Evidence for a base-pairing interaction between U6 small nuclear RNA and the 5' splice site during the splicing reaction in yeast

HITOSHI SAWA and JOHN ABELSON*

California Institute of Technology, Division of Biology, 147-75, Pasadena, CA 91125

Contributed by John Abelson, August 21, 1992

ABSTRACT U6 small nuclear RNA (snRNA) is an essential factor in mRNA splicing. On the basis of the high conservation of its sequence, it has been proposed that U6 snRNA may function catalytically during the splicing reaction. If this is the case, it is likely that U6 snRNA interacts with the splice sites in the spliceosome to catalyze the reaction. We have used UV crosslinking to analyze the interactions of U6 snRNA with the splice substrates during the yeast splicing reaction. Crosslinked products in which the central region of U6 snRNA was joined to the 5' splice site region of mRNA precursor and lariat intermediate were identified. The crosslinking sites were precisely located in one of these products. The results suggest a possible base-pairing interaction between U6 snRNA and the 5' splice site of the mRNA precursor.

In nuclear pre-mRNA splicing, a number of factors including five small nuclear ribonucleoproteins (snRNPs) assemble on the mRNA precursor (pre-mRNA) in an ordered pathway to form a large RNP complex called the spliceosome (1–3). In this complex, the intron is excised in a two-step reaction. The first step involves the cleavage at the 5' splice site and the formation of a lariat structure. In the second step, the 3' splice site is cleaved and concomitantly the 5' exon is ligated to the 3' exon. This two-step reaction is fundamentally similar to the self-splicing pathway of group II introns, which is known to be catalyzed by the intron itself (4, 5). This has led to the hypothesis that the mechanisms of pre-mRNA splicing and self-splicing are related and that pre-mRNA splicing is RNA-catalyzed (6–8). Although no conserved secondary structures have been reported in pre-mRNA introns, small nuclear RNAs (snRNAs) are known to participate in the splicing reaction as trans-acting components. Therefore, it has been argued that snRNAs might interact with each other and/or with pre-mRNA to form the structure that catalyzes the splicing reaction.

Several RNA–RNP interactions are known to be involved in mRNA splicing. U1 snRNA interacts with the 5' splice site by base pairing at an early stage in the spliceosome assembly pathway (9–11). This interaction is apparently destabilized after spliceosome formation (12–14). Recently U1 snRNA was also shown to base pair with the 3' splice site in Schizosaccharomyces pombe (15). U2 snRNA base pairs with the branch point (16–18). A study of cryptic splice site activation in Saccharomyces cerevisiae suggests that U5 snRNA interacts with the exon sequences at the 5' and 3' splice sites (19). U4 and U6 snRNAs base pair with each other in the U4/U6 snRNP (20–22). This base-pairing interaction is destabilized after the entry of this snRNP into the spliceosome prior to the first-step reaction (23–25). In the mammalian system, base pairing between the 5'-end region of U2 snRNA and the 3'-end region of U6 snRNA has also been demonstrated (26, 27).

A good deal of circumstantial evidence links U6 to the catalytic center of the spliceosome. It is the most highly conserved of the snRNAs (60% sequence identity between yeast and mammal) (28). Mutations in highly conserved sequences in U6 block the first step of the second step of the splicing reaction both in vitro and in vivo (29, 30). There is also the intriguing observation of introns in the U6 gene of certain fungi, suggesting an insertion of the substrate into the catalyst (31–33). If U6 is a part of the catalytic center, it should be in close proximity to the intron. Sawa and Shimura (34) demonstrated an interaction between U6 and the 5' splice site region in the mammalian system. For technical reasons, the specificity of the crosslink could not be precisely determined. In this report we have used the same technique to identify crosslinked products between U6 and the intron in the yeast system. The location of the crosslinks was mapped precisely in both molecules. The results indicate a potential for a base-pairing interaction between U6 snRNA and a conserved sequence near the 5' splice site of the intron. The formation of such a structure could be an essential prelude to catalysis, since it would position essential conserved residues in U6 RNA in close proximity to the 5' splice site.

MATERIALS AND METHODS

UV-Crosslinking Experiments. Preparation of yeast whole-cell extract from SS330 strains (35) and in vitro synthesis of actin pre-mRNA (36) were carried out as described. Standard splicing reactions were performed as described (24) in 40-μl reaction mixtures or in 1- to 2-ml reaction mixtures for large-scale crosslinking experiments. After incubation, the reaction mixtures were diluted 1:5 to 1:20 with buffer E [12 mM Hepes-KOH, pH 7.0/0.3 mM KCl/3 mM MgCl2/0.12 mM EDTA/12% (vol/vol) glycerol] to minimize nonspecific crosslinking. The diluted reaction mixtures were then UV-irradiated (wavelength, 254 nm; GTE, G15T8) for 2 min on a microtiter plate on ice at a distance of 3 cm from the lamp (≈4000 μW/cm²). The RNAs were recovered from the irradiated samples as described (34). Annealing of the crosslinked products with a biotinylated oligonucleotide and binding to streptavidin-agarose (GIBCO/BRL) were carried out as described (34). The oligonucleotide [U6α, complementary to nucleotides (nt) 84–99 of U6 snRNA] was biotinylated at its 5' end with Biotin-ON phosphoramidite (Clontech) during synthesis. After binding, agarose beads were washed six times with 5–10 vol of buffer G (0.6 M KCl/20 mM Hepes-KOH, pH 7.0/0.2 mM EDTA), and bound RNAs were eluted by incubation with buffer D (20 mM Hepes-KOH, pH 7.0/50 mM KCl/0.2 mM EDTA/20% glycerol) at 65°C for 3 min. Selected RNAs were analyzed by electrophoresis on a 5% polyacrylamide gel containing 8 M urea. For the large-scale crosslinking experiments, the bands corresponding to the crosslinked products were excised and RNAs were eluted

Abbreviations: snRNA, small nuclear RNA; nt, nucleotide(s); snRNP, small nuclear ribonucleoprotein.
*To whom reprint requests should be addressed.
from the gel in RNA elution buffer (0.5 M NaOAc, pH 5.3/1 mM EDTA/0.1% SDS).

Analysis of Crosslinked Products. RNase H-digestion experiments were performed as described (34). Primer-extension reaction mixtures with avian myeloblastosis virus reverse transcriptase (Promega) contained about 0.01 fmol of RNA and 40 fmol of primer and experiments were performed as described (37). After the reaction, the RNA template was hydrolyzed with 0.1 M NaOH at 65°C for 1 hr and the DNA extension products were separated electrophoretically on an 8% polyacrylamide gel containing 8 M urea. The radioactivity in the gel was detected using a PhosphorImager (Molecular Dynamics, Sunnyvale, CA).

RESULTS

Crosslinked Products of U6 snRNA. To analyze the interaction of U6 snRNA and pre-mRNA during the splicing reaction, the in vitro splicing reaction was carried out with a yeast extract and 32P-labeled actin pre-mRNA. Then the reaction mixture was UV-irradiated. After irradiation, the RNAs were isolated by proteinase treatment and phenol extraction. Crosslinked products of U6 snRNAs were extensively enriched by an affinity selection using a biotinylated oligonucleotide complementary to U6 snRNA and streptavidin-agarose. The bound RNAs were eluted and analyzed by gel electrophoresis (Fig. 1A).

In the reaction mixture incubated for 3 min, two crosslinked products (designated U6-preI and U6-preII) were observed in addition to pre-mRNA (Fig. 1A, lane 1). These products had quite similar electrophoretic mobilities and were difficult to reproducibly separate (data not shown). At 3 min, the first step of the reaction had not occurred, indicating that these products must be U6 crosslinked to pre-mRNA. The amounts of U6-preI and U6-preII decreased after a 10-min incubation (lanes 4 and 5). After the first step of the reaction had occurred (5 min), an additional U6 crosslink was observed (designated U6-2/3). This product was cleaved by RNase H when incubated with oligonucleotides specific for the intron (Fig. 1B, lane 3) or the 3' exon (lane 4) but was not cleaved with an oligonucleotide complementary to the 5' exon (lane 2). This suggests that U6-2/3 contains a crosslink between U6 snRNA and the lariat intermediate. To confirm this, a primer-extension experiment was carried out with uncrosslinked lariat intermediate or U6-2/3 purified by electrophoresis (Fig. 1C). The oligodeoxyribonucleotide primer in this reaction was complementary to the 3' splice site region. With both RNAs, extension by reverse transcriptase was blocked at the branch point, resulting in a 51-nt extension product. This clearly shows that U6-2/3 is a lariat structure. Because this product was cleaved by RNase H with the 3' exon oligonucleotide (Fig. 1B, lane 4), U6-2/3 contains a crosslink between U6 snRNA and the lariat intermediate, and not the lariat intron. These U6-crosslinked products could not be selected with biotinylated oligonucleotides complementary to U1, U2, U4, or U5 snRNAs (data not shown), suggesting that the products do not contain other snRNAs. Failure to detect crosslinks with more than one snRNA may be due to the low efficiency of crosslinking in our conditions. Less than 0.01% of the starting pre-mRNA was found in the U6-crosslinked products.

The fact that the crosslinked products described above are not selected when the purification is carried out in the absence of a complementary oligonucleotide (data not shown) confirms that they contain U6 sequences. These products were not formed when the splicing reaction was carried out in the absence of ATP (Fig. 1A, lane 6), indicating that formation of the spliceosome is required.

Mapping of the Crosslinking Sites by RNase H Digestion. The location of the crosslinks in U6-preI and U6-preII was determined by RNase H mapping. As mentioned above, the separation of these crosslinked products was not reproducible. This time we used a mixture of the products purified by electrophoresis. This mixture probably contained about equal amounts of U6-preI and U6-preII, because we always detected similar amounts of them when they were distinguishable on the gel (Fig. 1A and data not shown). The mixture of U6-preI and U6-preII was digested by RNase H directed by oligonucleotides complementary to a portion of the 5' exon or intron near the 5' splice site (oligonucleotides 5'E or Int). The digestion products were separated by electrophoresis using the digestion products of uncrosslinked pre-mRNA as markers (Fig. 2A). When oligonucleotide 5'E was used, a small fragment corresponding to the 5' portion of pre-mRNA relative to the cleavage site (5' fragment of pre-mRNA) was generated from the mixture of U6-preI and U6-preII and from uncrosslinked pre-mRNA. This indicates that the crosslinking sites are not in this fragment. The digestion product containing the 3' fragment of pre-mRNA was detected as a slowly migrating RNA (lane 5'E). There-
Digestion of snRNA by assuming oligonucleotide with oligonucleotide U6dl, the mobility of the cleavage product was similar to that of lariat intermediate (Fig. 2C). Therefore, the crosslinking site in U6-2/3 is also in the U6d1 region.

**Determination of the Crosslinking Sites by Primer-Extension Analyses.** To localize the crosslinking sites more precisely, we performed primer-extension experiments. In these experiments, reverse transcriptase was expected to stop 1 nt before the crosslink as in stops caused by the 2'-3' phosphodiester bond at the branch site (see Fig. 1C) and by chemically modified nucleotides (39). Mapping of crosslinks in U6 snRNA was performed using the 5'-end-labeled oligonucleotide U6a as a primer (Fig. 3A). With U6-2/3, five nearly consecutive strong stops were observed. These signals indicate that reverse transcriptase extension was blocked at positions 39 and 41–44 on U6 snRNA (indicated by small dots in the figure). A signal corresponding to a block at A40 was not detected. The primes were cleaved by RNase H. Consequently, the crosslinking sites of U6-preI and U6-preII should be in the U6d1 region. This is also the case with U6-2/3. When purified U6-2/3 was cleaved in response to oligonucleotide U6d1, the mobility of the cleavage product was similar to that of lariat intermediate (Fig. 2C). Therefore, the crosslinking site in U6-2/3 is also in the U6d1 region.

**Fig. 2.** RNase H-digestion experiments of the crosslinked products. (A) Pre-mRNA or the mixture of U6-preI and U6-preII purified after large-scale crosslinking experiment was digested by RNase H with oligonucleotides 5'E (complementary to nt 32 to -17 upstream of the 5' splice site) or Int (complementary to nt +41 to +56 downstream of the 5' splice site) or without oligonucleotides (lane no oligo). The digestion products were analyzed as in Fig. 1A. (B and C) Mixtures of U6-preI and U6-preII (B) or U6-2/3 (C) were digested by RNase H with U6a (complementary to nt 84–99 of U6 snRNA) or U6d1 (complementary to nt 28–54 of U6 snRNA) oligonucleotides or without oligonucleotides (lane no oligo). Digestion products were analyzed as in Fig. 1A.

Therefore, the crosslinking sites in U6-preI and U6-preII are downstream of the region complementary to oligonucleotide 5'E. In contrast, RNase H digestion directed by oligonucleotide Int efficiently produced the 3' fragment of pre-mRNA, indicating that the crosslinking sites are upstream of the Int region. Therefore, the crosslinking sites in U6-preI and U6-preII are between the 5'E and Int regions. These results also confirm that the U6-preI and U6-preII species contain intact pre-mRNA and not the intermediates or products of the splicing reaction, since RNase H digestion of the crosslinked products at either side of the crosslinking sites gave rise to linear RNA fragments with the same mobilities as those obtained by RNase H digestion of intact uncrosslinked pre-mRNA.

To analyze the crosslinking sites on U6 snRNA, the mixture of U6-preI and U6-preII was cleaved by RNase H with two oligonucleotides complementary to U6 snRNA (Fig. 2B). RNase H cleavage directed by oligonucleotide U6a, complementary to a region near the 3' end of U6 snRNA, led to a product with slightly increased mobility relative to the uncleaved crosslinked product. In contrast, digestion directed by oligonucleotide U6d1, which is complementary to the central region of U6 snRNA (see Fig. 4), gave rise to a product with the same mobility as uncrosslinked pre-mRNA. Since this cleavage product contains intact pre-mRNA, this result indicates that most of U6 snRNA was removed by digestion with oligonucleotide U6d1. This can be explained by assuming that oligonucleotide U6d1 hybridized with U6 snRNA on both sides of the crosslinking site and that both

**Fig. 3.** High-resolution mapping of the crosslinking sites by primer-extension analysis. The crosslinked products, pre-mRNA, and lariat intermediate were purified after large-scale crosslinking experiments. U6 snRNA was isolated from the irradiated reaction mixture by proteinase treatment and affinity selection with the biotinylated U6 oligonucleotide. These RNAs were used for primer-extension analyses with end-labeled oligonucleotides U6a (A) or Int (B) as primers. For each reaction, 0.01 fmol of pre-mRNA, lariat intermediate, or U6-2/3 or 0.004 fmol of U6 preI or U6-preII was used. Marker, DNA sequencing with the U6a (A) or Int (B) primers and letters adjacent to the gel was the same as in Fig. 1C. Small dots adjacent to the RNA sequences indicate positions of reverse transcriptase stops in the U6-2/3 template. In addition to the stops at positions 39–43 in U6, we also note another band corresponding to a block at position 19 (A, lane U6-2/3). There is no evidence for this minor crosslinked product in the RNase H-mapping experiment (Fig. 2C). Further experiments will be required to determine whether this is a significant crosslink site.
absence of this stop makes it unlikely that reverse transcriptase is fortuitously blocked at multiple positions near a crosslinking site. These blocks were not observed in uncrosslinked U6 snRNA affinity-purified from an irradiated splicing reaction mixture, excluding the possibility that they are due to nonspecific damage of U6 snRNA caused by the irradiation. These results suggest that U6-2/3 consists of five crosslinked products that were not resolved by electrophoresis and that there are crosslinking sites at G^37, A^{41}, A^{54}, C^{65}, and A^{44} in U6 snRNA. Consistent with the results of RNase H-mapping experiments (Fig. 2C), these nucleotides are in the U6d1 region (see Fig. 4). With U6-prefI, similar extension products were weakly observed, indicating that the crosslinking sites are located at the same positions as in U6-2/3. Although we failed to map precisely the crosslinking sites of U6-prefI by primer-extension analyses, RNase H-digestion experiments show that the crosslinking sites are in the U6d1 region (Fig. 2B).

To map the crosslinking sites on pre-mRNA or the lariat intermediate, we used oligonucleotide Int as a primer (Fig. 3B), because the crosslinking sites of U6-prefI and U6-prefII were shown to be upstream of the Int region (Fig. 2A). With U6-2/3, five consecutive signals were observed that indicate reverse transcriptase blocks at nt +4 to +8 downstream of the 5′ splice site. These blocks did not occur when uncrosslinked pre-mRNA or lariat intermediate was used as template. The results again suggest that U6-2/3 consists of five crosslinked products and indicate that the crosslinking sites are at residues U^{144}, G^{155}, U^{166}, U^{277}, and C^{288} in the intron (the numbers indicate relative positions from the 5′ splice site). Insufficient U6-pref and -prefII crosslinked products were obtained to allow mapping.

**DISCUSSION**

UV crosslinking has been widely used to characterize interactions within and between RNA molecules (40). Generally, the UV photodadducts join two bases that are in close proximity due to tertiary structure and not base pairs in a helix. For example, in a study of intranadrash crosslinks in UV-irradiated 50S ribosomes, Mitchell et al. (41) characterized 10 crosslinks but only 1 crosslink was between nucleotides in a helical region. Nonetheless, one can create a number of crosslinks in which nucleotides in RNA known by other criteria to be base paired are joined together in UV-induced crosslinks. These include crosslinks between U4 and U6 snRNAs and between U2 and U6 snRNAs, in the trypanosome system (42), and between U1 snRNA and the 5′ splice site of the intron and between U2 and the branch point of the intron, in the mammalian system (34). In general, these intermolecular crosslinks are in extremely low yield and the chemistry by which they form is unknown. If the crosslinks are most likely to form between stacked bases, then it is possible that UV crosslinks form during the pairing or dissociation of a helix, during rare structural perturbations of the helix, or when the helix geometry is altered by interaction with a protein. The presence of a UV crosslink between two bases most likely indicates that they are at least in transient proximity to each other but it neither proves nor disproves their occurrence in a base pair.

The above discussion is relevant to the results we have observed. We have isolated a collection of crosslinked species in which a nearly contiguous set of bases in U6 RNA is crosslinked to a complementary set of bases in the intron near the 5′ splice site. Since these crosslinks require ATP to form and since one of them occurs in a splicing intermediate, it is highly likely that nt 39–44 in U6 are in close proximity to nt +4 to +8 of the intron within the spliceosome. Since these nucleotides are complementary, it seems reasonable that they do pair (Fig. 4), although the UV crosslinks might not have occurred when they were in that form.

After we had obtained these results, we were excited to learn that Wassarman and Steitz (43) had also observed a similar crosslink between U6 and the 5′ splice site region of the intron. Their work was in the mammalian system and employed psoralen-induced crosslinking, which is more apt to join base-paired regions. The crosslink they observed was between C^{37} in U6 (equivalent to C^{43} in yeast U6 RNA) and a T1 oligonucleotide in the adenovirus intron that spans the region from positions +6 to +21 of the intron. If their crosslink is, in fact, between C^{37} and the first base of this oligonucleotide, their results could be identical to ours. However, they propose a different base-pairing scheme in which the highly conserved ACAGAG sequence in U6 pairs with the intron at positions +2 to +7, a pairing 5 base pairs out of register with ours. Since their crosslinked species is between pre-mRNA and U6 and the one we have characterized is between the lariat intermediate and U6, an alternative possibility is that both pairing schemes are correct and that there is a conformational switch between the two steps of splicing. In any case, the discovery of similar crosslinks produced by different techniques in two phylogenetically divergent splicing systems provides strong indication that there is a functional association between U6 RNA and the 5′ splice site.

If this interaction includes pairing with bases +4 to +7 as both we and Wassarman and Steitz (43) propose, then it would replace at least a portion of the base pairing known to occur between U1 snRNA and the 5′ splice site, and it would also replace the base pairing between U5 snRNA and the intron in this region proposed by Wassarman and Steitz (43) from their psoralen-crosslinking results. These mutually exclusive pairings appear to be a crucial feature of the splicing process. The interaction between U1 and the 5′ splice site is an earlier step in spliceosome assembly but its function may be restricted to splice-site recognition and this interaction may be replaced in time by the U6 interaction in the fully assembled spliceosome. It is attractive to hypothesize a role for the RNA helicases, known to be essential factors in splicing (2, 3), in the catalysis of these helical interactions. If both U6 and U1 interact with the same nucleotides in the 5′ splice site then the alteration of these nucleotides could lead to complicated phenotypes. It has long been known that G^5 -> A mutation in the intron (term 'A') in actin and ribosomal protein RP51A leads to aberrant splicing in vivo (44–46). In the first step of the aberrant splicing, cutting and
lariat formation take place 5 bases upstream in A′ actin pre-mRNA and 3 bases upstream in the A′ RPS24A pre-mRNA. Compensatory mutations in U1 did not suppress these aberrant cleavages (9, 10). It is an interesting possibility that the A′ mutation that disrupts the pairing between U6 and the intron causes the catalytic center to be positioned incorrectly and, therefore, to cut incorrectly. The suppression of the A′ mutation may thus require changes in both U1 and U6. It is likely, however, that the genetic dissection of the U6–intron interaction will be difficult because individual base substitutions in the region from G39 to A′′′ of U6 have no effect either in vitro or in vivo (29, 30). A triple mutation of the bases AAC (corresponding to nt 41–43 in yeast) to UUG in Xenopus U6 RNA severely inhibited splicing in the oocyte complementation assay (47, 48). It may be that single-base changes are not sufficient to disrupt the pairing between U6 and the intron.

The interaction proposed here between U6 snRNA and the intron would position the highly conserved ACAGAG sequence (nt 47–52) very close to the 5′ splice site (Fig. 4). Mutations in this sequence have strong effects on either the first or the second step of splicing, both in vitro and in vivo (29, 30). It is an exciting possibility that the U6–intron interaction is crucial to the formation of the active conformation of the spliceosome and that Fig. 4 represents a snapshot of the RNA secondary structure in a part of that complex at some point along the reaction pathway. It is important to know which point. From our results, it is likely that the interaction exists in similar form, both before and after the first step of splicing; however, there could be important but subtle changes that occur after the first step. Temperature-sensitive mutations in the PRP genes required for pre-mRNA splicing allow us to dissociate the process into a number of stages (3, 24). We should be able to use these tools to determine the time in the assembly and reaction pathway in which RNA interactions, as determined through crosslinks, occur. Due to the low efficiency, UV crosslinking will probably not be the tool of choice for these experiments. Wyatt, Sontheimer, and Steitz (personal communication) have employed the ligation tools developed by Moore and Sharp (49) to position base analogues containing photo-induced crosslinking groups into unique positions of pre-mRNA. Using this technique they have demonstrated specific crosslinks between pre-mRNA and U5 snRNA. It is clear that this technology can now be applied more broadly to provide a comprehensive picture of the RNA interactions in the spliceosome as it is assembled and as it functions.

We thank Suzanne Horvath and staff for synthesis of oligonucleotides, Y. Shimura for help in the early stage of this project, and C. Guthrie, E. Sontheimer, D. Wassarman, and members in our lab for critical comments on the manuscript. H.S. was supported by a long-term fellowship from the Human Frontier Science Program Organization. This work was supported by National Institutes of Health Grant GM 32637.