The secondary structure of a messenger RNA precursor probed with psoralen is melted in an *in vitro* splicing reaction

Paul L. Wollenzien¹, Prabhat Goswami, John Teare¹, Jozsef Szeberenyi* and Carlos J. Goldenberg

Department of Microbiology and Immunology, University of Miami School of Medicine, Miami, FL 33101 and ¹Department of Biochemistry, St Louis University Medical School, St Louis, MO 63104, USA

Received August 17, 1987; Revised and Accepted September 25, 1987

ABSTRACT

The secondary structure of the SP6/mouse insulin precursor RNA was determined by psoralen cross-linking experiments. A series of long-range contacts occur within the left half of the pre-mRNA that contains the intervening sequence. Multiple secondary structures for the pre-mRNA exist since some of the interactions share common sites. In splicing buffer but without the splicing extract added, many of these interactions are stable up to at least 50°C. These interactions, however, are dissociated during the *in vitro* splicing reaction. This dissociation requires ATP and it occurs during the first 30 min. of the splicing reaction. Pre-mRNAs containing psoralen cross-links in different locations within the RNA molecule were purified and used as substrates for *in vitro* splicing. Psoralen cross-links at any of the double-stranded regions resulted in complete inhibition of the splicing reaction. This indicates that destabilization of the secondary structure of the SP6/mouse insulin pre-mRNA is necessary for *in vitro* splicing.

INTRODUCTION

Higher eukaryotic genes frequently are interrupted by noncoding sequences (introns) which are first transcribed as parts of the pre-mRNAs and subsequently removed by the process of RNA splicing (reviewed in 1,2). The endonucleolytic cleavage of intron-exon junctions in *in vitro* cell-free systems occurs in a multicomponent ribonucleoprotein complex (spliceosome) (3,4,5). The assembly of the pre-mRNA in the spliceosome complex must position the exons and RNA intermediates in such a way as to promote accurate cutting and ligation to prevent incorrect splicing or shuffling of exons.

The 60S spliceosome complex is formed most likely by an ordered pathway of assembly of the pre-mRNA with heterogeneous nuclear ribonucleoproteins (hnRNPs), Sm small nuclear ribonucleoproteins (snRNPs) and probably other unknown factors (6,7,3,4,-5,8).

The information that directs the accurate splicing of a mRNA precursor must reside, in part, in the precursor itself. Extensive analysis of the sequence requirements for pre-mRNA splicing have shown that the 5' and 3' splice sites are the only essential sequence elements that determine the site at which cleavage occurs (reviewed in 9).

Recognition of these splice site sequences alone, however, is clearly not sufficient to specify the selection of splice sites in a precursor containing multiple intron and/or cryptic splice sites (10,11,12). It appears that the relative proximity of the donor and acceptor splice sites and the sequence context in which they are placed are important determinants in splice site selection (12). Furthermore, little is known about how RNA structures influence RNA splicing. In this regard, Solnick (13) recently showed that the introduction of inverted sequences, that formed a stable duplex region, in two adjacent introns of a pre-mRNA caused an alternative splice that excluded the exon in the resulting loop, both in vivo and in vitro. The relationship between the structure of isolated pre-mRNA and the structure of the packaged RNA precursor (hnRNP) in the cell remains to be determined. For example, the packaging of the hnRNA into hnRNPs in vivo may alter the RNA secondary structure in a way that might influence patterns of splicing.

The experiments reported here investigate the secondary structure of a mouse insulin mRNA precursor by using a nucleic acid cross-linking reagent, psoralen. Psoralen is a heterocyclic, aromatic agent that intercalates into accessible double-stranded RNA regions and, upon long-wave (365 nanometer) ultraviolet irradiation, undergoes a photoinduced reaction forming covalent interstrand cross-links. The intercalated psoralen becomes covalently attached to RNA via a cyclobutane linkage to pyrimidine bases (predominantly uridine in RNA). A single psoralen molecule can covalently attach to one or two nucleotides producing either monoadducts or cross-links, respectively (14). The cross-linked RNA molecules that result from this treatment can then be fractionated by gel electrophoresis under denaturing conditions (15), and the site of the cross-link determined by primer-extensions with reverse transcriptase.

We show here, using an in vitro splicing system, that the secondary structure of the mouse insulin pre-mRNA is perturbed upon incubation with the splicing extract. The changes follow a pattern that indicates a melting of double-stranded regions within the pre-mRNA before the endonucleolytic cleavage occurs. The melting of the pre-mRNA secondary structure requires ATP. The effect on in vitro splicing of psoralen cross-links located at various double-stranded regions of the pre-mRNA was assessed. These substrates were inactive for the splicing reaction.

MATERIALS AND METHODS

SP6 Transcriptions and in vitro splicing

The SP6/mouse insulin pre-mRNA labeled with α [^{32}P]nucleotides was prepared as described (5,16). Splicing reactions were carried out with a partially purified nuclear extract isolated from MOPC-315 mouse myeloma cells (17,18) or a nuclear extract isolated from HeLa cells (19). The experiments presented here were performed with a

HeLa nuclear extract. Both splicing extracts, however, melted the secondary structure of the SP6/mouse insulin pre-mRNA as assayed by photochemical cross-linking with amino methyltrimethyl psoralen (AMT).

AMT cross-linking

The ^{32}P -labeled pre-mRNA (12 ng, 4×10^5 cpm in 25 μl reactions) was incubated at 30°C in splicing buffer (12 mM Hepes (pH 7.9), 60 mM KCl 3.2 mM MgCl_2 1 mM DTT, 12% glycerol, 2.1% (w/v) polyvinyl alcohol, 1000 U/ml RNasin, 0.5 mM ATP, 20 mM creatine phosphate) without the splicing extract or plus 12 μl (60 μg protein) of splicing extract. The reactions were chilled on ice and AMT (0.02 μg per 12 ng of transcript plus 240 ng of tRNA) was added for an additional 10 min. This optimal ratio of AMT per pre-mRNA transcript was determined empirically. The intensity of the bands migrating above the transcript increased nearly linearly with increasing amounts of AMT below the optimal ratio (data not shown). The intensity of the cross-link RNAs did not increase linearly with AMT concentrations above the optimal and there was a significant background, not resolved into distinct bands (data not shown). The reactions were irradiated at 365 nm for 7 min in a high intensity mercury light irradiator as described (20). The samples were then digested at 37°C for 30 min. with Proteinase K (1 mg/ml) in the presence of 0.5% SDS and 20 mM EDTA. The nucleic acids were extracted with phenol-chloroform-isomyl alcohol (50:50:1) and the aqueous phase precipitated with 2.5 volumes of ethanol.

For monoaddition experiments, RNA (4 μg , 4.5×10^6 cpm) in splicing buffer was incubated with 10 μg of AMT at 4° for 15 min. under N_2 atmosphere and then irradiated for 60 min. with 390 nm light (3 mwatts/cm 2). The sample was ethanol precipitated, dried briefly, and redissolved in water before further use. Cross-linking on pre-mRNAs (150 ng) containing monoadducts was preformed in 50 μl of splicing buffer as described above.

For gel electrophoresis, the RNAs were resuspended in loading buffer (80% formamide, 90 mM Tris-HCl (pH 8.3), 90 mM Boric Acid and 2.5 mM EDTA), incubated at 50°C for 3 min. and electrophoresed on 8 M urea - 5% polyacrylamide gels at 8 watts for 12 hrs.

Primer-extension Analysis

The RNAs were eluted from the gel as described (5) and used as templates for reverse transcriptase analysis. The individual RNA fractions (approximately 4,000 cpm) were resuspended in water and dried with two pmoles of a selected 5' ^{32}P -end labeled synthetic oligonucleotide (see Fig. 4A). The nucleic acid were redissolved in 1 μl of 80% formamide, 0.4 M NaCl, 40 mM Pipes pH 6.8, 1 mM EDTA, and hybridized 15 min. at 45°C . The DNA-RNA hybrids were extended for 1 hr at 42°C with 7 units of reverse transcriptase in 20 μl reactions containing 40 mM KCl, 6 mM MgCl_2 , 50 mM Tris-HCl

pH 8.3, 10 mM DTT, 5 units RNasin and 0.5 mM of each deoxynucleotide. Purification of the cDNAs before electrophoresis was as described (21). Electrophoresis was in a 8 M urea - 12% polyacrylamide gel at 90 watts for 90 min. or 270 min.

For each of the utilized synthetic primers the SP6/mouse insulin pre-mRNA was sequenced with reverse transcriptase as above in the presence of deoxy to di-deoxy-nucleotides at a ratio of 10:1.

RESULTS

The pattern of AMT cross-linking in the pre-mRNA

We have used the psoralen cross-linking technique to monitor the internal structure of the pre-mRNA during the splicing reaction. The secondary structure of the SP6/mouse insulin pre-mRNA incubated under splicing conditions but in the absence of

Figure 1. Fractionation of the AMT cross-linked SP6/mouse insulin pre-mRNA incubated in the absence of the nuclear extract. A: ^{32}P -labeled SP6/mouse insulin pre-mRNA was incubated at 30°C for 30 min. in splicing buffer (see Materials and Methods) in the absence of the nuclear extract. The RNAs were then AMT cross-linked and electrophoresed on a 5% polyacrylamide gel. The twelve RNAs fractions that were reproducibly observed were eluted from the gel (5). The numbers on the right indicate the RNA fractions. 5T indicates a triplet and 8D a doublet. B: Equal aliquots of the twelve bands were then applied to a second denaturing gel in order to check for purity of the individual RNA fractions (lanes 1-12). An aliquot of the unfractionated psoralen cross-linked pre-mRNA was also subjected to electrophoresis (unnumbered lane). The band co-migrating with fraction 1 (lane 1), is present in all of the bands (lanes 2-12) and it represents a small amount of photo-reversion of the AMT reaction.

the nuclear extract was investigated by AMT cross-linking. The AMT cross-linked molecules have new identifiable bands of reduced electrophoretic mobility when electrophoresed on 8 M urea -5% polyacrylamide gels (Fig. 1A). The cross-linking conditions we have used (see Materials and Methods), have been optimized to produce a distinct pattern of cross-links. For these conditions, the number of AMT's incorporated per pre-mRNA was determined by using [^3H]-AMT; Table 1 shows that approximately 1.3 to 1.5 AMT's are covalently attached per pre-mRNA molecule. Two experiments indicate that the new electrophoretic species do not arise from intermolecular cross-linking. First, performing the cross-linking under conditions of higher RNA concentrations does not increase the intensity of the bands migrating above the non-crosslinked transcript. Second, under non-denaturing conditions the AMT cross-linked RNAs migrated as a

Figure 2. AMT cross-linking in pre-mRNA at a series of temperatures. ^{32}P -labeled pre-mRNA was reacted with AMT and 390 nm light to form monoadducts and then after the excess of AMT was removed (see Materials and Methods). The RNA was then cross-linked at a series of temperatures. Lane C: RNA that was cross-linked directly with 365 nm light (see Materials and Methods). Lane M: RNA with AMT monoadducts. The remaining lanes contain the monoadduct RNA incubated and irradiated with 365 nm light (see Materials and Methods) at the indicated temperatures. The numbers on the left indicate the RNA fractions as in Fig. 1.

single band with the mobility of the non-crosslinked RNAs on either polyacrylamide or agarose gels (data not shown).

For purified ribosomal RNA it was previously estimated (22) that the fraction of psoralen adducts are cross-linked about 1 in 12; therefore, statistically the pre-mRNA molecules should not have more than one cross-link each, but they may have more than one psoralen monoadduct. Furthermore, in the experiment shown in Fig. 2, we have used 390 nm light to exclusively produce AMT monoadducts on the SP6/mouse insulin pre-mRNA. Those RNA molecules have an electrophoretic mobility identical to the untreated RNA (Fig. 2, lane M). We can conclude, therefore, that the gel system is not separating on the basis of the monoaddition sites but rather on the basis of the cross-linking sites in each RNA molecule.

To investigate the thermal stability of the interactions that are the sites of the cross-links, the pre-mRNA containing AMT monoadducts was cross-linked at different temperatures (Fig. 2). The pre-mRNA was incubated with a high concentration of AMT, irradiated with 390 nm light and the excess of AMT removed by ethanol precipitation (see Materials and Methods). This treatment incorporated 0.9 AMT/pre-mRNA as determined by ^3H -AMT. The pre-mRNA containing the AMT monoadducts was then cross-linked at a series of temperatures (Fig. 2). Two conclusions can be reached from this experiment. First, the overall pattern of cross-linking at 4°C in directly cross-linked molecules (Fig. 2, lane C) and in molecules that contain monoadducts before they were cross-linked (Fig. 2, lane 4), remains the same but the intensity of some of the bands is changed. This similarity occurs in spite of the fact that in the monoaddition experiments, the molecules were incubated with a high concentration of AMT yielding presumably a larger number of intercalation sites occupied. Moreover, the pre-mRNA containing monoadducts have gone through a cycle of ethanol precipitation, redissolution in water, and renaturation before the cross-linking (see Materials and Methods). The pattern of bands on the gel shown in Figure 2 (lane C) is somewhat different than that shown in Figure 1, but this is due to slightly different denaturation conditions during electrophoresis, since analysis of the cross-linking pattern by reverse transcriptions (see Fig. 3) indicated that the same sites were cross-linked in the samples shown in Figure 2, lane C and in Figure 1A (data not shown). Second, the melting experiment (Figure 2), indicates that the pattern of cross-linking at 30°C is nearly identical to the pattern at 4°C. Interestingly, there are some bands that increase in intensity as the temperature is increased to 50°C; therefore we have not been able to analyze the melting of the different double-stranded regions according to a simple helix-coil model.

AMT cross-linking sites determined by reverse transcriptions.

To determine the sites of AMT cross-linking at the nucleotide level, the twelve most prominent and reproducible bands were excised, eluted (Fig. 1A), and the different

Figure 3. Reverse transcriptions of AMT cross-linked pre-mRNAs incubated in the absence or the presence of a splicing extract.

A) The SP6/mouse insulin pre-mRNA was incubated in splicing buffer (see Materials and Methods) in the absence of a nuclear extract, AMT cross-linked and the different fractions purified as in Figure 1B. The different fractions were used as templates for reverse transcriptions (see Materials and Methods). The primer used (primer A in Fig. 4A) was complementary to nucleotides 372 to 392 of the pre-mRNA. Lane M: ³²P-labeled markers of PBR322 digested with HpaII. Lane T: Pre-mRNA not treated with AMT was extended with reverse transcriptase and deoxynucleotides only. Lanes A, G, C and U: The uncross-linked pre-mRNA sequenced with the four deoxy and each of the dideoxynucleotides (Materials and Methods). The nucleotide sequence of this region of the molecule is shown on the right. Lane X: The pre-mRNA was incubated in splicing buffer minus the nuclear extract, subjected to AMT cross-linking and the total unfractionated RNA was used as a template for reverse transcriptions. Lanes 1 to 11: Indicate the purified AMT cross-linked fraction 1 to 11 (see Fig. 1B) used as templates for reverse transcriptions. AMT monoaddition sites are indicated in the sequence by a dot above the nucleotide which correspond to reverse transcription stops present in fraction 1 (lane 1) but absent in the pre-mRNA extended with deoxynucleotides (lane T). The strong reverse transcription stops detected in the fractions 1 to 11 are indicated in the sequence shown on the right by numbers corresponding to the fraction number in which they occur, underneath the corresponding nucleotide.

RSJ indicate the right splice site/junction.

B) Lanes 10' and 20': Splicing reactions were incubated at 30°C for 10 or 20 min. and treated with AMT (see Materials and Methods). The total RNAs from those reactions were then used as templates for reverse transcriptions. The primer used (primer F in Fig. 4A) was complementary to nucleotides 212 to 232 of the pre-mRNAs.

Lanes A, G, C, U: SP6/mouse insulin pre-mRNA sequenced in the presence of dideoxynucleotides (Materials and Methods). Lane T: Uncross-linked pre-mRNA extended with reverse transcriptase and deoxynucleotides only. The nucleotide sequence of this region of the molecule is shown on the right. AMT monoaddition sites are indicated by dots above the nucleotide where there is a reverse transcription stop in the 10' and 20' samples.

RNA fractions were used as templates for reverse transcription reactions. The purity of the fractionated cross-linked molecules is shown in Figure 1B. A band co-migrating with fraction 1 was present in all of the fractions (Fig. 1B, lanes 1 to 12); it represents

Figure 4. Summary of reverse transcriptions of the AMT cross-linked SP6/mouse insulin pre-mRNAs incubated in the absence of a splicing extract.

A) The AMT cross-linked eleven fractions (F1 to F11) purified as in Fig. 1B were reverse transcribed with six different oligonucleotide primers. The additional oligonucleotides are complementary to the following regions of the pre-mRNA: B, 102 to

118; C, 200 to 216; D, 137 to 153; and E, 396 to 412. Each primer is represented by a black box (A through F). The arrows indicate strong reverse transcriptase stops. Fraction 1 (F1) show the AMT monoaddition sites. The numbers in F1 represent nucleotides (1 = the initiation site) in the SP6/pre-mRNA. The line labeled 10' shows a splicing reaction in which the total RNA was reverse transcribed. The AMT monoaddition sites are indicated by asterisks in F1 and 10'.

B) The data in panel A was analyzed for the location of sequence complementarities that would generate the cross-linking pattern in each fraction. The complementarities are shown connected by heavy lines. In two fractions (F8 and 10) we cannot find complementary sites; it is possible that the cross links occur in tertiary structural sites. RSJ and LSJ indicate the right and left splice junction in the pre-mRNA.

a small amount of photoreversal of the AMT cross-linking reaction which occurs during purification.

Psoralen adducts are barriers to reverse transcription (23), the reverse transcriptase usually stops one nucleotide before the actual site of addition (Ericson and Wollenzien, unpublished results). Therefore, it is possible to examine the cDNAs made from each fractionated RNA template and conclude which nucleotides are involved either in cross-links or in monoadducts. In our system, since the fractionation of the RNA templates is based on the location of the cross-links, it was expected to detect reverse transcription stops at different specific nucleotides in the different fractions.

Figure 3A shows the cDNAs from total unfractionated cross-linked RNA (lane X), the fractionated RNA templates (lanes 1 to 11), and control templates either without (lane T) or with dideoxynucleotides (lanes A,G,C,U). The reverse transcriptions were initiated from the primer "A" complementary to a sequence at the right end (residues 372 to 392) of the SP6/mouse insulin pre-mRNA (total length 412 nucleotides see Fig. 4). The fraction 12 has not been included in this and subsequent figures since specific stops were not identified for it. From the autoradiogram shown in Figure 3A and from a longer electrophoresis of the same samples (data not shown), the nucleotides involved in the right parts of the cross-links in each of the RNA fractions can be determined. These sites are summarized on the right of the gel in Figure 3A. Note that for most fractions there are multiple stops, indicating that each RNA fraction must still contain multiple cross-linked species. To determine the location of the left parts of the cross-links in each fraction, additional primers were used to initiate reverse transcription just to the left of the original stops (data not shown). For fraction 2, only one strong reverse transcription stop was initially determined at nucleotide 154 (see Fig. 3A, lane 2 and Fig. 4A); however, a second strong stop at nucleotide 111 (see Fig. 4A) was detected when the cross-link in the RNA was photo-reversed with 254 nm light before the reverse transcriptions (data not shown).

For each RNA fraction, a series of sites were determined in this way that were candidates for being the nucleotides involved in the cross-links. This data is summa-

Figure 5. AMT cross-linked RNAs after the *in vitro* splicing reaction.

A) ^{32}P -labeled SP6/mouse insulin RNA precursor was synthesized, spliced *in vitro* (see Materials and Methods) and then fractionated by electrophoresis on a 8M urea - 6% polyacrylamide gel. The reactions were terminated at the times indicated: 15 min. (lane 2), 30 min. (lane 3), 120 min. (lane 4). Lane 1: pre-mRNA incubated at 30°C for 120 min. in the absence of nuclear extract.

B) The splicing reactions were terminated at the indicated times and then subjected to AMT cross-linking (see Materials and Methods): 10 min. (lane 3), 15 min. (lane 4), 30 min. (lane 5), and 180 min. (lane 6).

Lane 1: SP6/mouse insulin pre-mRNA incubated at 30°C for 30 min in the absence of nuclear extract. Lane 2: pre-mRNA incubated at 30°C for 30 min. minus the nuclear extract and subjected to AMT cross-linking. Lane 7: A splicing reaction was incubated for 30 min. at 30°C in the absence of ATP and subjected to AMT cross-linking. The numbers on the right indicate the RNA fractions as in Fig. 1. The number 2 below fraction 1 indicates RNA species 2 shown on 5A.

rized in Figure 4A. Since psoralen is known to form covalent interstrand cross-links in accessible base-paired RNA regions, we have examined all possible pairs of potential cross-linking sites, in each fraction, to determine which ones contained suitable complementary sequences. For all RNA fractions, except 8 and 10, we have been able to find pairs of nucleotides that are involved in complementary sequences that make them AMT cross-linking sites. These interactions are summarized in Figure 4B. In addition to the unassigned stops in fractions 8 and 10, there are other stops in several of the fractions that cannot be assigned in this way (Fig. 4B, fractions 7, 9 and 11).

The pattern indicates that in general the molecules that contain the largest cross-linked loops have the slowest mobility, but the location of the loop is also important since the cross-linked molecules in fraction 11 has a lower mobility than the molecule in fraction 9. In addition, the smallest loop in a molecule is 44 nucleotides in circumference (fraction 2, U110 x U153); indicating that the gel electrophoresis may not be separating cross-linked interactions smaller than this size.

Melting of the secondary structure regions of the pre-mRNA in an in vitro splicing reaction.

We have previously reported (5) that the SP6/mouse insulin pre-mRNA was accurately spliced in vitro by a partially purified nuclear extract isolated from MOPC-315 mouse myeloma cells. The ^{32}P -labeled pre-mRNA was incubated for different times at 30°C under conditions optimal for splicing and the processed products were resolved on a 8M urea - 6% polyacrylamide gel (Fig. 5A). The products of the splicing reaction have been previously characterized by primer extension and fingerprinting analysis (5). Besides the spliced RNA product (RNA species 3), the first exon (RNA species 5) and lariat RNAs (RNA species 2 and 4) were also detected (Fig. 5A).

To determine the secondary structure of the SP6/mouse insulin pre-mRNA during the splicing reaction, AMT cross-linking was performed on the RNA after it had been incubated at 30°C in the presence of the nuclear extract for increasing times. The pattern of cross-linking in the input pre-mRNA is shown in Figure 5B (lane 2). After incubation in the presence of the nuclear extract, a different pattern was detected (Fig. 5B, lanes 3 to 6). At 10 minutes of incubation, instead of a set of discrete cross-linked bands, a smear of radioactive material migrating above the transcript and a band co-migrating with band 2 were detected. It should be noted that the free 5' exon (RNA species 5) and the RNA containing the intron-3' exon lariat (RNA species 2) were detected first at 30 min. of incubation (Fig. 5A, lane 3). Thus, the disappearance of discrete AMT cross-linked bands in the presence of the nuclear extract occurred during the lag period of incubation, before the endonucleolytic cleavage. The first band migrating above RNA species 1 (band 2 of Fig. 1) disappeared after prolonged times of incubation (3 hrs. Fig. 5B, lane 6).

The absence of discrete bands representing cross-linked molecules in the presence of the nuclear extract (Fig. 5B, lanes 3 to 6) could be either because the sites were melted or because they were inaccessible to AMT by virtue of association with proteins or splicing factors. To investigate this alternative, reverse transcriptions were carried out on pre-mRNA that was first incubated with the nuclear extract for 10 min. or 20 min. before the incubation with AMT and irradiation. In both samples, stops at specific sites throughout the molecule indicated the addition of AMT adducts (Fig. 3B). Moreover, many of the monoadducts were at the same nucleotide residues, which were cross-

Figure 6. *In vitro* splicing reactions using as substrates AMT cross-linked per-mRNAs purified by electrophoresis on denaturing gels.

The ^{32}P -labeled SP6/mouse insulin pre-mRNA (1×10^8 cpm) was incubated in splicing buffer, AMT cross-linked (Materials and Methods) and the different fractions purified (5) as described in Figure 1B. Each purified cross-linked band (4×10^5 cpm) was then used as a substrate for an *in vitro* splicing reaction incubated at 30°C for 120 min (Materials and Methods). The numbers above each lane represent the different cross-linked RNA bands, according to the nomenclature followed in Figure 1B. Minus and plus the nuclear extract are indicated only for the RNA band 1. For the rest of the bands, the left lanes are minus and the right lanes are plus the nuclear extract.

linked by AMT in the naked pre-mRNA, as determined by the primer extension analysis. In these samples, the residue C154 was a strong reverse transcription stop (Fig. 3B and data not shown), characteristic of the cross-link U153 x 110 for fraction 2 (Figs. 1 and 4B). Because of the low frequency of those molecules after incubation with the splicing extract, it was not feasible to purify them for identification. We conclude that AMT can react throughout most of the pre-mRNA molecule after incubation with the nuclear extract. Therefore, the absence of discrete cross-linked bands in the pre-mRNA incubated in splicing reactions (see Fig. 5B) indicates a melting of the pre-mRNA secondary structure contacts as monitored by the gel electrophoresis assays.

Figure 7. *In vitro* splicing reactions using as substrates photoreversed AMT cross-linked pre-mRNAs.

Samples of the purified cross-linked ^{32}P -labeled pre-mRNAs (4×10^5 cpm) used for the experiment shown in Fig. 6 were irradiated at 254 nm ($6000 \mu\text{W}/\text{cm}^2$) for 30 min. at 4°C from a distance of 3 cm. in 100 μl TE buffer (10 mM Tris-HCl pH 7.4, 1 mM EDTA). The RNAs were precipitated with ethanol and used as substrates for *in vitro* splicing reactions incubated at 30°C for 120 min. (Materials and Methods). The numbers above each lane represent the different cross-linked RNA bands, as in Fig. 6, subjected to photoreversal. Minus and plus the nuclear extract are indicated only for the RNA band 2. For the RNA bands 3 to 6, the right lanes are minus and the left lanes are plus the nuclear extract. For the RNA bands 8 to 11, the right lanes are plus the nuclear extract.

Melting of the pre-mRNA during the splicing reaction is ATP dependent.

The perturbation or melting of the pre-mRNA secondary structure was ATP dependent. A reaction incubated in the presence of the nuclear extract for 30 minutes in the absence of ATP yielded the same pattern of AMT cross-linked RNAs as the pre-mRNA minus extract (Fig. 5B, compare lanes 2 and 7). This experiment also ruled out the possibility that the disappearance of the discrete low mobility bands in Fig. 5B was due to a trace RNase activity present in the nuclear extract.

Pre-mRNAs containing AMT cross-links were inactive for splicing.

The effect on *in vitro* splicing of AMT cross-links in double-stranded regions

TABLE I

RNA Fractions	Picomoles of ^{32}P -RNA	Picomoles of ^3H -AMT Incorporated into the RNA Fractions	AMT/pre-mRNA	Picomoles of ^3H -AMT incorporated into the RNA Fractions after Photoreversal	AMT/pre-mRNA after photoreversal
1	0.12	0.18	1.5	0.12	1.0
2	0.026	0.038	1.4	0.022	0.8
3	0.014	0.019	1.3	0.016	1.1
4	0.014	0.018	1.3	0.015	1.0
5	0.012	0.017	1.4	0.016	1.3
6	0.028	0.036	1.2	0.020	0.7
7	0.014	0.019	1.3	0.014	1.0
8	0.012	0.017	1.4	0.014	1.1
9	0.012	0.017	1.4	0.012	1.0
10	0.014	0.018	1.2	0.012	0.8
11	0.028	0.038	1.3	0.024	0.8

^{32}P -labeled pre-mRNA (1.8×10^6 cpm in 100 μl reaction, 1 picomole = 4×10^6 cpm) containing 1 μg of tRNA as a carrier was incubated in splicing buffer, the AMT was added as described in Materials and Methods. A similar reaction (100 μl) containing an equivalent amount of cold pre-mRNA was incubated with ^3H -AMT (specific activity 4.3 Ci/mM, 1.1 μCi) as described in Materials and Methods. The different RNA fractions were purified from the gel as described in the legend to Fig. 1, using the ^{32}P -labeled RNAs as markers. The cpm of the different RNA fractions were then determined. One picomole of ^3H -AMT incorporated into the RNA = 1×10^4 cpm. The same experiment was performed with ^{32}P -labeled pre-mRNA and ^3H -AMT. After subtracting the spill-over from the ^3H -channel identical results were obtained (data not shown). The Photoreversal of the RNA fractions were as described in the legend to Fig. 7.

located at various levels within the pre-mRNA was investigated. The ^{32}P -labeled pre-mRNAs were subjected to AMT cross-linking as described for the experiment of Figure 1. The cross-linked RNAs were then fractionated by denaturing gel electrophoresis (Fig. 1) and used as substrates for *in vitro* splicing reactions. Each purified cross-linked RNA band was incubated for 120 min. at 30°C under conditions optimal for splicing (see Materials and Methods). Figure 6 shows that all of the cross-linked pre-mRNAs were inactive for the splicing reaction. The RNA band 1 which most likely is a mixture of unreacted and AMT monoadducts (see Fig. 3) was the only one which generated the products of the splicing reaction (Fig. 6, lane 1).

To determine whether the inhibition of the splicing reaction with the cross-linked pre-mRNAs was due to an unspecific effect of the psoralen "per se", a reversal of the AMT cross-links RNAs by a short wavelength ultraviolet light was carried out. Each purified cross-linked RNA band was irradiated at 254 nm as described in the legend to

Figure 3. Secondary structure of SP6/mouse insulin pre-mRNA in *in vitro* splicing buffer determined by AMT cross-linking. The proposed complementarities where AMT is able to cross-link are shown in heavy diagonal bars. The molecule must exist in multiple conformations since different complementarities share the same sequences. The boxed numbers indicate the fractions in which the particular cross-link was identified. For fraction 2 the stem between nucleotides 122-127 and 136-142 was predicted by computer analysis (see text).

Fig. 7 and subsequently incubated under conditions optimal for splicing (see Materials and Methods). Figure 7 shows that all of the cross-linked RNA bands subjected to a photochemical reversion migrated in a denaturing gel electrophoresis as the untreated pre-mRNA and in the presence of the splicing extract were able to generate the products of the splicing reaction. Furthermore, the experiment presented in Table I shows that under the conditions of photoreversion used in our experiments the pre-mRNAs still contain covalently attached AMT molecules, about 60-70% of the amount before the reversal. Taken together, these results indicate that the AMT at the cross-links sites do not have any effect on the splicing reaction and that destabilization or unwinding of the double-stranded regions, at least for some of the fractions (see Discussion), which were detected by AMT cross-linking are necessary for *in vitro* splicing.

DISCUSSION

The pre-mRNA can adopt many secondary structure conformations.

Secondary structure interactions in the SP6/mouse insulin pre-mRNA assayed by AMT cross-linking are shown in Figure 8. Judged by the pattern of AMT cross-linking, the pre-mRNA is able to assume at least seven conformations in splicing buffer in the absence of the nuclear extract. The base-pairing interactions that were detected by the cross-linking were located in the 5' 220 nucleotides of the pre-mRNA surrounding the intron (Fig. 8). They are most likely high affinity sites for AMT intercalation at the input of AMT concentration that was used; all of them involve sites where AMT reacts between uridine residues in weak or interrupted helices. Therefore, it is possible that other conformations involving secondary structure interactions in other regions of the pre-mRNA were not detected by our analysis of the twelve most frequent bands. Psoralen does not necessarily intercalate into every double-stranded RNA region and also cross-linking requires the presence of appropriate pyrimidine residues in opposite strands (24). However, in an experiment in which the RNA was incubated with a 30-fold higher concentration of AMT and the AMT monoaddition was limited by the time of exposure to 390 nm light, the pattern of cross-linking detected by gel electrophoresis was the same except that the frequency of several of the bands was altered (see Fig. 2). In this experiment the RNAs containing monoadducts had to be renatured before cross-linking (see Fig. 2). The fact that a similar cross-linking pattern resulted as compared with RNAs directly cross-linked (Fig. 2, lanes 4 and C) indicates that the RNAs renatured into its original conformations. Therefore, although only a fraction of the RNAs are cross-linked at any particular location, they are most likely indicative of the conformation of all of the molecules. An additional observation is that the smallest loop in a molecule resolved by gel electrophoresis was 44 nucleotides in circumference (Fraction 2 in Fig. 1), thus there may be additional short-range base-pairing interactions that have been cross-linked but were not detected by this technique.

In a further effort to understand the interactions that have been detected by AMT cross-linking, computer analysis (25,26) of the SP6/mouse insulin pre-mRNA was used to predict secondary structures for the molecule. The entire molecule folded into the structure of the lowest free energy did not predict any of the sites that were actually found to be cross-linked even though there are "high affinity" AMT cross-linkable sites in this predicted structure. However, if the first 160 nucleotides were folded, the structure contained a base-pairing interaction between nucleotides 105-113 and 149-156 in which U110 could be cross-linked to U153, as has been found in fraction 2 (see fraction 2 in Fig. 8). The predicted structure contained also base-pairing interaction between nucleotides 122-127 and 136-142 (see fraction 2 in Fig. 8). In addition, the pattern of reverse transcription suggests that twelve of the stops in fraction 1 might

result from AMT adducts at short-range cross-linking sites; examination of predicted short-range structures reveals base-paired interactions at nucleotides 300 to 324 and 323 to 353 that could account for ten of the AMT adducts.

All of the long-range AMT cross-links involve the left half of the pre-mRNA including and surrounding the intervening sequence. They would act to shorten the distance between the donor and acceptor splice sites although in many instances the splice sites would not be correctly positioned within the secondary structure. It must be emphasized that we do not know whether these interactions have any functional importance. It is possible, although unlikely, that the peculiar distribution of AMT cross-linkable long-distance contacts in this molecule is due to the distribution of UpU and CpU dinucleotides in the molecule.

The pre-mRNA secondary structure is melted before the first endonucleolytic cleavage of the splicing reaction.

We have demonstrated here that the SP6/mouse insulin pre-mRNA upon incubation with the splicing extract undergoes a rearrangement of its secondary structure. Our results indicate that during a splicing reaction melting or unwinding of the secondary pre-RNA structure occurs in the presence of ATP. It is likely that ATP is hydrolyzed during the unwinding reaction. In the presence of the imidonucleotide AMP-PNP, which has a non-hydrolyzable β - γ phosphodiester bond, the unwinding of the pre-mRNA secondary structure did not occur (data not shown). An ATP-dependent RNA unwinding activity (helicase) that is involved in the formation of an RNP particle during a splicing reaction has been previously suggested by experiments introducing regions of secondary structure in the pre-mRNA (13,27) or by measuring the susceptibility of the pre-mRNA to RNAase-A in the presence of the splicing extract (28). The use of ATP as a source of energy necessary to melt out mRNA secondary structure has been described recently (29) in an *in vitro* translation system, for the eukaryotic initiation factor (eIF)-4A, which is enhanced by eIF-4B and eIF-4F under certain conditions (30).

An RNA unwinding activity was recently reported (31,32) in fertilized frog eggs. This activity was present at low levels in *Xenopus* oocytes, increased during oocyte maturation, and it was at high levels throughout early embryogenesis. The activity, however, was assayed only on intermolecular regions of RNA:RNA hybrids. The nature of the reactions that the unwinding activity of fertilized frog eggs mediates is unknown.

Trans-splicing *in vitro* of two RNAs associated by a segment of a complementary sequence in the intron was recently demonstrated (27,33). Introduction of a region of secondary structure between two separated exons affected the efficiency of the *in vitro* splicing reaction. The associated RNA substrates were 15% active compared to reactions containing the equivalent precursor RNA (27). The preannealed double-stranded region was maintained in at least some of the RNA molecules present in the splicing reaction

(27). Therefore, it is quite possible that for the trans-splicing reaction melting of the particular introduced stem is not necessary. Most importantly, though, it was observed that a large fraction of the preannealed RNA precursor does dissociate during in vitro splicing in an ATP-dependent process, suggesting the presence in the nuclear extract of an helicase activity (27).

We have demonstrated here that pre-mRNAs containing AMT cross-links were not active as substrates for the splicing reaction. Photoreversal of the cross-link sites, however, yielded pre-mRNAs still containing covalently attached AMT molecules that were able to serve as active substrates for the splicing reaction (see Figs. 6 and 7). The cross-links in the RNA molecules of fractions 4 to 6 and 11 constrain the donor splice site to be within a closed loop, so they may be inactivated for this reason. The cross-links in the molecules in fractions 2 and 3 occur within the intervening sequence well away from the splice sites and branch point. However, these cross-links produce fixed loops in the intron of 44 and 54 nucleotides in circumference, so the intervening sequence (118 nucleotides) is effectively shortened to 74 and 64 nucleotides, respectively. It is known that minimum intron sizes of about 80 nucleotides and 100 nucleotides are needed for in vivo splicing (34) and in vitro splicing (35), so this may be the reason for the inactivation of the intron in molecules from fractions 2 and 3. Our results showed that the base-paired interactions in fraction 2 are resistant to be melted during the splicing reaction (Fig. 5B), it is possible that not all of the double-stranded regions within a pre-mRNA are necessarily melted during splicing. Alternatively, those base-pair interactions are transiently melted, and then re-established again in the spliceosome. We do not know yet whether fraction 2 is unable of spliceosome formation or the inhibition of splicing is at a later step. The crosslinks in fractions 7 and 9 occur between the exon sequences; the complete intervening sequence in those cases is sequestered in circular RNA of 181 and 216 nucleotides, respectively. It is surprising then that these molecules do not appear to be spliced. This result implies that in vitro splicing is inhibited if the RNA is topologically constrained.

The smear seen in electrophoresis of RNA samples from splicing reactions subjected to AMT cross-linking may result from a large number of different cross-linked molecules (intramolecular or intermolecular interactions with other RNA components of the spliceosome) that are not resolved into sharp bands. Intermolecular interactions between hnRNAs and snRNAs have been previously demonstrated in vivo by psoralen cross-linking experiments (36).

The secondary structure of the pre-mRNA we have studied was determined in a splicing buffer. Folding of the pre-mRNA might be a different situation in the nucleus in the presence of the transcription apparatus and RNA binding proteins. After a transcript is synthesized, RNA binding proteins might actively promote removal of stems

exposing regions of the pre-mRNA for binding of other splicing components like snRNPs. The formation of such splicing committed complex would generate a splicing site in the pre-mRNA. In this model, splice sites can become optional when they are hindered or sequestered in regions of the pre-mRNA structure which were not melted. For example, binding of the hnRNPs proteins or snRNPs to certain regions of the RNA might be impaired either by the particular secondary structure of the pre-mRNA or by the presence of specific trans-splicing factors.

ACKNOWLEDGEMENTS

This work was supported by NIH grants AI-24479 to C.J.G and GM 35410 to P.W.

*Present address: Department of Biology, Medical School, Pecs, Hungary

REFERENCES

1. Padgett, R.A., Grabowski, P.J., Konarska, M.M., Seiler, S., and Sharp, P.A. (1986). *Ann. Rev. Biochem.* 55, 1119-1150.
2. Green, M.R. (1986). *Ann. Rev. Genet.* 20, 671-706.
3. Frendwey, D. and Keller, W. (1985). *Cell* 42, 355-367.
4. Grabowski, P.J., Seiler, S.R., and Sharp, P.A. (1985). *Cell* 42, 345-353.
5. Kaltwasser, G., Spitzer, S.G. and Goldenberg, C.J. (1986). *Nucleic. Acids. Res.* 14, 3687-3701.
6. Mount, S.M., Pettersson, I., Hinterberger, M., Karmas, A., and Steitz, J.A. (1983). *Cell* 33, 509-518.
7. Kramer, A., Keller, W., Appel, B., and Luhrman, R. (1984). *Cell* 38, 299-307.
8. Choi, Y.D., Grabowski, P.J., Sharp, P.A., and Dreyfuss, G. (1986). *Science* 231, 1534-1539.
9. Rogers, J.H. (1985). *Int. Rev. Cytol.* 93, 188-279.
10. Kuhne, T., Wieringa, B., Reiser, J., and Weissmann, C. (1983). *Embo J.* 2, 727-733.
11. Lang, K.M. and Spritz, R.A. (1983). *Science* 220, 1351-1355.
12. R. Reed and T. Maniatis. (1986). *Cell* 46, 681-690.
13. Solnick, D. (1985). *Cell* 43, 667-676.
14. Cimino, G.D., Gamper, H.B., Isaacs, S.T., and Hearst, J.E. (1985). *Ann. Rev. Biochem.* 54, 1151-1193.
15. Wollenzien, P.L. and Cantor, C.R. (1982). *J. Mol. Biol.* 159, 151-166.
16. Goldenberg, C.J. (1984). *Proc. Natl. Acad. Sci. U.S.A.* 81, 4707-4711.
17. Goldenberg, C.J. and Hauser, S.D. (1983). *Nucleic Acids Res.* 11, 1337-1348.
18. DiMaria, P.R., Kaltwasser, G., and Goldenberg, C.J. (1985). *J. Biol. Chem.* 260, 1096-1102.
19. Krainer, A.R., Maniatis, T., Ruskin, B., and Green, M.R. (1984). *Cell* 37, 415-427.
20. Isaacs, S.T., Shen, C.K.J., Hearst, J.E., and Rapoport, H. (1977). *Biochemistry* 16, 1058-1064.
21. Qu, L.H., Michot, B., and Bachellerie, J.P. (1983). *Nucleic Acids Res.* 11, 5903-5920.
22. Wollenzien, P., Hearst, J.E., Thammana, P., and Cantor, C.R. (1979). *J. Mol. Bio.* 135, 255-269.
23. Youvan, D.C. and Hearst, J.E. (1982). *Anal. Biochem.* 119, 86.
24. Garrett-Wheeler, E., Lockard, R.E., and Kumar, A. (1984). *Nuc. Acids Res.* 12, 3405-3423.
25. Zuker, M. and Stiegler, P. (1981). *Nucl. Acids Res.* 9, 133-148.
26. Jacobson, A.B., Good, L., Simonetti, J., and Zuker, M. (1984). *Nucl. Acids Res.* 12, 45-54.

27. Konarska, M., Padgett, R.A., and Sharp, P.A. (1985). *Cell* 42, 165-171.
28. Ruskin, B. and Green, M.R. (1985). *Cell* 43, 131-142.
29. Ray, B.K., Lawson, T.G., Kramer, J.C., Cladaras, M.H., Grifo, J.A., Abramson, R., D., Merrick, W.C., and Thach, R.E. (1985). *J. Biol. Chem.* 260, 7651-7658.
30. Grifo, J.A., Abramson, R.D., Satler, C.A., and Merrick, W.C. (1984). *J. Biol. Chem.* 259, 8648-8654.
31. Rebagliati, M.R. and Melton, D.A. (1987). *Cell* 48, 599-605.
32. Bass, B.L. and Weintraub, H. (1987). *Cell* 48, 607-613.
33. Solnick, D. (1985). *Cell* 42, 157-164.
34. Wieringa, B., Hofer, E., and Weissmann, C. (1984). *Cell* 37, 915-925.
35. Padgett, R.A., Konarska, M.M., Aebi, M., Hornig, H., Weissmann, C., and Sharp, P.A. (1985). *Proc. Natl. Acad. Sci. USA* 82, 8349-8353.
36. Calvet, J.P. and Pederson, T. (1981). *Cell* 26, 363-370.