The structure of a pre-mRNA molecule in solution determined with a site directed cross-linking reagent

John Teare and Paul Wollenzien*
Edward A.Doisy Department of Biochemistry and Molecular Biology, St Louis University Medical Center, St Louis, MO 63104, USA

Received November 10, 1989; Revised and Accepted January 12, 1990

ABSTRACT

We describe the use of site specific psoralen (SSP) to determine the solution structure of a segment of the human beta globin pre-mRNA. In these experiments, SSP is first delivered as monoadducts to specific nucleotides in the pre-mRNA and subsequently used to form intramolecular RNA-RNA cross-links. The use of this reagent greatly decreases the number of the cross-linked products as compared to generalized psoralen cross-linking. The experiments confirm the locations of previously determined aminomethyltrimethylpsoralen (AMT) cross-links in the human precursor mRNA. In addition, new cross-links consistent with an alternative secondary structure and a small number of cross-links that represent higher order interactions have been determined. Altogether, 42 of 47 cross-links identified in this analysis can be accounted for in a small number of alternative secondary structures and higher order interactions. The site directed cross-linking technique will be useful for the precise determination of RNA secondary and tertiary structures under a variety of experimental conditions.

INTRODUCTION

Psoralen is a planar, aromatic compound that intercalates into double stranded regions of nucleic acids. Irradiation of psoralen-nucleic acid complexes with long wave ultraviolet light results in the formation of covalent psoralen adducts as well as cross-links in which two segments of nucleic acid are covalently joined through a psoralen molecule. Because of these properties psoralen has been used as a probe for RNA and DNA secondary and tertiary structures *in vivo* and *in vitro* (1,2).

Usually the cross-linking experiments are done by incubating the nucleic acid with psoralen in the dark to form intercalated complexes and then irradiating with light of the appropriate intensity and wavelength to form cross-links. Several disadvantages are inherent in this approach. For molecules of even modest size, a large variety of photoproducts are produced which must be separated as well as possible and identified individually during the analysis. This usually necessitates a large starting sample. In addition, since it is not possible to control the binding constants or accessibility of the sites in the nucleic acid for psoralen it is a matter of chance if sites of interest are also psoralen cross-linking sites.

For these reasons, we have developed a method for introducing site directed psoralen cross-links into single stranded RNA molecules. A site specific psoralen derivative (SSP, ref. 3) is placed as a covalent monoadduct at a targeted site in the RNA. The RNA molecule is allowed to refold in an appropriate buffer and the SSP is irradiated to form cross-links. We have previously used AMT cross-linking, UV cross-linking and chemical probing experiments to determine the secondary structure of human and rabbit beta globin pre-mRNA in solution (4). In this report we describe the analysis of cross-link locations of SSP after delivery to selected sites in a human pre-mRNA. This will determine whether SSP indicates the same structure as AMT and UV cross-linking and will provide additional information about alternate secondary structures in the RNA molecule. We find that the reduction in the number of psoralen photoproducts by targeting the SSP to specific regions of the pre-mRNA simplifies the analysis of cross-link location. The patterns of cross-linking support the secondary structure model and provide some information for the higher order structure of the molecule in solution.

MATERIALS AND METHODS

Materials

2-pyridyldithioethylmethylamidodiethoxyethanemethylamino-methyltrimethylpsoralen (site specific psoralen, SSP) was purchased from HRI Associates. 1-methylimidazole and cystamine were from Aldrich and 1-ethyl 3,3 dimethylanimopropylcarbodiimide (EDC, protein sequencing grade) was from Sigma. SEP-PAC C18 reverse phase cartridges were obtained from Waters Associates and DEAE Toyopearl 650M ion exchange resin was from Toyo Soda MFG. Nucleotides, deoxynucleotides, and dideoxynucleotides were from Pharmacia. T-7 RNA polymerase was purified according to Davanloo et al. (5). Avian Myeloblastosis Virus (AMV) reverse transcriptase was from Life Sciences and RQ1 DNase I was purchased from Promega. Oligonucleotides were

* To whom correspondence should be addressed

synthesized on an Applied Biosystems 380A DNA synthesizer and were gel purified before use.

In vitro Transcription A 495 nucleotide run off transcript of the human beta globin pre-mRNA encoding the first exon, the first intron, and most of the second exon was synthesized with T7 RNA polymerase and BamH I-linearized pTZ19R-HBG template as previously described (4).

Addition of Cystamine to the 5′ End of the DNA Oligonucleotides 2000 pmol of [^{32}P] 5′-oligonucleotide was reacted with cystamine as described (6). The product, 5′cystamine-oligonucleotide, was purified as previously reported (7) except that the DEAE ion exchange and the C18 reverse phase columns were doubled to 300 μl and 60 to 70 μl respectively. The eluate from the C18 column in 50% methanol/50 mM triethylammonium carbonate, pH 7.2 (TEAC) was dried to 100 μl and was spun through a G-50 Sephadex column (3 mL volume) equilibrated to 50 mM TEAC. The 5′cystamine-oligonucleotide was dried and resuspended in 10 mM Tris, HCl, pH 7.2, 1 mM EDTA (TE).

Addition of SSP to the Oligonucleotide 400 to 500 pmol of 5′cystamine-oligonucleotide in 100 μl TE was treated with 10 mM DTT for 1 hour at 23°C. The reaction mixture was spun through a G-50 Sephadex column equilibrated with 50 mM TEAC/50 μM DTT. To the eluate (100 μl), 100 μl of methanol and 12 to 15 nmol of SSP (in ethanol) were added and the solution was incubated as for 3 hours at 37°C. The reaction mixture was dried to 100 μl and spun through a G-50 Sephadex column equilibrated with 50 mM TEAC. The eluate was dried to 10 μl. 20 μl formamide was added and the SSP-oligonucleotide was electrophoresed on a 20% polyacrylamide, 8.3 M urea gel. Autoradiography of the gel revealed a prominent new band (compared to the starting oligonucleotide) that is the SSP-oligonucleotide. This was cut out, eluted, and concentrated as previously described (7).

Hybridization and Monoaddition of SSP-oligonucleotide to the RNA 30 to 50 μg of [^{32}P] UTP labeled human beta globin pre-mRNA was dried with an equal molar amount of SSP-oligonucleotide. The sample was resuspended in 10 μl of R-loop buffer (80% formamide, 0.4 M NaCl, 40 mM Pipes, pH 6.8, 1 mM EDTA) and hybridized for 15 min. at 45°C. After cooling on ice for 5 minutes, 390 μl of TE with 50 mM NaCl was added and the solution was incubated for 10 minutes at 4°C. The sample either was irradiated for 1 hour at 4°C using light from a high intensity mercury light irradiator (8) that has passed through 1.5 cm of a 0.85 M $CuCl_2$ solution (9) or was irradiated under N_2 atmosphere for 2 hours at 4°C with 380 nm light (3 mW/cm^2). The RNA/SSP-oligonucleotide monoadduct was ethanol precipitated and resuspended in dH_2O.

Removal of the Oligonucleotide 25 to 50 μg of the RNA/SSP-oligonucleotide monoadduct was digested with 50 units of DNase I in 300 μl DNase I buffer (40 mM Tris-HCl, pH 7.9, 10 mM NaCl, 6 mM $MgCl_2$) with 10 mM DTT for 40 minutes at 37°C. 5 μl of 1.0 M 2-mercaptoethanol (in TE, pH 7.2) was added and the RNA was phenol extracted, ether extracted, and twice ethanol precipitated. The RNA/SSP monoadduct was stored in dH_2O at −20°C.

Cross-Linking of the RNA/SSP Monoadducts 20 to 50 μg of RNA/SSP monoadduct was incubated at 10 min. at 4°C in 500 μl of spliceosome assembly buffer and cross-linked for 10 min. at 4°C with 320−380 nm light (100 mW/cm2, ref 8). The cross-linked RNA was ethanol precipitated, resuspended in dH_2O, and electrophoresed on a 5% polyacrylamide/8.3 M urea gel at 36 volts/cm for 4 hours. The cross-linked RNA bands were visualized by autoradiography, excised, and eluted as described (4).

Determination of Cross-Link Location The locations of the SSP cross-links in the purified RNA fractions were determined by a series of primer extension reactions as previously described (4). In the present experiments, 1 to 10 ng of cross-linked purified RNA, 2 pmol of ^{32}P-labeled oligonucleotide primer and 2 units of AMV reverse transcriptase were used in each 20 μl reaction. Control reactions on linear RNA were terminated with dideoxynucleotides to indicate the nucleotide sequence. The reactions were phenol extracted, treated with 160 mM NaOH for 30 min. at 55°C, ethanol precipitated, and redissolved in 4 μl dH_2O. 1 μl was denatured and electrophoresed on 8% denaturing gels and viewed by autoradiography. The primers used in these experiments were complementary to the following regions: 175−193, 245−263, 332−348, 429−449, 459−477, and 482−495.

RESULTS

Eight DNA oligonucleotides were designed and synthesized for the purpose of placing SSP monoadducts on specific uridines in a human beta globin pre-mRNA. The structure of SSP, SSP-oligonucleotides, and the steps used deliver the SSP to target sites on the RNA molecule are shown in Figure 1. We have previously shown that SSP coupled to a DNA oligonucleotide in this way can be attached efficiently to the RNA target site by hybridization and irradiation of the complex with 365 nm light (7). In these reactions, the sequence of the DNA oligonucleotide is chosen so that the RNA has one unpaired nucleotide adjacent to the target uridine; this arrangement enhances the reaction of the SSP at the target uridine presumably by creating a favorable intercalation site (7). In the present experiments, RNA/SSP-oligonucleotide complexes were irradiated first with 380 nm light in order to produce SSP monoadducts. Since the structure of the complex has been designed to bring the furan side of the SSP close to the RNA target site and since 380 nm light preferentially activates the furan side for monoaddition, a significant fraction of complexes should result in addition of SSP monoadducts to the RNA target site. The DNA oligonucleotide can then be removed by DNase I digestion and reduction. SSP monoadducts made in this way should be capable, at a later time, of absorbing 340−370 nm light to form intramolecular RNA-RNA cross-links.

Target sites were chosen by inspecting the secondary structure model of the human beta globin pre-mRNA that we previously proposed (4). This model was derived from the pattern of chemical reactivity of bases, analysis of the sites of AMT and UV cross-linking, and consideration of the minimum free energy of possible structures. All of the targeted uridines were intended to be in positions to form cross-links in the main secondary structure; in addition, many of the targeted uridines would also be in positions to form cross-links in alternative secondary structural elements. Figure 2 shows the sequences of the SSP-oligonucleotides and their target sites in the RNA.

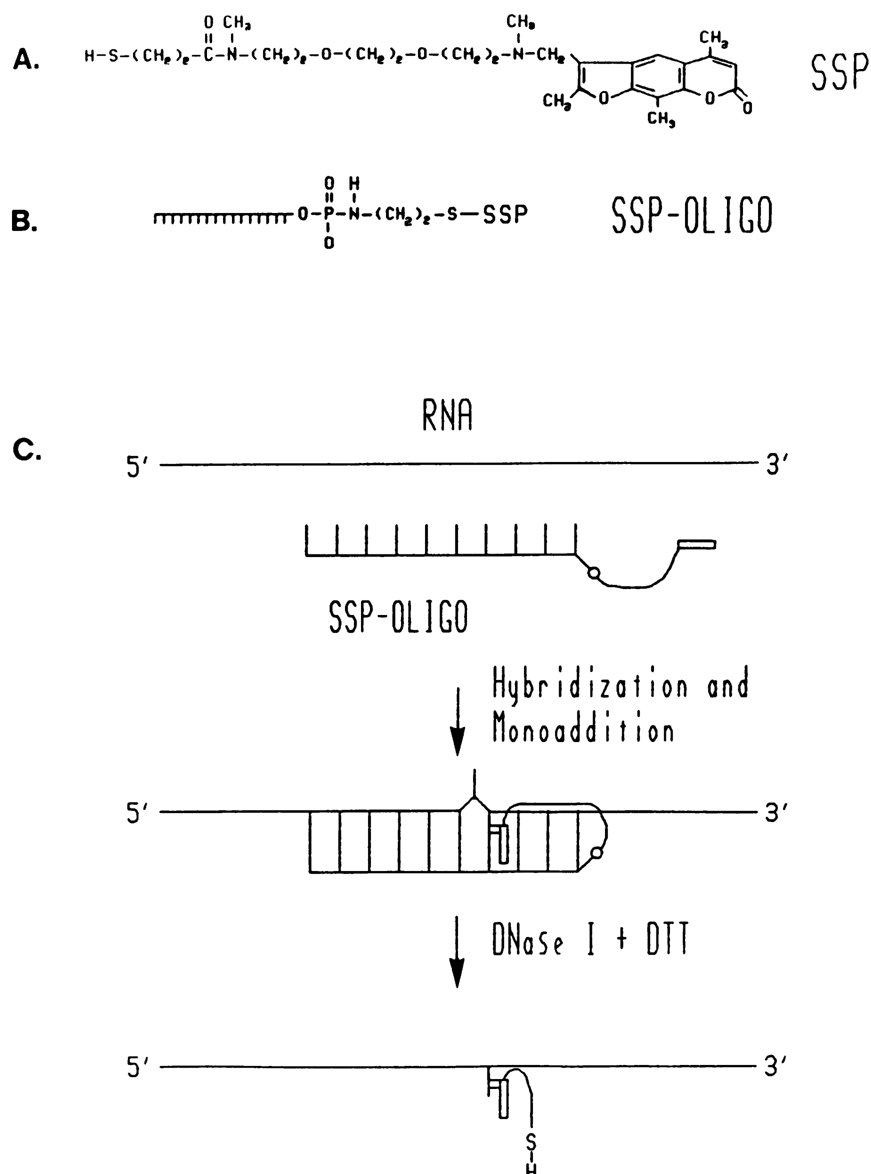


Figure 1. Synthesis of the site specific psoralen (SSP) monoadduct on the RNA target site. (A) The structure of SSP. (B) The structure of the junction in the SSP-oligonucleotide. The DNA oligonucleotide is represented by the horizontal line with the perpendicular bars. The 5′ phosphate of the oligonucleotide is modified with a thioethylamine group and is then coupled to SSP through a disulfide linkage. (C) The SSP-oligonucleotide is hybridized to the RNA target site followed by intercalation of the SSP into the heteroduplex between the base pairs adjacent to the bulge. The complex is irradiated with 380 nm light resulting primarily in furan side monoaddition of the SSP onto the 3′ face of the RNA target uridine. The oligonucleotide is removed by treatment of the RNA/SSP-oligonucleotide monoadduct with DNase I under reducing conditions.

Analysis by primer extension assays using AMV reverse transcriptase and synthetic DNA oligonucleotide primers showed that six of the eight SSP monoadducts were located primarily on the targeted uridines (Fig. 2). A psoralen monoadduct or cross-link is an absolute stop for an elongating reverse transcriptase, the pattern of stops providing information on the stereochemistry of the psoralen adduct (10). A double stop one nucleotide before and at the pyrimidine involved in the psoralen adduct indicates that the adduct is located on the 5′ face of the pyrimidine base; a single stop one nucleotide before the pyrimidine involved in the adduct indicates that the adduct is on the 3′ face of the base. The distribution of SSP monoadducts indicates SSP is usually adding to the 3′ side of the target base. Therefore, SSP is probably intercalated between base-pairs adjacent to the bulge position.

There are sequence dependent variations in the efficiency of the site directed monoadducts. Estimates of efficiencies using a gel shift assay ranged from 10% for SSP-92 to 40% for SSP-264 (data not shown). Most of the target sequences showed only low levels of monoaddition at locations other than the target uridine. SSP-241 and SSP-262 contain UpG and UpA dinucleotides within the heteroduplex region which are high efficiency psoralen photobinding sites (11,12,13) and, as shown in figure 2, compete with the intended site for intercalation and photoaddition of SSP. In fact, SSP-262 was originally designed to direct SSP to U265, but since the addition to U262 was so efficient, it was renamed. In the case of SSP-386, efficiency of monoaddition at the target uridine was increased by creating two adjacent, unpaired nucleotides in the heteroduplex at the target site. Efficiency of

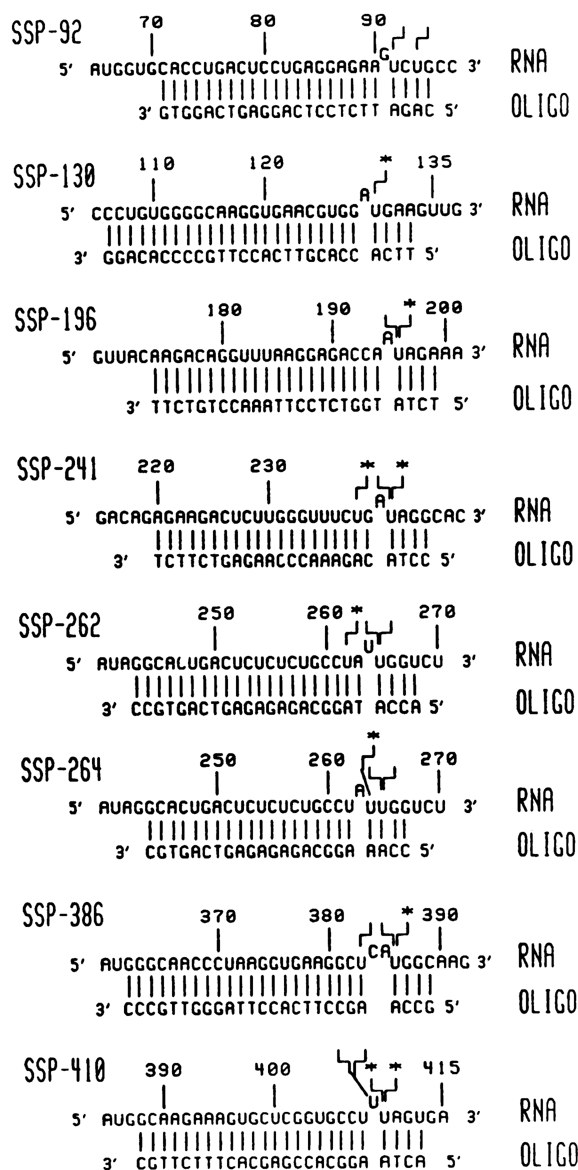


Figure 2. Structure of the target sites and DNA oligonucleotides used in the SSP placement. The locations of the SSP monoadducts are indicated by the broken lines; the left or right orientation of the break in the line indicates the location of the adduct on the 5′ or 3′ face of the base. Strong monoaddition sites are marked with an asterisk. SSP-241 allowed efficient monoaddition at 2 sites; U238 and U241. SSP-262 was designed to place the adduct at U265, however, the monoaddition occurred primarily at U262. Where there is ambiguity in the base pairing at the target site, such as with SSP-410, there was less specificity for monoaddition at a single target uridine.

addition to the 3′ side of U386 was less than 5% with one nucleotide unpaired and increased to 25% with two nucleotides unpaired. This did not change the location of the adduct.

Pre-mRNA containing each of the SSP monoadducts was treated to remove the DNA oligonucleotide and was redissolved in spliceosome assembly buffer at 4°C. After they were cross-linked with 365 nm light, the individual cross-linked species were separated according to their different electrophoretic mobility on denaturing gels (4). The major bands were purified and the locations of the cross-links in each fraction were determined by a series of primer extension experiments (4). Three criteria were used to judge whether pairs of reverse transcription stops represent cross-linking sites: (i) that the reverse transcription stops were fraction specific, (ii) that the stops were *not* detected with the same primer since in an authentic cross-link the 3′ site will be an absolute stop (10) and will prevent the reverse transcriptase from reaching the 5′ site, and (iii) that the loop size predicted by the cross-link was consistent with the electrophoretic mobility of the molecules. In addition, if there was base-pairing complementarity between the two sites and this was supported by the pattern of chemical modification (see ref. 4), then the cross-link was attributed to a secondary structure interaction.

Figure 4 shows an example in which two primer extension experiments were done on cross-linked fractions from SSP monoadducts placed on U241 and U410 in the pre-mRNA. These two uridines were cross-linked together by AMT (4); thus we expected that a SSP monoadduct at either position would result in cross-linking to the complementary position. The reverse transcription stops in fractions 1 and 2 of the SSP-410 experiment and fractions 7 and 8 of the SSP-241 experiment (Fig. 4, lanes

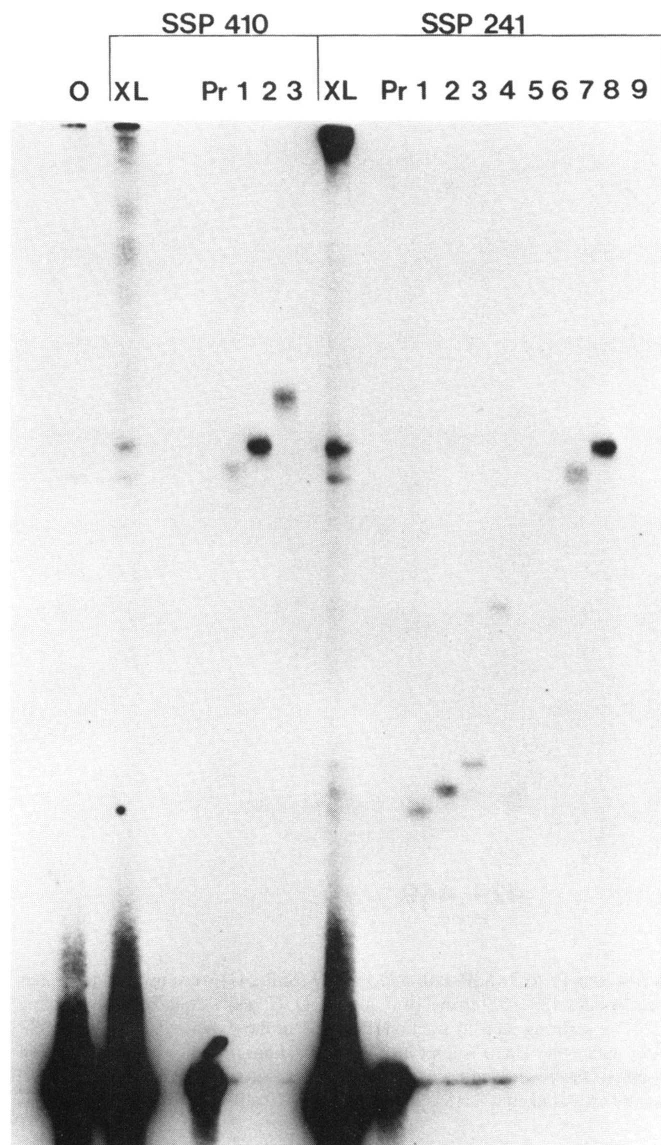


Figure 3. Gel electrophoresis of SSP cross-linked RNA. An autoradiogram of a 5% polyacrylamide/8.3 M urea gel is shown . Lane O; SSP-410 RNA before re-cross-linking with 365 nm light. For SSP-410 and SSP-241, Lanes XL; SSP monoadduct RNA in spliceosome assembly buffer at 4°C re-cross-linked with 365 nm light (unfractionated). Lanes Pr, Cross-linked RNAs that co-migrate with linear RNA. Lanes 1 through 3 (SSP-410) and Lanes 1 through 9 (SSP-241) are the purified cross-linked RNA molecules.

1 and 2, SSP-410 and lanes 7 and 8, SSP-241) indicate that cross-linking between these two positions can occur with monoadducts placed at either nucleotide. We also observed additional cross-links from these two positions. Fraction 3 of the SSP-410 data set contained two cross-linked species; one to U196 and one to U184. These cross-links are identical to previously identified AMT cross-links. The monoadduct at U241 gave many additional cross-links. The fraction Pr (molecules which co-migrate with linear RNA) and fraction 1 gave reverse transcription stops that indicated cross-links to positions U264 and U265 respectively. These cross-links correlate well with an alternative structure proposed for this region of the precursor (Fig. 6, and refs. 4 and 14). On the other hand, crosslinks found in the SSP-241 experiment in fractions 1, 2, 3, and 6 do not occur between complementary sequences and are not consistent with chemical probing data on this RNA (4). We speculate that some of these cross-links may represent higher order interactions in the RNA (see below).

The data for the cross-link locations in the human beta globin pre-mRNA using the 8 site directed SSP reagents is summarized in Figure 5 and in Table 1. It should be noted that in some of the fractions that were isolated the cross-linking sites were not identified because the low yield of cross-linked molecules made interpretation of the primer extension experiments very difficult.

Two cross-links out of the 46 that we have analyzed give evidence for an unusual psoralen transfer event. In the first, fraction 9 of the SSP-241 data set shows SSP adducts at U184, U241, and U410 (Fig. 4). The monoadduct was originally placed at position U241 but the cross-links U184 × U241 or U241 × U410 would not result in loop sizes large enough to retard migration on a denaturing gel by 70% (RF= 0.28), (Fig. 3, lane 9). In fact, fraction 9 of SSP-241 and fraction 3 of SSP-410 have identical RF values (Fig. 3). Thus the cross-link found in the SSP-241 experiment must be U184 × U410, suggesting that the SSP monoadduct at position U241 was transferred to U410 during the 365 nm irradiation and then mediated a cross-link between U410 and U184. In the second unusual cross-link, in fraction 6 of the SSP-386 data set (Table 1), the adduct originally at U386, was transferred to U264, and then mediated a cross-link between U264 and U376. In both cases, a residual reverse transcription pause occurred at the original target site perhaps because the transfer reaction, being catalyzed by light, caused some damage to the base. It should be emphasized that this type of transfer is unusual and can be recognized by the pattern of reverse transcription stops and electrophoretic mobility of the cross-linked molecules.

From the data obtained from the SSP experiments and previously reported chemical probing data (4), two alternative structures for the 3′ exon domain for the human precursor mRNA can be proposed (Fig. 6). Altogether, 29 cross-links from the 8 data sets were incorporated into the alternative secondary structural models.

Twelve additional cross-links may result from the higher order structure of the molecule. The cross-links are specific (i.e. a single cross-link per fraction), and are present at moderate frequency. They usually occur between sequences which have two base-pairs at the site of the psoralen intercalation, but otherwise have little complementarity and are not protected from chemical reaction. At 4°C and in the presence of Mg^{++}, these interactions appear to be specific and stable enough to allow SSP cross-linking. Such interactions could arise from the folding of the secondary structures on top of each other. A summary of the locations of these cross-links is shown in Table 1 and Figure 6.

DISCUSSION

We have described the use of a site directed cross-linking reagent to study the secondary structure of a large RNA molecule. The synthesis of the SSP-oligonucleotides is both convenient and efficient. Hybridization of the reagent to the RNA target sequence and irradiation of the complex with 380 nm light usually results in efficient and specific furan-side monoaddition of SSP to the intended uridine nucleotide. The oligonucleotide can be removed from the RNA by treatment with DNase I and DTT leaving the SSP molecule attached at the target site. Cross-linking of the SSP monoadduct RNA under native conditions results in a small set

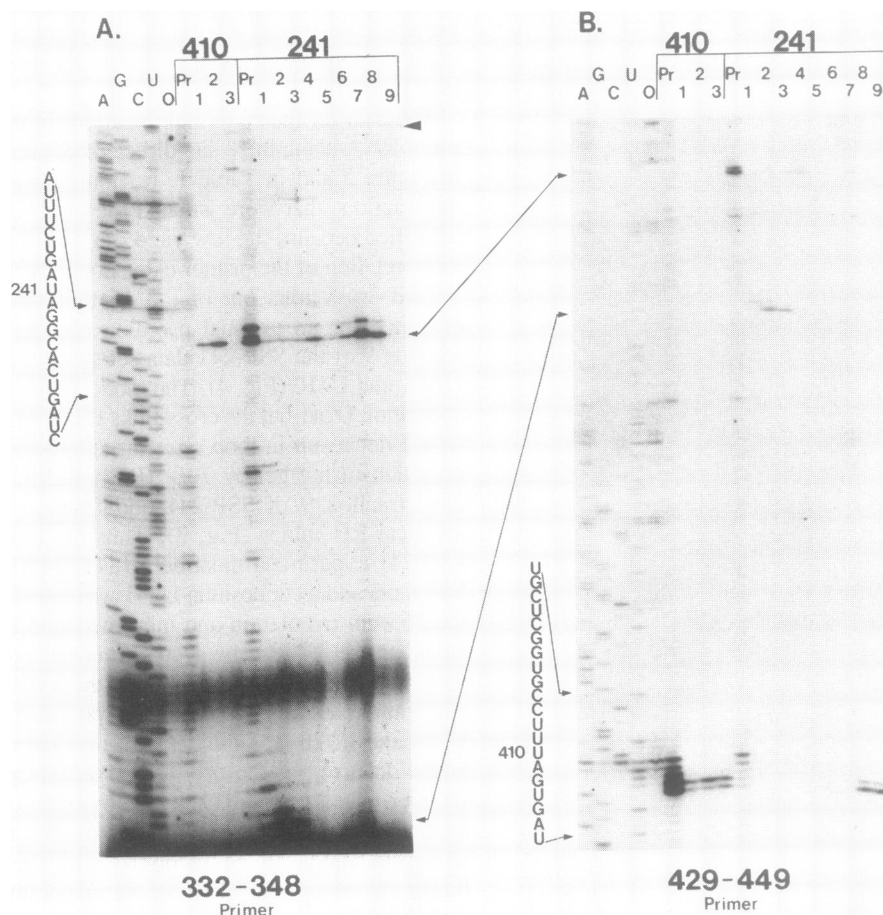


Figure 4. Primer extension analysis of SSP cross-linked RNA molecules. The RNAs in fractions Pr to 3 (SSP-410) and Pr to 9 (SSP-241) were used as templates for reverse transcriptase. The primers hybridized to nucleotides 332−348 (panel a) and to nucleotides 429−449 (panel b). Lanes A, G, C, and T contain the sequencing reactions and lane O contains extension products of control linear RNA. Nucleotide U241 is a strong stop in fractions 1 and 2 in the experiment in which SSP was first placed on U410. This indicates cross-linking between U410 and U241. Likewise, nucleotide U410 is a strong stop in fractions 7 and 8 in the experiment in which the SSP adduct was first placed on nucleotide U241. The arrow in panel A points to the reverse transcription stop at A185 (one nucleotide 3′ to U184) in fraction 9 of SSP-241 that resulted from a SSP transfer from U241 to U410 and subsequently cross-linked to U184. Lane 3 of SSP-410 also shows this stop at A185.

of cross-linked molecules that can be easily purified on single dimension denaturing gels and analyzed by primer extension experiments. The number of cross-linked products from each SSP experiment ranged from two to ten with an average of seven. At face value this would seem to indicate that any particular position investigated yielded at least seven different secondary structures. However, by considering the two major alternative secondary structures in the 3′ half of the molecule, a number of sites in which there are small rearrangements in the base-pairing pattern, and the number of higher order structure contacts, it is possible to account for more than 85% of the cross-links that we have determined.

The efficiency of re-cross-linking is dependent on two variables. First, the psoralen moiety of SSP requires the precise alignment of the 3,4 pyrone double bond with the 5,6 double bond of a pyrimidine on the opposite strand in order to form a cross-link (1). Such alignment will depend upon the local structure of the duplex formed with the target site and will vary with different sequences. The second variable is the possibility that alternative structures in the RNA could sequester SSP in a secondary structure without a cross-linkable pyrimidine on the opposite strand or could exclude the target sequence containing SSP from being involved in a secondary structure rendering it inaccessible for cross-linking. As can be seen in Figure 3 and Figure 5, most of the SSP adducts result in multiple cross-links indicating that the RNA can fold into multiple conformations. It is possible, although unlikely, that other conformations exist that would not allow SSP to cross-link.

The infrequency of SSP transfer (2 out of 46 identified cross-links) indicates that event is rare. It might be dependent on a variety of unknown factors such as specific sequence or structure of the RNA, stereochemistry of the adduct, equilibrium constants between alternative structures (see below), and solution conditions. The reaction may be one result of a general isomerization reaction reported by Tessman et al. (15). The authors noted that irradiating 8-methoxypsoralen (8-MOP) furan side monoadducts in calf thymus DNA with 342 nm light resulted in a larger number of pyrone side monoadducts than was expected. Their analysis, however, was the quantitation of the different photoproducts after complete digestion of the DNA, thus a transfer event that involved furan side adducts moving from one nucleotide to another could not have been detected.

All but one of the previously determined AMT cross-links (4) involving the target sites were verified in the present experiments; this indicates that SSP behaves similarly to AMT. The one exception was an AMT cross-link between U410 and C440 (4).

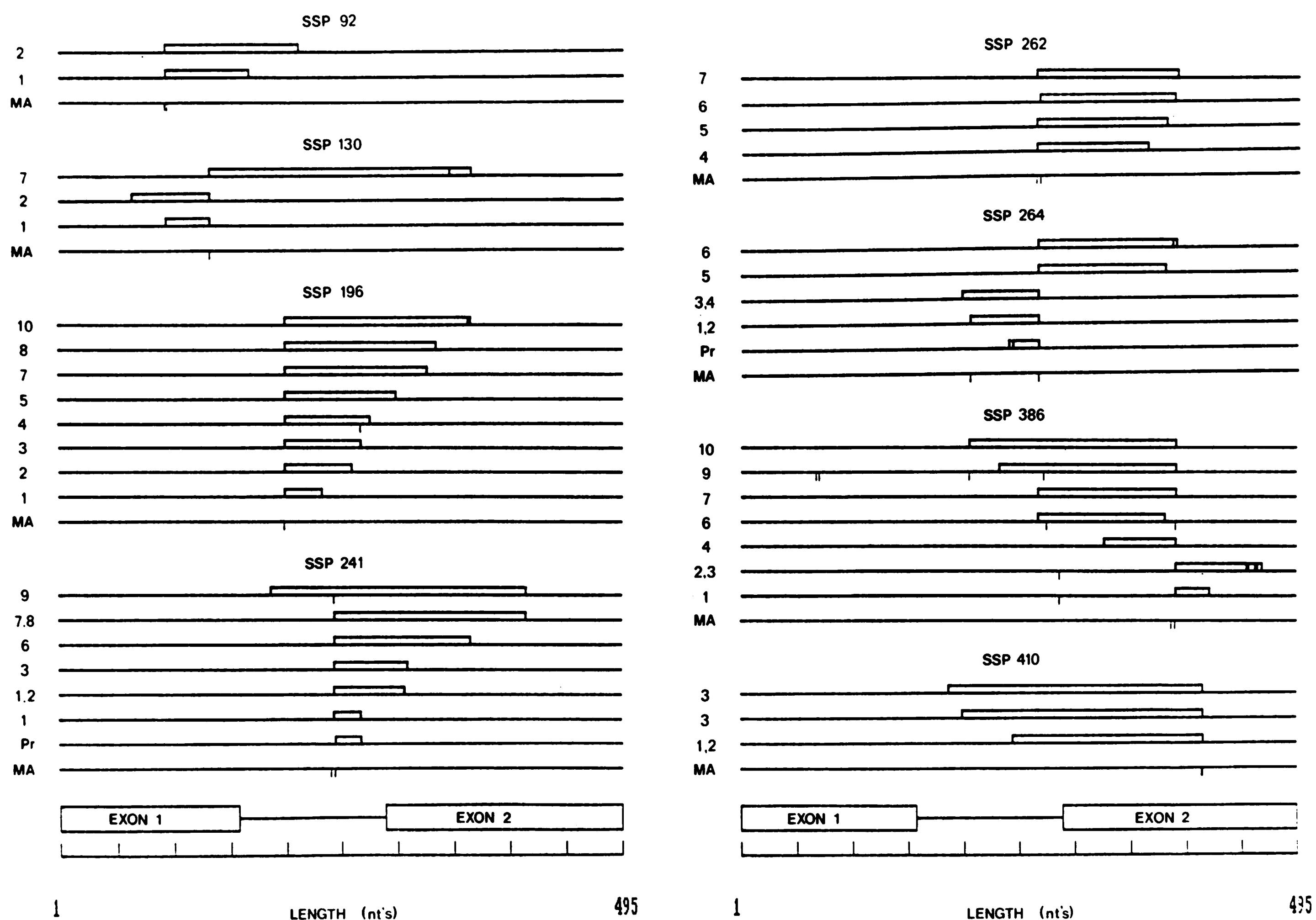
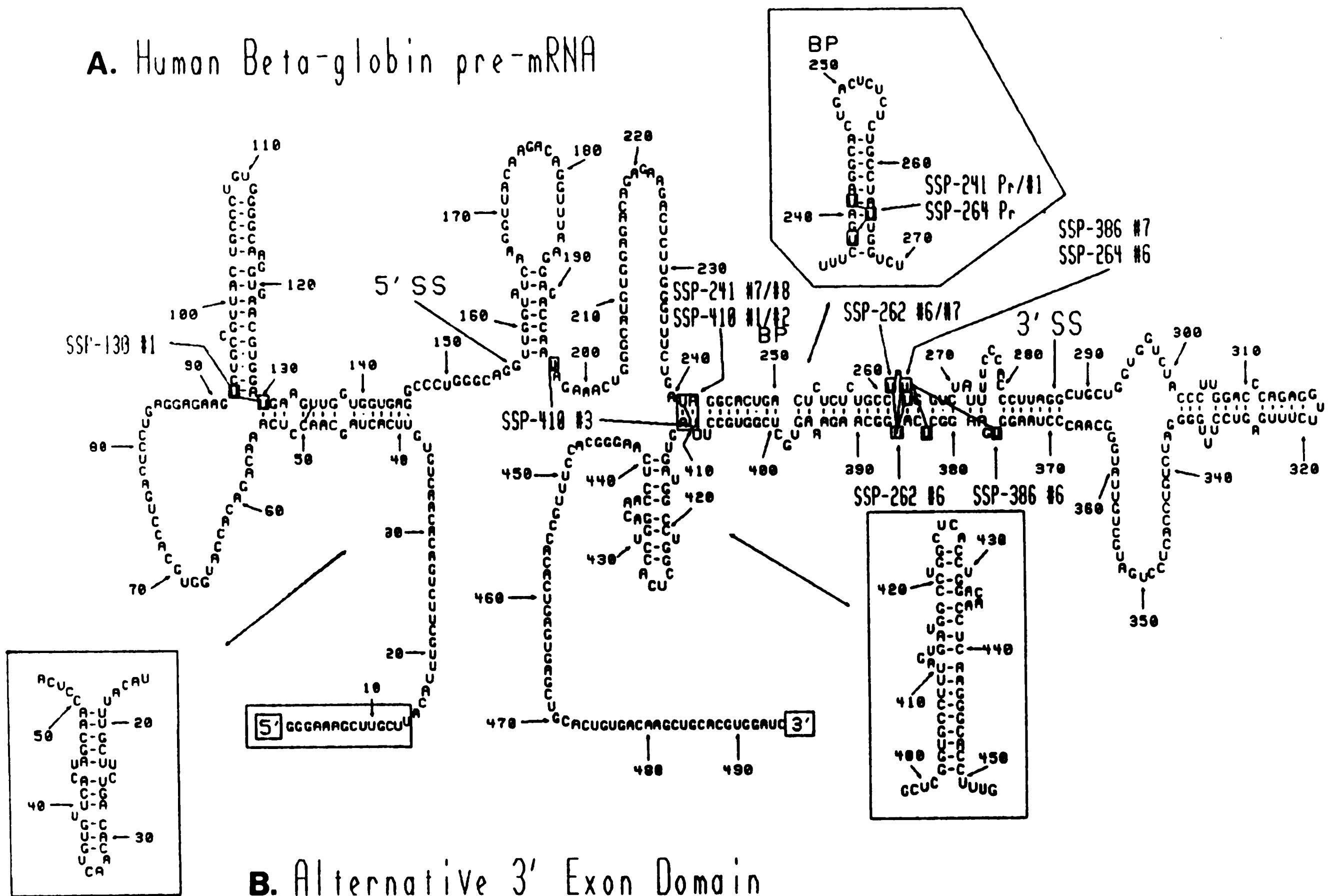


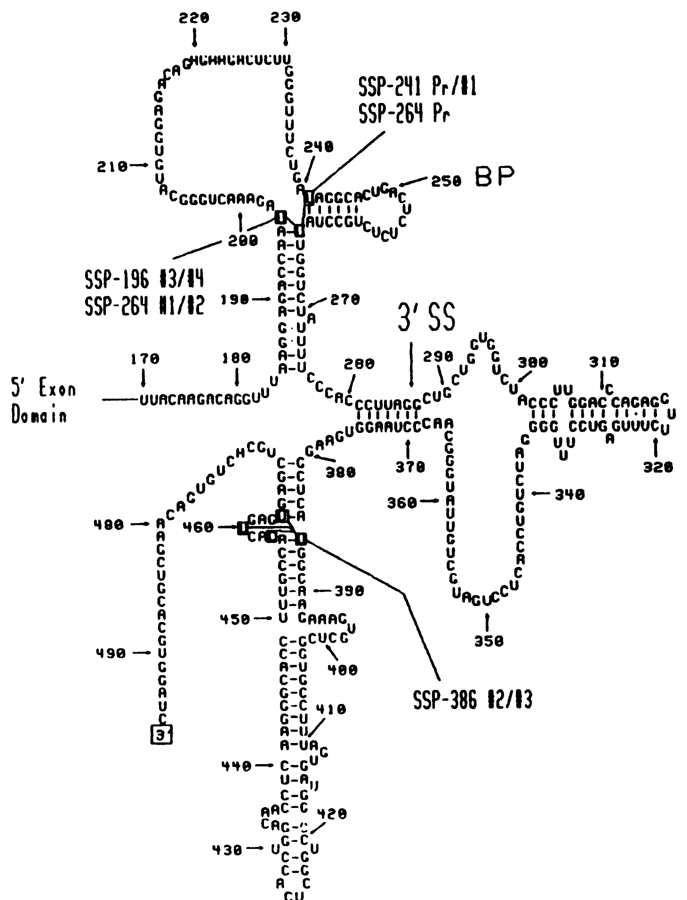
Figure 5. Summary of the locations of SSP cross-linking. The locations of the SSP monoadducts are shown as the short lines pointing down (MA). The cross-links are shown as bars connecting short lines pointing up. Some fractions contain more than one cross-link and others may have reverse transcription stops that have not been identified as cross-links (see text).

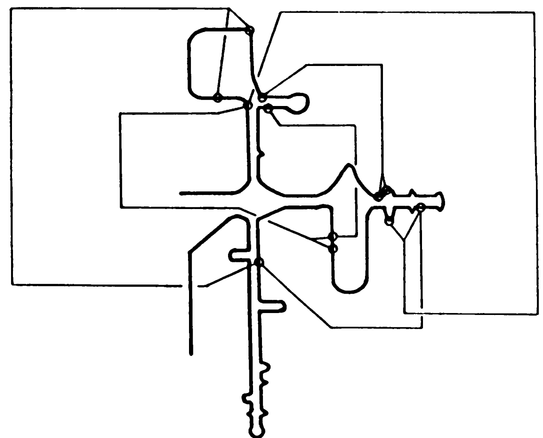
Table 1: Summary of the SSP Cross-linking Sites

SSP	Monoaddition Site	Cross-link Assignments 2°	Higher Order	Unassigned
92	92, 94	–	–	164, 209
130	130	63, 92	342, 361	–
196	196	264, 272, 295	323, 359, 361	230, 256
241	238, 241	264, 265, 410, *184–410	303, 306, 361	–
262	262	376, 384, 386	359	–
264	264	196, 203, 238, 241, 376, 383, 386	–	–
386	386	264, *264–376, 450, 452, 458, 460, 464	203, 230, 323	416
410	409, 410	184, 196, 241	–	–

* Cross-links that resulted from SSP phototransfer (see text)


A. Human Beta-globin pre-mRNA
BP
5′ SS
3′ SS
BP
SSP-130 #1
SSP-241 #7/#8
SSP-410 #1/#2
SSP-410 #3
SSP-262 #6/#7
SSP-241 Pr/#1
SSP-264 Pr
SSP-386 #7
SSP-264 #6
SSP-262 #6
SSP-386 #6
5′
3′


B. Alternative 3′ Exon Domain
SSP-241 Pr/#1
SSP-264 Pr
BP
SSP-196 #3/#4
SSP-264 #1/#2
3′ SS
5′ Exon Domain
SSP-386 #2/#3
3′


C.

The SSP monoadduct at U410 did not cross-link to C440. Since we have detected few SSP cross-links involving cytosines, it may be that SSP is less efficient than other psoralens in making adducts with cytosine. In one case, SSP was used to verify a small hairpin structure in the first exon in which we had not previously detected an AMT cross-link; this structure was found in fraction 1 of the SSP-130 data set, in which U92 was cross-linked to U130. However, a SSP monoadduct placed on the opposite side of this structure (SSP-92) failed to cross-link to the other side at U130. This asymmetry may be due to stereochemical constraints of the SSP adduct within this region of the secondary structure that allow only U130 furan side to U92 pyrone side cross-links but not the converse. Additional cross-links were identified that support secondary structures predicted on the bases of chemical probing data (4). These cross-links include U92×U130, U241×U265, U238×U264, and U241×U264.

A subset of cross-links were identified that are the result of small changes in the base pairing of the secondary structure shown in Figure 6, panel A. The cross-links U264×U376 and U264×U383 located in the duplexed region 258–287/369–390 in the 3′ exon domain are mutually exclusive to the cross-link U264×U386 located in the same region. These cross-links occur because of small changes in the base pairing pattern in this interval of the large hairpin. An equilibrium between these small alternative interactions is indicated by a phototransfer of SSP from U386 to U264 with subsequent cross-linking to U376 (see below). An additional duplexed region, 196–203/232–238, is suggested by the cross-link U196×U410. This base pairing interaction would bring U196 close to the inside end of the large, extended hairpin of the 3′ exon domain (Fig. 6, panel A) and, with a small change in the base pairing of 408–411/240–242, would allow cross-linking between U196 and U410.

Other SSP cross-links occurred at locations not previously identified as being involved in secondary structures. The alternative cross-links in the SSP-196, SSP-264, and SSP-386 data sets along with the base specific chemical probing data (4) were used to formulate a more extensive alternative secondary structure for the 3′ exon domain than was previously suspected. The new version of the alternate secondary structure includes a larger base-pairing interaction in the interval between 380 and 468 and a new interaction between the segments 185–195 and 264–275.

An interesting feature of the alternative structures is that the difference in the total number of base-pairs in the two main structures is only five. In addition, the structures may interconvert by a series of small changes in which there are compensations in the lengths of alternative duplex regions. The occurrence of alternative cross-links at one SSP location indicates that the alternative structures either exist in separate populations of molecules that renature along different pathways when re-introduced into native conditions or are in conformational equilibrium with each other. The very nature of the unusual light catalyzed SSP transfer reaction suggests that there is an equilibrium between some of these alternative structures. This equilibrium must be rapid enough at 4°C to allow transfer of SSP from the target site to a second site in one secondary structure and to allow a conformational switch in the RNA so that cross-linking can occur from the second site to a third site in an alternative structure during the 10 minute irradiation time.

There are some cross-links which neither correlate with the existing structural models nor provide convincing evidence for alternative secondary structures. Although it is possible that these cross-links represent weak secondary structure interactions that do not cause a strong chemical protection pattern, it is more likely that these cross-links represent higher order structure of the RNA. These cross-links include U241×C303, U241×U306, U241×U361, U196×U323, U196×U332, U262×U359, U203×U386, U230×U386, U323×U386, and U239×U386. Superimposition of these cross-links on the secondary structure model (Fig. 6, panel C.) reveals a pattern that suggests interactions between the secondary structure 'arms' in the 3′ exon domain. The folding of the arms on top of each other may involve a limited number of specific contacts, some of which could allow SSP cross-linking between secondary structures. We chose to call these higher order interactions rather than tertiary interactions because it is unlikely that they are as specific or contribute to the overall stability of the molecule in the same way as do the tertiary interactions in tRNA. Two other cross-links, U130×U342 and U130×U361 also suggest higher order interactions between the 5′ and the 3′ exon domains. Such cross-links would be expected to be more temperature labile than the cross-links that occur in secondary structures. For SSP-196, the cross-links to positions U230, U256, U264, and U272 are stable at 60°C in spliceosome assembly buffer while the cross-link to position U323 (fraction #7) is not (unpublished results).

The significance of these structures for pre-mRNA splicing or splice site selection has not yet been determined. Since SSP re-cross-linking will be efficient if the pre-mRNA is assembled into a RNP, the SSP technique will be able to determine the fate of the RNA structures during an *in vitro* splicing reaction. Because the location of SSP attachment can be controlled, it should be possible to find placement sites that will not interfere with splicing but will nevertheless provide needed structural information. The structural models presented in this paper will be a guide for these experiments.

ACKNOWLEDGEMENTS

These experiments were supported by an American Cancer Society Institutional Research Grant (IN-124). Dr. Tomasz Heyduk is thanked for helpful discussions on thiol chemistry and Gwen Ericson is thanked for helping revise the manuscript.

REFERENCES

1. Cimino, G.D., Gamper, H.B., Isaacs, S.T., and Hearst, J.E. (1985) Ann. Rev. Biochem. *54*, 1151–1193.
2. Wollenzien, P.L. (1988) In: Psoralen DNA Photobiology, Vol. 2, (ed. F.P. Gasparro), pp 51–66, CRC Press, Boca Raton, Fl.
3. Saffron, W.A., Goldenberg, M., and Cantor, C.R. (1982) Proc. Nat'l. Acad. Sci. USA *79*, 4594–4598.
4. Teare, J. and Wollenzien, P.L. (1989) Biochemistry *28*, 6208–6219.
5. Davanloo, P., Rosenberg, A.H., Dunn, J., and Studier, F. W. (1987) Proc.

Figure 6. Structural models for the human beta globin pre-mRNA. (A) The secondary structure model of the RNA, including alternative structures, proposed in ref. 4. is shown. Nucleotides cross-linked by SSP are connected and boxed. Other cross-links are shown that indicate local changes in the base pairing in the middle of the extended hairpin in the 3′ exon domain (SSP-386, fraction #6) or new alternate interactions that do not globally disrupt the secondary structure model (SSP-410, fraction #3). (B) An alternative 3′ exon domain (starting from nucleotide 170) proposed on the basis of alternative cross-links from SSP-196, SSP-241, SSP-264, and SSP-386 data sets and the previous chemical probing data (ref. 4). (C) The insert shows a line drawing of the alternative 3′ exon structure with the locations of higher order structure cross-links connected by lines. These cross-links are listed in the text.

Nat'l. Acad. Sci. USA *81*, 2035–2039.
6. Chu, B.C.F. and Orgel, L. (1988) Nuc. Acids Res. *16*, 3671–3691.
7. Teare, J. and Wollenzien, P.L. (1989) Nuc. Acids Res. *17*, 3359–3372.
8. Isaacs, S.T., Shen, C.J., Hearst, J.E., and Rapoport, H. (1977) Biochemistry *6*, 1058–1064.
9. Hearst, J.E., Isaacs, S.T., Kanne, D., Rapport, H., and Straub, K. (1984) Quart. Rev. Biophysics *17*, 1–44.
10. Ericson, G. and Wollenzien, P.L. (1988) Analyt. Biochem. *174*, 215–223.
11. Shi, Y.-B., Gamper, H., and Hearst, J.E. (1988) J. Biol. Chem. *263*, 527–534.
12. Shi, Y.-B., Gamper, H., and Hearst, J.E. (1987) Nuc. Acids Res. *15*, 6843–6854.
13. Shi, Y.-B., Gamper, H., Van Houten, B., and Hearst, J.E. (1988) J. Mol. Biol. *199*, 277–293.
14. Hall, K.B., Green, M.R., and Redfield, A.G. (1988) Proc. Nat'l. Acad. Sci. USA *85*, 704–708.
15. Tessman, J.W., Isaacs, S.T., and Hearst, J.E. (1985) Biochemistry *24*, 1669–1676.