliceosomal RNA has so far been revealed.

[4] "Closeness" in UV cross-linking is discussed by Budowsky and Abdorashidova in Vol. 37 of this series [Eds.].

E. Cross-linking of Site-specifically Photoactivatable Premessenger RNAs

Recently, intrinsic photolabels placed within pre-mRNA have been used to study the question of splice-site recognition and the dynamics of RNA in the spliceosome (*15*, *16*). In this approach, a photoactivatable 4-thiouridylate residue (4-sU) is built into the pre-mRNA substrate at a unique position near one of the splice sites. At times after splicing is begun, the reaction is pulsed with UV light, whereupon elements of the spliceosome near the substituted residue become cross-linked to the pre-mRNA at the substituted uridine. Analysis of the cross-linked material identifies the cross-linked partner, and primer extension can be used to map the sites of cross-linking. This methodology is extremely informative due to the unique placement of the photoactivatable nucleotide as well as the potential to "chase" cross-linked material through the splicing pathway (*16*).

III. RNA–RNA Interactions Early in Spliceosome Assembly

A. Interaction of U1 snRNA with the Pre-mRNA

The secondary structure of U1 snRNA is well illuminated by phylogenetic variation among the metazoans, but U1 from *S. cerevisiae* shares limited similarity with more typical U1 (*13*), including the conserved 5′-most stem loop. The sequence near the 5′ end of U1 is invariant and engages in base-pairing with 5′ splice sites during splicing in order to identify the region containing the bond to be attacked in the first step (*1–6*; Fig. 2). The precise definition of the splice site is not made relative to the U1 complementary sequence in the pre-mRNA; rather U6 seems to help make this decision later (*17*, *18*). The extent of complementarity between the 5′ splice site and the 5′ end of U1 influences its efficiency in splicing. The lower efficiency of some mutant 5′ splice sites can be increased by suppressor U1 snRNAs with

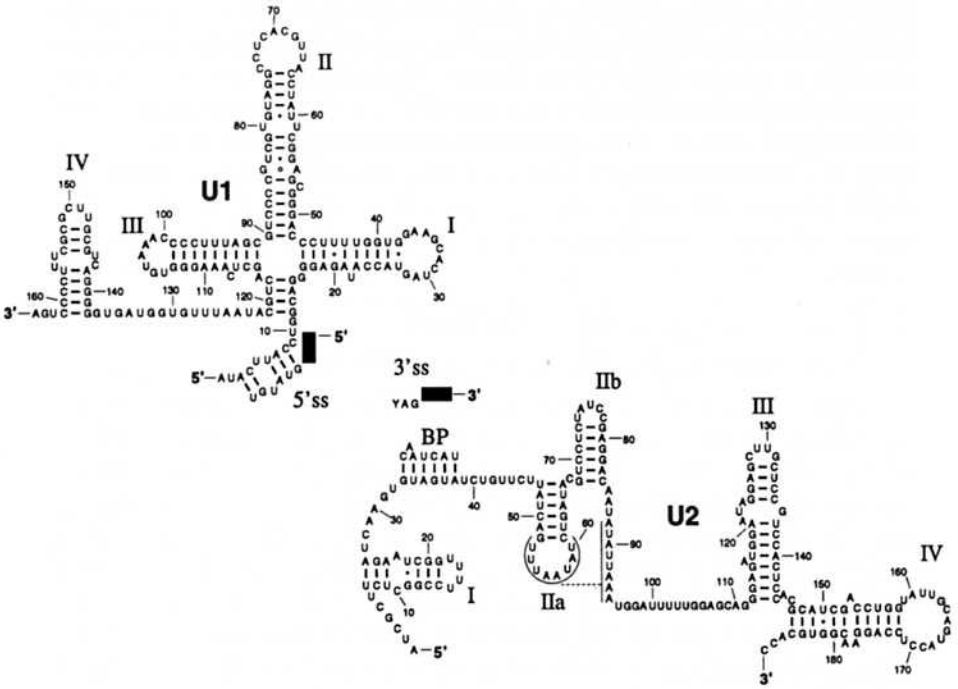

FIG. 2. RNA–RNA interactions in the prespliceosome. The sequences of human U1 and U2 are folded to show internal interactions as well as interaction between U1 and a 5′ splice site (5′-GUAUGU-3′) and between U2 and a branchpoint sequence (5′-UACUAAC-3′). The intron sequences connecting the 5′ splice site to the branchpoint and the branchpoint to the 3′ splice site are not shown. Exon 1, 5′ black bar; exon 2, 3′ black bar.

compensatory mutations in their 5′ ends that improve pairing to the mutant splice site (*19–21*). Because later requirements for splicing are superimposed on the 5′ splice-site consensus sequence, U1 suppression experiments are not effective at all positions in the U1 5′ splice-site helix.

U1 also contains a sequence (nucleotides 9–11) complementary to the 3′ splice site (Fig. 2), that is important in the identification of the 3′ splice site in the processing of introns in *Schizosaccharomyces pombe*. This interaction influences the first step of splicing in the "AG-dependent" class of introns *in vivo* (*22, 23*). In *S. cerevisiae*, spliceosome assembly and the first step of splicing are not strictly dependent on a 3′ splice-site AG (*1, 2*), and a test of this interaction in *S. cerevisiae* did not identify its role (*24*). However, complementarity between U1 nucleotides 9 and 10 and the last two nucleotides of the first exon (essentially extending the U1–5′ splice-site helix to include some of exon 1) did improve splicing efficiency at debilitated 5′ splice sites

(*24*). The negative results in *S. cerevisiae* do not strictly exclude a contribution to splicing, but the generality of the interaction between U1 nucleotides 9 and 10 and the 3′ splice site has yet to be established.

The function of U1 in identifying the 5′ splice site may be bypassed in certain instances. U1 depletion in splicing extracts of the nematode *Ascaris* blocks *cis*-spicing but does not block *trans*-splicing (*12*). To date, trypanosomes are known to carry out only *trans*-splicing; neither *cis*-introns nor U1 snRNA have been discovered in these organisms. Artificial *cis* constructs carrying SL RNA sequences in place of the 5′ end and the 3′ splice site of adenovirus or globin introns can be spliced in mammalian extracts in which U1 snRNA has had its 5′ end trimmed or blocked using complementary oligonucleotides (*25, 26*). Spliceosome assembly occurs in response to the addition of model 5′ splice-site oligonucleotides to extracts, apparently without the participation of the 5′ end of U1 snRNA (*27, 28*). These studies demonstrate that mechanisms may exist to bypass the need for U1–5′ splice-site interactions in certain instances.

B. Structure of U2 snRNA in snRNPs and Early Splicing Complexes

The second snRNP to bind the pre-mRNA substrate is U2. The stable binding of U2 snRNP is ATP dependent and requires branchpoint (yeast) or polypyrimidine tract (mammalian) sequences in the pre-mRNA (*1–6*). Although in mammalian extracts U2 snRNP can form complexes with substrates lacking a 5′ splice site, substrate commitment experiments argue that normally the binding of U2 snRNP occurs after formation of the U1-containing "commitment complex" (*1, 2*). Binding of U2 snRNP to the commitment complex forms the "prespliceosome." As analyzed by nondenaturing gel electrophoresis, the prespliceosome (complex III or B in yeast; complex A in mammals) contains U2 snRNP, but not the U4/U6.U5 snRNP (*1, 2*). The association of U1 snRNP with the prespliceosome is less stable and more difficult to demonstrate by native gel electrophoresis, but is detectable using less stringent conditions (*1, 2*).

Studies on U2 snRNA structure reveal that U2 snRNA must adopt different secondary structures during the course of its function in splicing. Chemical probing in yeast (*29*), and of both the 12-S and 17-S forms of the U2 snRNP in mammalian extracts (*30*) indicate that the RNA is folded as shown in Fig. 2. Pairing of nucleotides 7–14 with 19–26 (human numbering) forms "stem I," a structure consistent with chemical probing data (*29, 30*). Nucleotides forming this stem are highly conserved, and phylogenetic variation consistent with Watson–Crick pairing is observed only in the equivalent of the 14–19 base–pair in kinetoplastid organisms (*13, 31*). Genetic experiments designed to test for a contribution of the stem structure to function in

splicing have been negative, although the nucleotides composing the stem have other roles (*32–34*). An unusual conserved feature is the presence of non-Watson–Crick base appositions in the core of the stem. Hyperstabilization of the stem blocks splicing, revealing a requirement for appropriately tuned stability (*32*). These and other data (see below) indicate that this part of U2 must unfold later in splicing.

A phylogenetically supported stem-loop formed by pairing U2 nucleotides 47–52 with 61–66 (human numbering, "stem-loop IIa") (*13*, *31*; Fig. 2) is essential for U2 function (*31*). Also supported by phylogenetic variation is a base-pairing between the 8-base "loop IIa" (nucleotides 53–60) and a sequence downstream (the "conserved complementarity," nucleotides 88–95; dotted line in Fig. 2) (*13*, *31*). Genetic experiments do not demonstrate a function for the complementarity to the loop (*29*), and structure-probing experiments show that the loop and the conserved complementarity are accessible to chemical probes in the bulk of snRNPs. Mutations in yeast that destabilize stem IIa cause the RNA to be misfolded so that the conserved complementarity is paired to the loop IIa (*35*), causing cold-sensitive splicing and growth defects that can be suppressed by destroying the conserved complementarity (*36*).

Analysis of yeast splicing extracts made from cold-sensitive U2 strains shows that prespliceosome formation is blocked at restrictive temperatures (*35*). Contrasting results obtained in HeLa-cell extracts show that oligonucleotides directed against loop IIa and adjacent stem sequences allow spliceosome assembly and the first step of splicing, but interfere with the second step of splicing *in vitro* (*37*). Chemical cross-links can be formed within human U2 RNA that are compatible with the pairing of the loop to the conserved complementarity (*38*). One consistent interpretation of all the data argues that this region of U2 adopts more than one structure during the functional cycle. First, stem IIa is required for a spliceosome assembly step, then, after the first step of splicing but before the second step, this part of the RNA refolds and becomes accessible to the oligonucleotide. How or why this region of U2 might rearrange remains to be determined.

C. Interaction between U2 and Pre-mRNA

U2 base-pairs with the intron branchpoint region. Mutations in nucleotides 33–38 (human numbering) of U2 that increase complementarity to mutant branchpoint regions greatly improve their use in yeast (*39*) and mammals (*40*,*41*). In yeast, certain mutations in the branchpoint interaction region of U2 cause dominant slow-growth defects (*42*). Presumably this phenotype is due to a partial function of the mutant U2 snRNP that prevents wild-type U2 snRNP from helping remove introns at a rate consistent with growth, and suggests that stable complexes involving the mutant snRNP

sequester or deplete a limiting substoichiometric splicing factor *in vivo*. Because branchpoint complementarity to U2 snRNA is not required for the dominant phenotype, steps leading to stable addition of U2 snRNP may occur prior to recognition of the branchpoint by base-pairing to U2 in yeast *in vivo* (*42*).

Although there is an intrinsic branchpoint binding ability of mammalian U2 snRNPs (*43*), stable U2 binding to the branchpoint region in mammalian extracts is ATP dependent and is assisted by U2 auxiliary factor (U2AF) through adjacent polypyrimidine tracts in the intron (*1*, *5*). Thus, in both yeast and mammals there is evidence that base-pairing between U2 and the branchpoint region is not the only factor that contributes to the initial selection of the U2 binding site in pre-mRNA.

Psoralen cross-linking experiments demonstrate that the U2–branchpoint interaction is established early in the course of splicing in mammalian extracts, well before the first chemical step (*44*). Surprisingly, the U2–branchpoint cross-link occurs independently of ATP, suggesting that this RNA–RNA interaction may take place before U2 snRNP binding is defined as "stable" by the biochemical assays normally employed (*44*). Many complexities of mammalian branchpoint selection remain to be explained, such as why very poor complementarity to U2 is tolerated in some mammalian introns (*45*), and how several different adenosine residues in the same region can be used as branchpoints (*46*). Although the conserved yeast branchpoint is often used as a model because of extensive complementarity to U2 and the provocative adenosine, there is as yet little direct evidence to support the extent of interaction commonly depicted (Fig. 2), despite evidence for the bulged attacking residue (*47*).

IV. RNA–RNA Interactions in the Assembling Spliceosome Prior to the First Chemical Step

A. Extensive Interaction between U4 and U6 snRNAs

The prespliceosome is the substrate for the next step of spliceosome assembly, addition of the U4/U6·U5 tri-snRNP (*1–6*). The tri-snRNP is formed by the association of the U4/U6 snRNP with the U5 snRNP (*1–6*). A site of association between U4 and U6 *48*, *49*) was first mapped by psoralen cross-linking (*50*). In the U4/U6 snRNP (and the tri-snRNP) the interaction is so stable that deproteinized U4–U6 hybrids are efficiently recovered after cold phenol extraction and can be analyzed by nondenaturing gel electrophoresis (*49*, *50*; Fig. 3). U6 lacks the Sm binding-site consensus sequence;

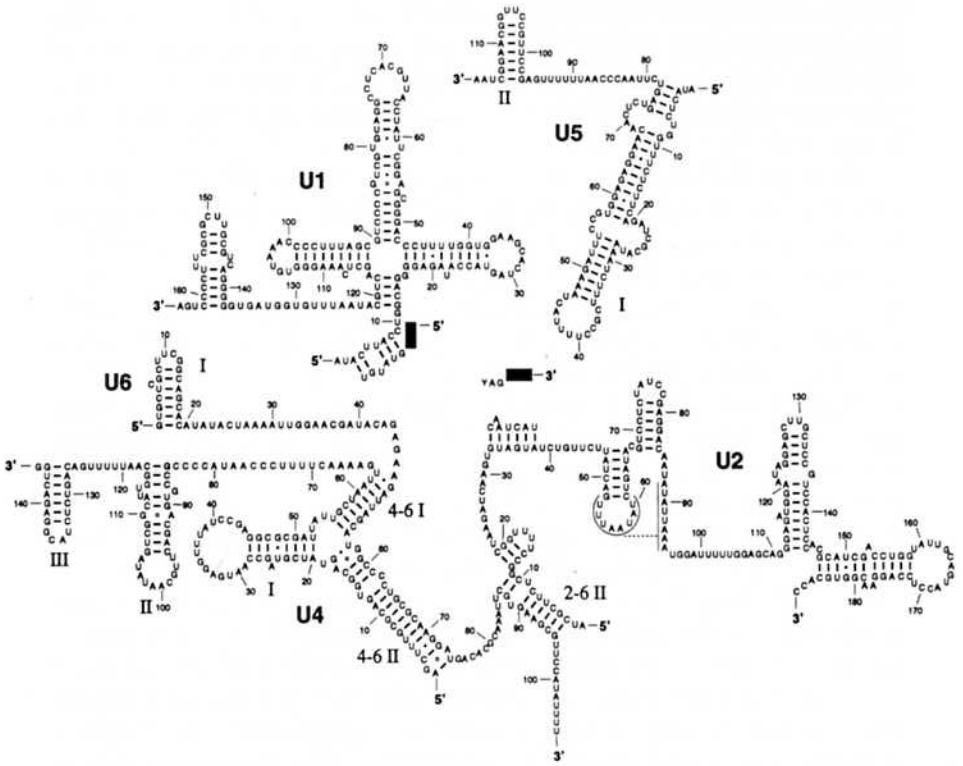

FIG. 3. RNA–RNA rearrangements in the assembling spliceosome. The sequences of the human spliceosomal snRNAs are folded to show interaction established during spliceosome assembly. Extensive interaction between U4 and U6, as well as between the 5′ end of U2 and the 3′ end of U6, is shown. The intron sequences connecting the 5′ splice site to the branchpoint and the branchpoint to the 3′ splice site are not shown. Exon 1, 5′ black bar; exon 2, 3′ black bar.

however, it appears in Sm antibody immunoprecipitates of snRNPs by virtue of its interaction with U4 (*6*).

Phylogenetic data allowed the development of a model for the interaction between U4 and U6 snRNAs (*51*; Fig. 3). There are two stem regions involving U4 and U6, separated by an internal U4 stem structure that creates a Y junction. The U6 sequences that interact with U4 in the U4/U6 snRNP and in the tri-snRNP are also required to participate in other RNA–RNA interactions later in spliceosome assembly. The constraints on the U6 sequence to maintain the ability to participate in both sets of interactions must account for some of the high degree of U6 sequence conservation through evolution.

B. Structure of U5 snRNA

The most prominent feature of U5 snRNA is the conserved 11-nucleotide loop at the end of a conserved stem structure (loop I in Fig. 3). The loop is accessible to solvent, as shown by chemical probing experiments (*52*). The extended stem is punctuated by an internal loop containing an invariant CCG sequence on the 3′ side and a longer, less well-conserved sequence on the 5′ side (*13*, *53*; Fig. 3). At the base of the stem to the 3′ side is the Sm binding site, and following the Sm site there is a 3′ terminal stem loop. An extra sequence can be found at the 5′ end of U5 in mammalian cells, encoded by variant genes (*54*). The single yeast U5 gene produces two RNAs that differ at their 3′ ends, with the shorter version lacking the 3′ stem loop (*13*). Yeast U5 also has an extra stem loop projecting from the main stem near the conserved internal loop (*13*, *53*). Genetic experiments have identified an interaction between the nucleotides of the U5 loop and the exon sequences of pre-mRNA (discussed in Sections V,E and VII,A).

C. Interaction between the 5′ End of U2 and the 3′ End of U6

Psoralen cross-linking of snRNAs in mammalian cells first identified an interaction between the 5′ end of U2 and the 3′ end of U6 (*55*; helix 2–6 II in Fig. 3). The potential to form this interaction is phylogenetically conserved from humans to yeast and trypanosomes. Genetic studies in transfected mammalian cells have measured the activity of a suppressor U2 (able to recognize a splicing substrate with a mutant branchpoint) carrying second mutations in the sequence complementary to U6. Such second mutations block suppressor U2 function, but function was regained by transfection of a U6 gene containing a compensatory mutation that restored the helix (*56*, *57*). In yeast, U2 or U6 mutations in the component strands of the helix do not block growth, suggesting that this interaction is not absolutely essential for splicing (*42*, *58–60*). Its broad conservation and the results in mammalian cells suggest that it contributes to the efficiency of splicing.

In the assembling spliceosome, the interaction between the 5′ end of U2 and the 3′ end of U6 may help the tri-snRNP bind to the prespliceosome, or serve to position the assembling RNA elements within the complex (*1*). Splicing extracts probably contain abundant U2·U4/U6·U5 snRNP complexes in the absence of added pre-mRNA, because psoralen cross-linking experiments reveal the double cross-linked U4–U6–U2 snRNA product in amounts expected given the efficiency of the individual U4–U6 and U6–U2 cross-linked species (*61*). For this reason, U6 is shown paired with U4 and U2 simultaneously (Fig. 3).

D. The Departure of U1

Demonstrating association of U1 with splicing complexes containing other snRNPs has been technically challenging (*1*). The weak association of U1 with splicing complexes, as well as the interaction of U6 with the 5′ splice site, has led to the proposal that U1 leaves the splicing complex some time after defining the 5′ splice site. Thus, although Fig. 3 is drawn to contain all the splicing snRNAs currently known, evidence for a stable complex containing all of them simultaneously is scant. In a direct test of the involvement of U1 in the splicing reactions, assembled yeast spliceosomes blocked at the PRP2 step were isolated and stripped of detectable U1 snRNA (R. J. Lin, personal communication). Such spliceosomes could be chased through the splicing reactions on addition of appropriate factors in the absence of U1, indicating that U1 is not necessary for the catalytic steps of splicing. Recently, a mutually exclusive interaction between U6 and the 5′ splice site has been identified (*27*, *28*). These observations, as well as the weak biochemical association of U1 with assembled spliceosomes, suggest that U1 snRNP leaves the spliceosome prior to the first step of splicing.

V. Rearrangement of Spliceosomal RNA Structure during the Catalytic Steps of Splicing

A. An Interaction between U6 and the 5′ Splice Site

UV cross-linking first identified an interaction between U6 and pre-mRNA (Figs. 4 and 5), as well as between U6 and the lariat intermediate, near substrate sequences at the 5′ splice site in both yeast (*62*, *63*) and mammalian (*16*, *61*) extracts. Two distinct base-pairing models between U6 and the 5′ splice site region were tested, and one, in which the 4, 5, and 6 positions of the splice site pair with the first three nucleotides of the invariant ACAGAGA box of U6, is supported by compensatory mutations in both RNAs (*17*, *18*; Fig. 5). This interaction would seem to be mutually exclusive with pairing of the 5′ splice site with U1 (see Fig. 4; compare Fig. 3 with Fig. 5), and it is inferred that after identification of the 5′ splice site region by U1, U6 takes over from U1 and assists in specifying the precise bond to be attacked in the first step of splicing. This interaction explains why certain suppressor U1–5′ splice site mutations do not function: U6 interactions with the mutant substrate remain perturbed and result in inaccurate specification of the bond to be cleaved (*17*, *18*).

Model oligonucleotides representing the 5′ splice site can induce the assembly of U2·U4/U6·U5 complexes, even when the 5′ end of U1 is

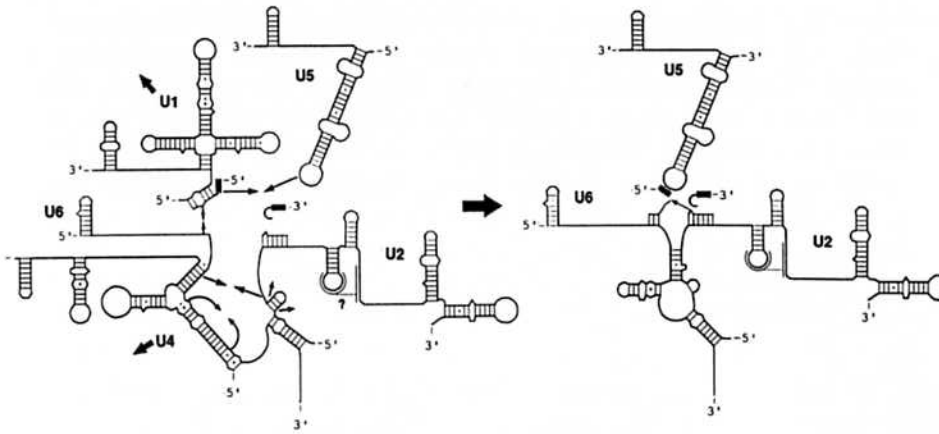

FIG. 4. Rearrangement of snRNA structure during activation of the spliceosome. A representation of human snRNAs is given in Fig. 3, showing changes that occur between assembly and activation. During this time, U1 leaves the 5′ splice site to U6, U5 establishes contact with exon 1, and U4 leaves U6, which refolds on itself, as well as establishing new contacts with U2.

blocked (*27*, *28*). Studies using 4-sU-substituted transcripts show that the second nucleotide of the intron (the U of the conserved GU) can be cross-linked to the third A of the U6 ACAGAGA sequence, but only after the first step of splicing (*16*). This suggests that the interaction between U6 and the 5′ splice site region may be altered between the first and second steps of splicing.

B. U4 Is Unwound from U6

Before the first catalytic step of splicing occurs, the association of U4 snRNA with the splicing complex becomes greatly weakened or lost (*1–5*; Fig. 4). As with U1, the question of whether U4 is necessary for the catalytic steps of splicing has been addressed by assembling yeast spliceosomes blocked before the first step and stripping them of U4 snRNA. Spliceosomes lacking U4 can carrying out the splicing reactions on addition of the appropriate factors (*11*), arguing that U4 snRNA does not participate in the catalytic steps of splicing.

Because of the extent of interaction between U4 and U6, the destabilization of U4 at physiological temperatures must be an active process (*51*). In addition, the dramatic loss of such a significant segment of structured RNA argues that U6 snRNA must adopt other structures in the absence of U4, and that these structures may be particularly important to the catalysis of the splicing reactions (*51*).

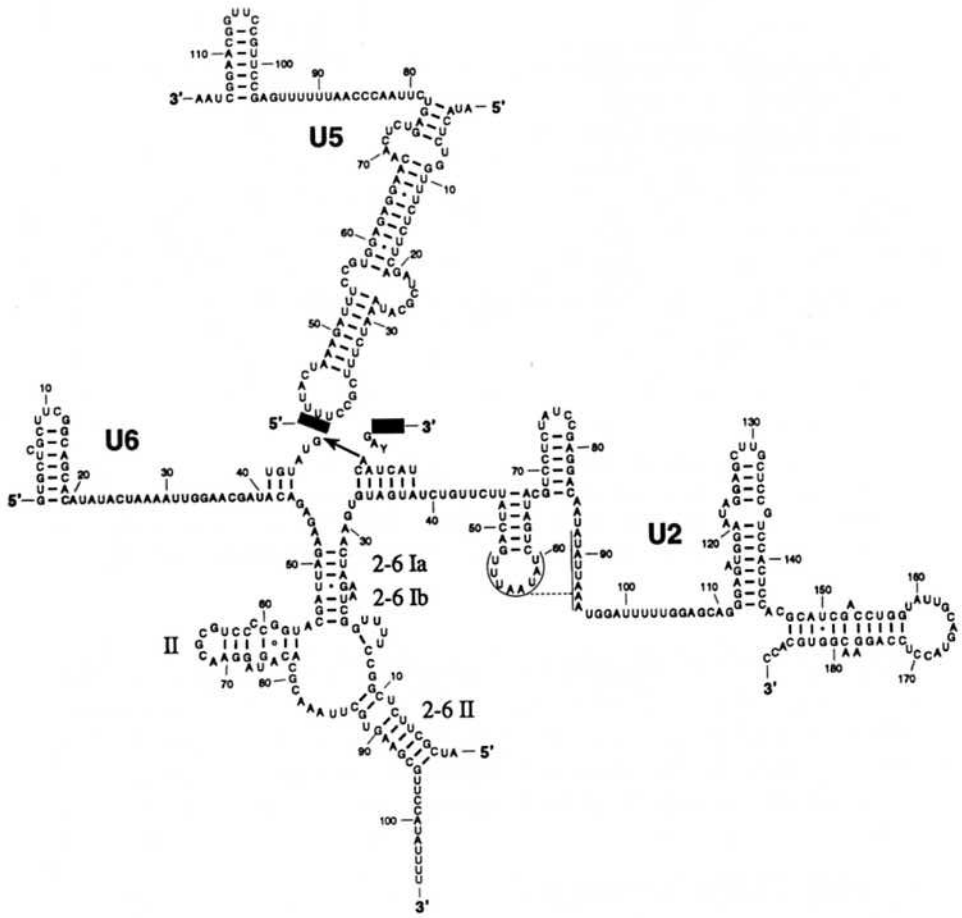

FIG. 5. RNA–RNA interactions in the spliceosome after rearrangement. Sequences of human U2, U6, and U5 are folded to show interactions important for the catalytic steps of splicing. As modeled in Fig. 4, an internal stem has formed in U6 (residues 57–78), and residues 49–55 of U6 interact with residues 20–28 of U2 to generate the U2–U6 helix Ia and Ib structures. In addition, U5 residues 40–42 interact with the last three residues of exon 1. The arrow indicates the attack of the 2′ hydroxyl of the branchpoint on the 5′ splice site. The intron sequences connecting the 5′ splice site to the branchpoint and the branchpoint to the 3′ splice site are not shown. Exon 1, 5′ black bar; exon 2, 3′ black bar.

C. The Refolding of U6

What happens to the U6 structure when U4 leaves? Phylogenetic, genetic, and structure-probing data indicate formation of a conserved internal stem in U6 snRNA when U4 is not bound (*64*, *65*; Figs. 4 and 5). The

sequences of U6 that form this internal stem (sometimes called "the 3′ stem") are made up of those that form stem II of the U4/U6 structure (Figs. 3–5). The internal stem of U6 is balanced in stability relative to the U4/U6 interaction (*64*, *65*). Mutations that hyperstabilize the stem result in a cold-sensitive growth phenotype in yeast (*65*). Suppressors of the hyperstabilizing mutations are similar to suppressors of a cold-sensitive mutation in U4 that destabilizes the U4/U6 interaction (*66*). These results argue that the relative stabilities of the U4/U6 interaction and the U6 internal stem are so exquisitely tuned that neither has a large stability advantage over the other. This might be expected of nucleic acids that must interconvert between structures.

D. Establishment of Additional Interactions between U2 and U6

Formation of the internal U6 stem replaces many but not all of the U4/U6 interactions. The region of U6 that participates in U4/U6 stem I is highly conserved and shows complementarity to U2 nucleotides 22–28 (human numbering), including two bulged residues (4–6 I in Fig. 3). Genetic experiments in yeast show that complementarity between U6 and U2 in these regions is required for function (*33*; Figs. 4 and 5). Mutations in several of the U2 residues that form these interactions are lethal and interfere with both steps of splicing *in vivo* (*33*). *In vitro*, such mutations preferentially block the second step (*34*). The extended interaction is referred to as U2/U6 helix I, with the longer of the two segments designated helix Ia and the shorter helix Ib (2–6 Ia and 2–6 Ib in Fig. 5). Mutation of the corresponding residues in mammalian U6 (*67*, *68*) and U2 (*32*) cause slight loss of function. Mutations that hyperstabilize U2 stem I (Fig. 2), which must be disrupted to form both the U2/U6 helix II and helix I interactions, can block splicing (*32*), suggesting that U2/U6 helix I is also important for mammalian splicing. The dependence on helix I and helix II for efficient splicing seems different in yeast and mammalian cells: mammalian splicing is more sensitive to changes in helix II than in helix I, but yeast growth is more dependent on helix I than on helix II. These differing dependencies could be due to the assays used, or to intrinsic differences in the splicing machineries of the two organisms.

E. U5 Interacts with the First Exon

The highly conserved U5 loop nucleotides interact with the exon regions of the pre-mRNA (Figs. 4 and 5). In yeast, improvement in Watson–Crick complementarity between U5 U residues 96–98 (40–42 in human U5) and the three exon nucleotides just upstream of the splice site positively influences 5′ splice site cleavage (*69*, *70*). U5 genes from a library in which the U5 loop sequence was randomized were selected on the basis of their ability to

suppress a 5′ splice-site mutation. The activities of a number of such suppressor U5 genes are explained by the hypothesis that residues 96–98 (40–42 in human U5) stabilize an interaction between the splicing machinery and exon 1 of the mutant substrate (*69*, *70*).

Using mammalian extracts and a pre-mRNA substrate carrying a photoactivatable 4-sU residue as the next to last (*15*) or last nucleotide of the first exon, cross-links can be detected between the substituted exon residue and U40 and U41 of U5 (*16*). The time course of appearance of these cross-links places them before the first step of splicing, but after formation of a detectable U1–substrate cross-link (*16*). As the splicing reaction proceeds, the early cross-link with U1 becomes less efficient as the U5 cross-link appears. This pattern is consistent with establishment of contacts between U5 and the exon sequences as U1 is displaced from the 5′ splice site (Figs. 4 and 5). Presumably the interaction between U6 and the 5′ splice site is also established at this time, because it becomes detectable with similar kinetics using a substrate with the photoactivatable residue at a different site (*16*).

F. The First Chemical Step

Once the appropriate structures are formed, the transesterification reaction can proceed. The 2′ OH selected as the nucleophile attacks the phosphate at the 5′ end of the intron (arrow, Fig. 5), releasing the first exon and creating a (5′)A 2′–5′ G(3′) dinucleotide that is uniquely present in spliceosomes having completed the first step of splicing. The free exon may be held in position in part through its interaction with U5 (*16*, *69*, *70*).

VI. Conserved Residues in the Core of the Assembled Spliceosome

Figure 6 shows a model for the secondary structure elements formed by rearrangement of the snRNAs during formation of the spliceosome. The sequence of human snRNAs is used to create the model. Conserved secondary structure elements that vary widely in primary sequence are represented by stem loops. Nucleotide differences between human and yeast (*13*) are indicated by arrows. Because the lengths of mammalian and yeast U2, U6, and U5 differ, two numbering systems are superimposed on the model. It is obvious that the most highly conserved residues are in the core of this folded RNA structure, as are the sites of the chemical reactions. Variation is restricted to peripheral elements, or is consistent with the maintenance of secondary structure elements (Fig. 6).

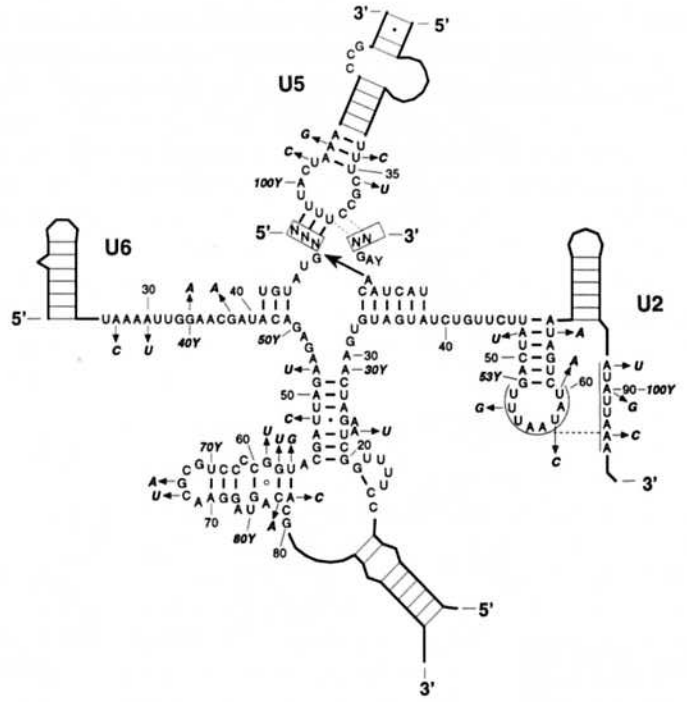

FIG. 6. Sequence variation between yeast and humans in core RNA elements of the spliceosome. The human sequence is folded as in Fig. 5. Regions of structural conservation, but little primary sequence conservation, are represented without sequence. The highly conserved sequences are shown, and substitutions in the yeast snRNAs are indicated by arrows and bold italics. Standard numbering is human, numbers followed by *Y* indicate the number of the homologous nucleotide in the yeast spliceosome.

VII. Interactions during the Second Catalytic Step

A. U5 Interacts with the Second Exon

Splicing of yeast pre-mRNAs is blocked by changes in either the first or last G of the intron. If an A is selected as the 5′ intron nucleotide in the first step, then the second step is blocked, producing "dead-end" lariats. This defect can be suppressed by mutations in U5 that improve Watson–Crick complementarity between U5 nucleotides 95 and 96 (human nucleotides 39 and 40) and the first two nucleotides in the second exon (*70*; Fig. 7). Suppression is not limited to dead-end lariats formed by improper 5′ splice-site selection: substrates blocked at the second step by mutation of the 3′ splice

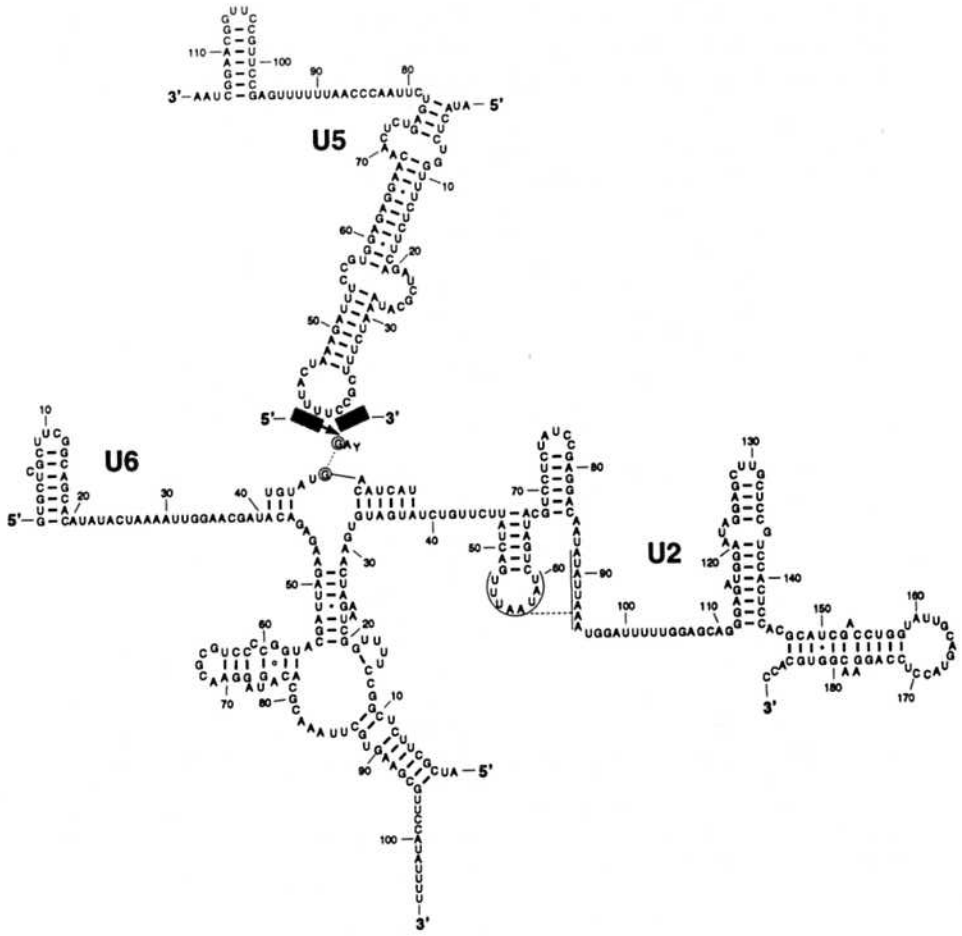

FIG. 7. RNA–RNA interactions during the second step of splicing. Human sequences are folded to show new interactions that form between U5 residues 39 and 40 and the first two nucleotides of the second exon. An interaction between the first and last nucleotides of the intron (circled) is indicated by a dashed line. The intron sequences connecting the 5′ slice site to the branchpoint and the branchpoint to the 3′ splice site are not shown. Exon 1, 5′ black bar; exon 2, 3′ black bar.

site AG to AA can also be processed in cells carrying U5 with improved complementarity to the second exon. These observations indicate that interactions between U5 and the second exon are important in the selection of the 3′ splice site for the second step of splicing (*70*). U5 position 96 (human U5 residue 40) complementarity to the last nucleotide of exon 1 improves the efficiency of (aberrant) 5′ splice-site cleavage, and complementarity of the

same U5 nucleotide to the first nucleotide of exon 2 improves efficiency of 3′ splice-site cleavage when either the branch or the 3′ splice site is noncanonical (*69*, *70*). This suggests the register of the U5 loop with respect to exon 1 may change between the first and second step (*70*).

In mammalian cell extracts, arguments for interaction between U5 and both exons are more direct. Pre-mRNA substrates carrying a single 4-sU residue at the last nucleotide of exon 1 can be cross-linked to U5 before the first catalytic step, and the cross-linked material can be chased through the splicing reactions (*16*). Furthermore, substrates carrying the 4-sU substitution at the first nucleotide of the second exon become cross-linked to U5 only after the first step of splicing, consistent with a role for the U5 loop in binding the second exon (*16*).

B. Second-Step Function of Nucleotides in the U6 ACAGAGA Sequence

The ability to reconstitute splicing extracts depleted of endogenous U6 snRNA using synthetic U6 allows mutational analysis of U6 function in splicing *in vitro*. Among other things, the data identify nucleotides that, when altered, allow spliceosome assembly and the first step of splicing, but block the second step in both yeast (*71*) and mammalian (*68*) systems. The third A and (in yeast) the following G of the ACAGAGA sequence (A51, G52 in yeast, *71*; A45 in human U6, *68*) appear to be specifically required for the second step because their mutation results in a strong accumulation of splicing intermediates but little or no product. Cross-linking data, using 4-sU substitution at the U of the 5′ splice site GU dinucleotide, show that, after the first step of splicing in human extracts, A45 of U6 is very near the 5′ splice site (*16*).

C. U2·U6 Helix Ia Has an Important Second-Step Function

Nucleotides in U6 and U2 interact by Watson–Crick base-pairing to comprise helix Ia (*33*). In reconstituted yeast splicing extracts *in vitro*, mutation of either a conserved pyrimidine (C58 yeast, *71*; U52 human, *68*) or an invariant A (A59 yeast, *71*) in U6 produces a strong block to the second step. Mutations in yeast U2 G26 or A27 also allow the first step but block the second step of splicing *in vitro* (*34*). These residues form the two base pairs of the U2·U6 helix Ia nearest the two bulged bases that separate helix Ia from helix Ib (2–6 Ia, Figs. 5 and 7, *33*). This structure may form prior to the first step as proposed (*33*); however, its first step function (if any) is not essential (*34*). Considering the dynamic properties of the spliceosome, the possibility remains that critical elements of this structure do not form until after the first step.

An interaction between the bulged U2 residue A25 and the U6 G52 in the U6 ACAGAGA sequence has been proposed, on the basis of suppression data in yeast (Fig. 7; 72). Suppression could be indirect, but if it represents direct contact between the two bases, it suggests that the U2·U6 helix I may be folded back onto the ACAGAGA element. The role of these nucleotides in the second step is mentioned above; here, suppression of 3′ splice-site mutations provides additional support for a second-step function (72). As the secondary-structure elements of the spliceosome become defined, tertiary-structure elements will provide spatial relationships between the known helices. In this case, a model for the spatial relationship between these regions also links the key nucleotides specifically involved in the second step (72).

D. The Identities of the Branchpoint Bases and Last Intron Base Influence the Second Step

The first intron base and the base at the branch can influence the efficiency of the second catalytic step (*1–5*). Normally, the branch is A(2′–5′)GU. First-step reactions leading to, for example, A(2′–5′)A, A(2′–5′)U, or C(2′–5′)G branches are blocked for the second step of splicing in yeast (*2–4*). In the case of A(2′–5′)AU branches, the second step can proceed with modest efficiency if the 3′ splice site is AC rather than AG, arguing that the branched bases influence 3′ splice site selection (Fig. 7; 73). Branched C(2′–5′)G substrates can be encouraged to participate in the second step by virtue of the dominant suppressor activity of certain PRP16 mutations (*2–4*). In mammalian cells, A(2′–5′)G is the most common branch dinucleotide, but branches with U(2′–5′)G are also observed. In certain instances when G carries the attacking 2′ hydroxyl [forming a G(2′–5′)G branched dinucleotide], the second step is inhibited *in vitro* (*47*). Whether the branchpoint base influences the second step by virtue of a base-specific interaction has not been investigated.

VIII. Modified Nucleotides in Splicing

Nucleotide modification is common in structural RNAs such as tRNA and ribosomal RNA, and the spliceosomal RNAs are no exception (75). Figure 8 shows the known nucleotide modifications in vertebrate snRNAs, folded as they might be in the activated spliceosome. Strikingly, the majority of the modified nucleotides are located in the core of the structure. Numerous invariant U residues in U2 and U5 are modified to pseudouridine, adding a potential hydrogen bond donor to the 5 position of the pyrimidine ring at all

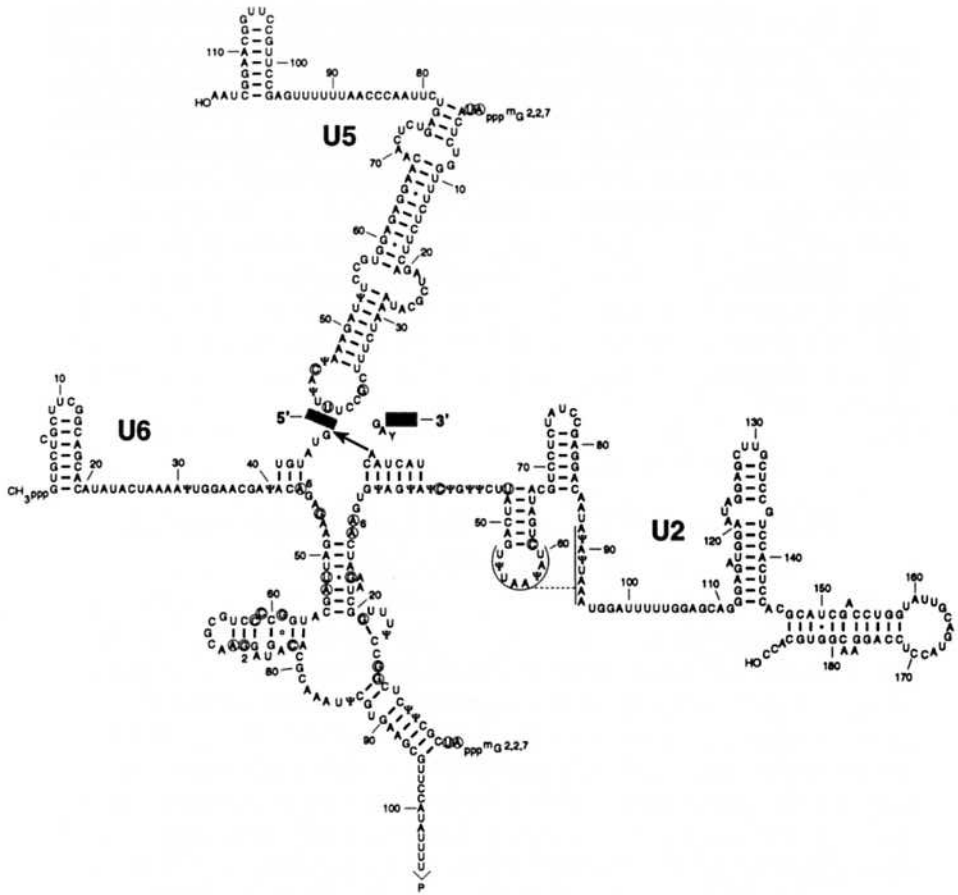

FIG. 8. Sites of nucleotide modification in the activated spliceosome. Figure 4 is redrawn here to show the modifications internal to and at the ends of the human spliceosomal snRNAs. Greek letter psi (ψ), pseudouridine; A6, N^6-methyladenine; G2, N^2-methylguanosine; circled residues, 2′-OCH_3 sugar modification. The intron sequences connecting the 5′ splice site to the branchpoint and the branchpoint to the 3′ splice site are not shown. Exon 1, 5′ black bar; exon 2, 3′ black bar.

these residues. The A43 in the conserved ACAGAGA sequence of U6 is methylated on the N6 position, possibly restricting certain interactions with this base (Fig. 8). Likewise, the function of the invariant A30 of U2 could be influenced by N6 methylation. An N2 methyl group on G71 of U6 in the central stem is also present. Methylation of 2′-OH groups of the ribose moieties is also common in the conserved core.

So far, no nucleotide modification in the spliceosomal snRNA has been ascribed a function. As suggested, methylation of potential hydrogen bond donors may prevent competing nonfunctional RNA–RNA interactions from forming. Alternatively, such modifications may increase the hydrophobic character of certain RNA structural elements and help exclude water from the catalytic apparatus as it folds. Finally, methylation of 2′ hydroxyls may protect adjacent phosphodiester bonds from attack, as well as preventing certain 2′-OH groups from being misidentified as nucleophiles (Fig. 8). Little information is available about modifications in the yeast spliceosomal snRNAs; however, mutation of many yeast U2 U residues equivalent to those present as pseudouridine in mammalian U2 is not lethal (*31*, *34*; D. Yan and M. Ares, unpublished), suggesting that many modifications may not be essential.

IX. Release of Spliced mRNA and Regeneration of snRNA Structures

Once the splicing reaction is complete, the mRNA must be released from the splicing complex, the snRNA interactions with the intron and exons must be melted, and snRNA structures important for spliceosome assembly must be regenerated, if to be used for another round of splicing. Release of spliced mRNA requires PRP22 protein (*74*), possibly to disrupt interactions between the spliced exons and U5. U2 and U6 remain associated with the intron product as indicated by cross-linking; these interactions must also be disrupted to release the snRNA from the intron. In addition, the base-pairs between U2 and U6 must be disrupted to release these snRNAs from each other.

Following its release, U6 must unfold and reassociate with U4 to regenerate the U4/U6 snRNP. Proteins that may be candidates for factors that function in this way have been identified by genetic approaches in yeast. Mutations in the PRP24 gene act as recessive suppressors of a U4 mutation that destabilizes the U4/U6 stem I (*66*). The PRP24 protein binds U6 snRNA, but not the U4/U6 complex, except in the mutant U4 strain. One hypothesis is that PRP24 is required to regenerate the U4/U6 snRNP and subsequently U4/U6·U5 snRNP for spliceosome assembly (*66*). It is difficult to tell from the genetic studies performed thus far whether the defect in these mutants is in regeneration of sufficient amounts of U4/U6·U5 snRNP or is manifested during spliceosome assembly or function of the tri-snRNP components once assembled.

X. Conclusions and Perspectives

The structure of snRNA appears to contribute to the function of the spliceosome in several different ways, serving to recognize substrate, position attacking groups, and possibly stabilize the association of intermediates. If the snRNAs of the assembled and activated spliceosome represent a ribozyme, then what is their relationship to other ribozymes, especially the self-splicing group-II introns with which they share mechanistic similarity (*1*, *76*)? Divalent metal ions (notably Mg^{2+}) are critically important in the function of ribozymes, both for folding and possibly catalysis (*77–79*). If the spliceosome is a metalloribozyme, then how and by what RNA elements are the catalytic metals positioned? Thus far, the demonstration of a catalytic activity related to splicing by RNA derivatives of the snRNAs in the absence of protein has not been possible. How the proteins of the spliceosome enhance the operation of its RNA elements is a matter of conjecture: despite the sequence similarity between RNA helicases and some splicing factors, no RNA helicase activity has been identified (*1–5*). Thus, the picture of the spliceosome that emerges from the available data remains blurred in several important areas. Further technical breakthroughs in purification and biochemical analysis of splicing components and in resolving events during splicing will be necessary to refine this picture and address many of the pressing mechanistic questions.

Acknowledgments

We thank R. J. Lin, Hitten Madhani, and Christine Guthrie for communicating results prior to publication. We thank Doug Black, Michelle Haynes, Scott Seiwert, and Al Zahler for making many valuable comments on the manuscript. Supported by National Institute of General Medical Sciences Grant GM40478 and a Research Career Development Award to M.A.

References

1. M. Moore, C. Query and P. Sharp, *in* "The RNA World" (R. F. Gesteland and J. F. Atkins, eds.), p. 303. CSHLab Press, Plainview, New York, 1993.
2. B. Rymond and M. Rosbash, *in* "The Molecular and Cellular Biology of the Yeast Saccharomyces" (E. W. Jones, J. R. Pringle and J. R. Broach, eds.), p. 143. CSHLab Press, Plainview, New York, 1992.
3. S. Ruby and J. Abelson, *Trends Genet.* **7**, 79 (1991).
4. C. Guthrie, *Science* **253**, 157 (1991).
5. M. R. Green, *Annu. Rev. Cell Biol.* **7**, 559 (1991).
6. J. Steitz, D. L. Black, V. Gerke, K. A. Parker, A. Kramer, D. Frendewey and W. Keller, *in*

"Structure and Function of Major and Minor Small Nuclear Ribonucleoprotein Particles" (M. L. Birnsteil, ed.), p. 100. Springer-Verlag, Berlin and New York, 1988.
7. I. Mattaj, *in* "Structure and Function of Major and Minor Small Nuclear Ribonucleoprotein Particles" (M. L. Birnsteil, ed.), p. 100. Springer-Verlag, Berlin and New York, 1988.
8. R. Luhrmann, *in* "Structure and Function of Major and Minor Small Nuclear Ribonucleoprotein Particles" (M. L. Birnsteil, ed.), p. 71. Springer-Verlag, Berlin and New York, 1988.
9. M. Moore and P. Sharp, *Nature* **365,** 364 (1993).
10. S. Burgess and C. Guthrie, *Cell* **73,** 1377 (1993).
11. S. L. Yean and R. J. Lin, *MCBiol* **11,** 5571 (1991).
12. T. Nilsen, *Annu. Rev. Microbiol.* **47,** 413 (1993).
13. C. Guthrie and B. Patterson, *ARGen* **22,** 387 (1988).
14. J. Puglisi, J. Wyatt and I. Tinoco, *JMB* **214,** 437 (1990).
15. J. Wyatt, E. Sontheimer and J. Steitz, *Genes Dev.* **6,** 2542 (1992).
16. E. Sontheimer and J. Steitz, *Science* **262,** 1989 (1993).
17. C. Lesser and C. Guthrie, *Science* **262,** 1982 (1993).
18. S. Kandels-Lewis and B. Seraphin, *Science* **262,** 2035 (1993).
19. Y. Zhuang and A. M. Weiner *Cell* **46,** 827 (1986).
20. P. Siliciano and C. Guthrie, *Genes Dev.* **2,** 1258 (1988).
21. B. Seraphin, L. Kretzner and M. Rosbash, *EMBO J.* **7,** 2533 (1988).
22. R. Reed, *Genes Dev.* **3,** 2113 (1989).
23. C. Reich, R. VanHoy, G. Porter and J. A. Wise, *Cell* **69,** 1159 (1992).
24. B. Seraphin and S. Kandels-Lewis, *Science* **262,** 2035 (1993).
25. J. Bruzik and J. Steitz, *Cell* **62,** 889 (1990).
26. S. Seiwert and J. Steitz, *MCBiol* **13,** 3135 (1993).
27. K. Hall and M. Konarska, *PNAS* **89,** 10969 (1992).
28. B. Konforti, M. Koziolkiewicz and M. Konarska, *Cell* **75,** 863 (1993).
29. M. Area and A. H. Igel, *Genes Dev.* **4,** 2132 (1990).
30. S.-E. Behrens, K. Tyc, B. Kastner, J. Reichelt and R. Luhrmann, *MCBiol* **13,** 307 (1993).
31. M. Ares and A. H. Igel, *in* "The Molecular Biology of RNA" (T. Cech, ed.) p. 13. Alan R. Liss, New York, 1989.
32. J. Wu and J. Manley, *MCBiol* **12,** 5464 (1992).
33. H. Madhani and C. Guthrie, *Cell* **71,** 803 (1992).
34. D. McPheeters and J. Abelson, *Cell* **71,** 819 (1992).
35. M. Zavanelli and M. Ares, *Genes Dev.* **5,** 2521 (1991).
36. M. Zavanelli, J. Britton, A. H. Igel and M. Ares, *MCBiol* **14,** 1689 (1994).
37. S. Barabino, B. Sproat and A. Lamond, *NARes* **20,** 4457 (1992).
38. B. Datta and A. M. Weiner, *JBC* **267,** 4497 (1992).
39. R. Parker, P. Siliciano and C. Guthrie, *Cell* **49,** 229 (1987).
40. Y. Zhuang and A. M. Weiner, *Genes Dev.* **3,** 1545 (1989).
41. J. Wu and J. Manley, *Genes Dev.* **3,** 1553 (1989).
42. L. Miraglia, S. Seiwert, A. H. Igel and M. Ares, *PNAS* **88,** 7061 (1991).
43. K. Nelson and M. Green, *Genes Dev.* **3,** 1562 (1989).
44. D. Wasserman and J. Steitz, *Science* **257,** 1918 (1992).
45. K. Hartmuth and A. Barta, *MCBiol* **8,** 2011 (1988).
46. J. Noble, Z.-Q. Pan, C. Prives and J. Manley, *Cell* **50,** 227 (1987).
47. C. Query, M. Moore and P. Sharp, *Genes Dev.* **8,** 587 (1994).
48. P. Bringmann, B. Appel, J. Rinke, R. Reuter and H. Theissenn *et al.*, *EMBO J.* **3,** 1357 (1984).
49. C. Hashimoto and J. Steitz, *NARes* **12,** 3283 (1984).

50. J. Rinke, B. Appel, M. Digweed and R. Luhrmann, *JMB* **185,** 721 (1985).
51. D. Brow and C. Guthrie, *Nature* **334,** 213 (1988).
52. D. L. Black and A. Pinto, *MCBiol* **9,** 3350 (1989).
53. D. Frank, H. Roiha and C. Guthrie, *MCBiol* **14,** 2180 (1994).
54. E. Sontheimer and J. Steitz, *MCBiol* **12,** 734 (1992).
55. T. Hausner, L. Giglio and A. M. Weiner, *Genes Dev.* **4,** 2146 (1990).
56. B. Datta and A. M. Weiner, *Nature* **352,** 821 (1991).
57. J. Wu and J. Manley, *Nature* **352,** 818 (1991).
58. P. Fabrizio, D. McPheeters and J. Abelson, *Genes Dev.* **3,** 2137 (1989).
59. H. Madhani, R. Bordonne and C. Guthrie, *Genes Dev.* **4,** 2264 (1990).
60. E. Shuster and C. Guthrie, *Nature* **345,** 270 (1990).
61. D. Wasserman and J. Steitz, *PNAS* **90,** 7139 (1993).
62. H. Sawa and Y. Shimura, *Genes Dev.* **6,** 244 (1992).
63. H. Sawa and J. Abelson, *PNAS* **89,** 11269 (1992).
64. T. Wolff and A. Bindereif, *Genes Dev.* **7,** 1377 (1993).
65. D. Fortner, R. Troy and D. Brow, *Genes Dev.* **8,** 221 (1994).
66. K. Shannon and C. Guthrie, *Genes Dev.* **5,** 773 (1991).
67. B. Datta and A. M. Weiner, *MCBiol* **13,** 5377 (1993).
68. T. Wolff, R. Menssen, J. Hammel and A. Bindereif, *PNAS* **91,** 903 (1994).
69. A. Newman and C. Norman, *Cell* **65,** 115 (1991).
70. A. Newman and C. Norman, *Cell* **68,** 743 (1992).
71. P. Fabrizio and J. Abelson, *Science* **250,** 404 (1990).
72. H. Madhani and C. Guthrie, *Genes Dev.* **8,** 1071 (1994).
73. R. Parker and P. Siliciano, *Nature* **361,** 660 (1993).
74. M. Company, J. Arenas and J. Abelson, *Nature* **349,** 487 (1991).
75. R. Reddy and H. Busch, *in* "Structure and Function of Major and Minor Small Nuclear Ribonucleoprotein Particles" (M. L. Birnsteil, ed.), p. 1. Springer-Verlag, Berlin and New York, 1988.
76. A. M. Weiner, *Cell* **72,** 161 (1993).
77. M. Yarus, *FASEB J.* **7,** 31 (1993).
78. A. Pyle, *Science* **261,** 709 (1993).
79. T. Steitz and J. Steitz, *PNAS* **90,** 6498 (1993).
80. K. Maschhoff and R. Padgett, *NARes* **20,** 1949 (1992).
81. P. Sharp, *Cell* **77,** 805 (1994).
82. T. Nielsen, *Cell* **78,** 1 (1994).
83. H. Madhani and C. Guthrie, *Annu. Rev. Genet.* in press (1994).